University of Mohamed Boudiaf - M'sila

FACULTY OF TECHNOLOGY

DEPARTMENT OF ELECTRONICS

Serial Number ……..…

Registration Number……..……

Thesis

Presented for obtaining the diploma of

LMD DOCTORATE

Field: Electronics

Specialty: Micro-electronics

TITLE

Ab-initio and artificial intelligence based methods for materials physical properties prediction.

Presented by

BOUZATEUR Inas

Supported on: / /

In front of the following jury Committee:

Full Name			**Quality**
BOUCHAMA Idris	**Professor**	Univ - **M'sila**	Chairman
BENNACER Hamza	**Professor**	Univ - **M'sila**	Supervisor
OUALI Mohamed Assam	**MCA**	Univ - **M'sila**	Co- Supervisor
DJOUDI Lakhdar	**Professor**	Univ -**Tissemsilt**	Examiner
HAMDANI Mostefa	**MCA**	Univ -**Tissemsilt**	Examiner
HARHOUZ Ahlam	**MCA**	Univ - **M'sila**	Examiner
LADJAL Mohamed	**Professor**	Univ- **M'sila**	Guest

Academic Year : 2023/2024

Dedication

Not all letters can find the right words... Not all words can express gratitude, love, respect, and appreciation... Therefore, I dedicate this humble work to:

My dear parents

No dedication can express love, esteem, and respect. You have done more than parents could do to ensure that their children follow the right path in their lives and studies. This work is the result of your years-long sacrifices for my education.

My only brother, Salah El Dine, was the first companion in this research journey. Allah saves him.

I am grateful to **my adorable sisters** Aya, Israa, and Anfel for their unwavering support and love. They have been my soul mates.

My dearest friends, DEBAB Zineb, AISSAOUI Lena, HADJI Maroua and Her Sisters , and my cousin CHAMI Manar . They encouraged and supported me during all these research years. My deepest attachment and heartfelt gratitude to you.

To all my family members, **BOUZATEUR and GOURGACHE.**

Especially my **grandparents** and **my paternal aunt, Jamila.** May Allah have mercy on them.

I want to express my gratitude to everyone who holds a special place in my heart, even those I haven't mentioned yet.

Acknowledgements

"The teacher who walks in the shadow of the temple, among his followers, gives not of his wisdom but rather of his faith and his lovingness. If he is indeed wise he does not bid you enter the house of his wisdom, but rather leads you to the threshold of your own mind."

Khalil Gibran

Over all, we are grateful to Allah for giving us the strength, health, patience, and courage to complete this work and achieve this goal.

With these few lines, I would like to express my gratitude and deep appreciation to my thesis supervisor, **Prof. BENNACER Hamza**, who has supervised me during these four years and for his wise advice on the treatment of my thesis subject as well as his permanent availability. His help and competence have been of great help and support in the writing of this thesis.

I would like to express my sincere thanks to **Dr. OUALI Mohamed Assam**, my thesis co-supervisor, who taught me a great deal during my work. I appreciated his great scientific skills, his availability, and his kindness.

I would like to express my sincere thanks to the members of the jury for this thesis: **Prof. BOUCHAMA Idris**, who kindly agreed to chair the jury for this thesis, **Prof. DJOUDI Lakhdar, Dr. HAMDANI Mostefa,** and **Dr. HARHOUZ Ahlam** for agreeing to examine my research work.

Special thanks to **Prof. LADJAL Mohamed,** Prof**. BENAMARA Zineb**, and **Dr. HEDJAR Hakima** for their help, encouragement, and advices during my research work.

I'll never forget my dear parents' assistance and continued support in helping me reach my goal.

Abstract

Creating novel and inexpensive compounds that can meet both current and projected needs is becoming increasingly important due to the rapid speed of industrial change. However, traditional methods of rationally finding new materials with a specific set of features have grown challenging and costly because of the rise in material structural and functional complexity. This crucial fact has paved the way for the application of intelligence methods in this field. This dissertation focuses on two fundamental axes of research: The first focus on developing an intelligent prediction and identification technique for various material properties, and the second is on using the DFT theory to calculate and analyze the lattice parameter and band gap energy properties of the proposed compound, Ba_2BiTaS_6. To achieve this goal, we have addressed several points: First, we came up with a new model that uses artificial neural networks (ANN) and the particle swarm optimization algorithm (PSO) to get rid of problems with local minima in ANN models while keeping the quality of the fitting. By predicting the band gap properties, this method speeds up the search for new chalcopyrite materials in photovoltaic solar cells with better resolution. The model has two separate parts: an ANN sub-system that makes predictions using low-resolution training data; and an "error model" sub-system that was added to deal with resolution quality issues and show uncertainty in the primary model. Furthermore, we presented a comparative analysis of optimization algorithms to understand and quantify ANN's performance in guiding the search process towards better solutions over all feasible solutions. We used an efficient technique based on ANN-PSO and fuzzy logic-PSO to predict the lattice constants of pseudo-cubic and cubic perovskites. We predicted the lattice parameter of double perovskite compounds using the extreme learning machine (ELM). Finally, we used the FP-LAPW method in the WIEN2k environment, which is based on the DFT theory, to calculate and analyze the lattice parameter and the band gap energy properties of Ba_2BiTaS_6.

Résumé

La création des nouveaux composants peu coûteux et capables de répondre aux besoins actuels et futurs devient de plus en plus cruciale en raison de la rapidité des changements industriels. Cependant, les méthodes traditionnelles de recherche rationnelle de nouveaux matériaux avec un ensemble spécifique de caractéristiques sont devenues difficiles et coûteuses en raison de l'augmentation de la complexité structurelle et fonctionnelle des matériaux. Ce constat essentiel a ouvert la voie à l'idée d'utiliser des méthodes d'intelligence artificielle dans ce domaine. La recherche présentée dans cette thèse se concentre sur deux axes fondamentaux : Le premier se concentre sur développement d'une technique intelligente de prédiction et d'identification de plusieurs propriétés des matériaux et le second sur l'utilisation de la théorie de la fonctionnelle de la densité (DFT) pour calculer et analyser le paramètre de réseaux et la bande interdite du composé Ba_2BiTaS_6 proposé. À cette fin, plusieurs points ont été abordés. Premièrement, nous avons proposé une structure innovante exploitant les réseaux de neurones artificiels (ANN) et l'algorithme d'optimisation par essaim de particules (PSO) pour éliminer les problèmes de minima locaux dans les modèles ANN tout en maintenant la qualité de l'ajustement. Cette approche accélère l'identification de nouveaux matériaux de chalcopyrite dans les cellules solaires photovoltaïques avec une résolution améliorée en prédisant les propriétés de la bande interdite. Le modèle comprend deux sous-systèmes parallèles : un système de prédiction conventionnel avec des données d'entraînement de basse résolution comme premier sous-système ANN, et un second sous-système de "modèle d'erreur" introduit pour aborder les préoccupations de qualité de résolution et représenter l'incertitude dans le modèle principal. De plus, nous avons présenté une analyse comparative des algorithmes d'optimisation afin de comprendre et de quantifier la performance des ANN dans la conduite du processus de recherche vers de meilleures solutions parmi toutes les solutions possibles. Des techniques efficaces basées sur ANN-PSO et la logique floue-PSO ont été utilisées pour prédire les constantes de réseau des pérovskites pseudo-cubiques et cubiques. La machine d'apprentissage extrême (ELM) a été utilisée pour prédire le paramètre de réseau des composés à double pérovskite. Enfin, nous avons utilisé la méthode FP-LAPW dans l'environnement WIEN2k, basée sur la théorie DFT, pour calculer et analyser le paramètre de réseaux et la band énergie de Ba_2BiTaS_6.

ملخص

إن إنشاء مواد جديدة وبتكلفة منخفضة قادرة على تلبية الاحتياجات الحالية والمتوقعة أصبح أكثر أهمية نظرًا للسرعة الفائقة للتغيرات الصناعية. ومع ذلك، فإن الطرق التقليدية لاكتشاف مواد جديدة ذات خصائص محددة أصبحت صعبة ومكلفة بسبب الزيادة في التعقيد البنيوي والوظيفي للمواد. هذه الحقيقة الأساسية أفسحت المجال لفكرة استخدام أساليب الذكاء الاصطناعي في هذا المجال.تركز الأبحاث المقدمة في هذه الرسالة على محورين أساسيين: 1) تطوير تقنية ذكية للتنبؤ وتحديد العديد من خصائص المواد؛ 2) استخدام نظرية الكثافة الوظيفية (DFT) لحساب وتحليل الخصائص البنيوية والإلكترونية للمركب المقترح(Ba_2BiTaS_6) . لتحقيق هذا الهدف، تم تناول عدة نقاط. أولاً، اقترحنا بنية مبتكرة تستفيد من الشبكات العصبية الاصطناعية (ANN) وخوارزمية اسراب الطيور (PSO) للقضاء على مشاكل النقاط الأدنى المحلية في نماذج ANNمع الحفاظ على جودة الملاءمة. تُسرع هذه الطريقة من تحديد مواد الكالكوبيرايت الجديدة في الخلايا الشمسية الكهروضوئية مع تحسين الدقة عن طريق التنبؤ بخصائص فجوة الطاقة. يحتوي النموذج على نظامين فرعيين متوازيين: نظام تنبؤ تقليدي ببيانات تدريب منخفضة الدقة كنظام فرعي أولANN ، ونظام فرعي ثاني "نموذج خطأ" تم إدخاله لمعالجة مخاوف جودة الدقة وتمثيل عدم اليقين في النموذج الرئيسي.بالإضافة إلى ذلك، قدمنا تحليلًا مقارنًا لخوارزميات التحسين لفهم وقياس أداء ANN في توجيه عملية البحث نحو حلول أفضل من بين جميع الحلول الممكنة. تم استخدام تقنيات فعالة تعتمد على ANN-PSOوالمنطق الضبابي PSO-للتنبؤ بثوابت الشبكة للبيروفسكايت المكعب والبيروفسكايت الزائف. كما تم استخدام آلة التعلم المتطرفة (ELM) للتنبؤ بمعامل الشبكة لمركبات البيروفسكايت المزدوجة.أخيرًا، استخدمنا طريقة FP-LAPW في بيئة WIEN2k، المعتمدة على نظريةDFT ، لحساب وتحليل الخصائص البنيوية والإلكترونية لـBa_2BiTaS_6.

Table of Contents

CHAPTER III OVERVIEW OF DFT APPROACH

CHAPTER IV RESULTS AND DISCUSSION

List of Tables

List of Figures

List of Abbreviations

AI Artificial intelligence.

ML Machine Learning.

DL Deep Learning.

ANN Artificial Neural Networks.

FFNN Feed-Forward Neural networks.

RNN Recurrent neural networks.

CNN Conventional Neural Network.

FIS Fuzzy Inference System.

TSK Takagi-Sugeno-Kang.

GD gradient descent.

EAs Evolutionary Algorithms.

PSO Particle Swarm Optimization.

ICA Imperialist Competitive Algorithm.

IWO Invasive Weed Optimization Algorithm.

DFT Density Functional Theory.

LDA Local Density Approximation.

LSDA+U Local Spin Density Approximation.

GGA Generalized Gradient Approximation).

MBJ Beck-Johnson modified (modified Beck-Johnson).

FP-LAPW Linearized augmented plane waves at full potential.

SCF Self-Consistent Cycle.

ZB Brillouin Zone.

LC Lattice Constant .

Eg Energy band gap .

GENERAL INTRODUCTION

Introduction

Designing and discovering new materials and structures is a challenging task. Hence, advancing new approaches to replacing materials that are often costly, harmful to health and environment, but remain vital to the development of modern technologies, is necessary. The field of material science and engineering has expanded significantly with the help of the periodic table elements in forming binary, ternary, or quaternary compounds with unique properties. It is worth noting, however, that according to natural law, combining two distinct elements does not result in combining their properties but rather constructs new physical properties [1].

In today's fast-paced world of technological advancements, it is increasingly important to design new and cost-effective compounds that can meet both present and future demands. For an extended period, experimental procedures have been the foundation of materials research, providing direct observation and measurement of the Physico-chemical characteristics of materials. For instance, the surface morphology may be seen at high resolution using scanning electron microscopy (SEM), whereas X-ray diffraction (XRD) can be used for understanding crystal structure and phase identification. Nuclear magnetic resonance (NMR) spectroscopy provides important insights into the dynamics and structure of molecules [2]. However, the most substantial difficulty to reaching the latter remains the issue of time. The arduous and time-consuming process of creating and testing many candidates, when only a few are actually required, has reached an unsustainable point in terms of time and cost [3, 4]. The integration of computers and digital computing has brought about a significant transformation to the vision of researchers regarding the issue. It has especially enabled the formulation of physical problems in mathematical equations, which are designed to cater to specific requirements [5].On these grounds, numerical calculation and simulation has been of paramount importance throughout recent decades. Large-scale simulations are among the most demanding requirements for technological development in a variety of fields. Accordingly, a sub-discipline of theoretical physics, based on numerical or computational physics, has emerged. The field focuses on implementing numerical algorithms to solve physical problems for which a theory already exists [6]. It is sometimes regarded as an intermediate branch between theoretical and experimental physics.

Due to the constant advancement in computer technology, numerical modeling techniques have evolved rapidly over the past few decades. These modeling techniques combine calculations

of various types, operating at different scales that range from atomic dimensions to the macroscopic scale of matter. For instance, the study of various properties of solid materials, such as stability, structural, electronic, electrical, magnetic, optical, and mechanical properties, has become possible with the help of treatments based on the quantum mechanics of physics. In particular, "ab-initio", first-principles quantum simulations have significantly contributed to this progress. By utilizing the density functional theory (DFT) formalism and the linearized augmented plane wave (LAPW) method [7], these simulations enable the determination of many physicochemical properties of matter under extreme conditions that are often inaccessible in experimental studies.

Over the past 50 years, functional density theory (DFT) has gained significant popularity. It has been most widely used for theoretical calculations of the physical properties of condensed matter [8]. Particularly, DFT uses its density to simulate each interacting system of intrinsic fermions rather than its many-body wave function to determine the ground state properties of a specific substance [9, 10]. This has enabled the prediction for a range of molecular properties, including vibrational frequencies, molecular structures, atomization and ionization energies, electric and magnetic characteristics, and reaction pathways [11]. However, the computational cost of such approaches is high due to the complex structure of the connections between particles. This has limited its application in most cases to a smaller set of materials and has made large-scale exploration of materials relatively impractical [12]. Thus, using Artificial intelligence (AI) methods in relevant contexts has emerged to solve this issue. These models use a variety of techniques to accurately predict a range of materials properties. Computational algorithms for machine learning (ML) provide advantages over other materials computational techniques for material properties [13]. This attracts researchers to rely on it to accelerate the discovery of new materials with higher performance [14]. This represents the main challenge that this thesis addresses. Mainly, it builds on this idea and proposes a novel approach that uses Artificial intelligence and machine learning approaches in material science applications. Our approach relies on machine learning for the specific purpose of accelerating data-driven material property discovery, guided by the expertise of domain experts.

In recent years, perovskites and chalcopyrite have gained unprecedented attention due to their remarkable physicochemical properties. Becoming the subject of the most engaging research topics. They are used for multiple purposes and are mostly regarded as the fundamental compounds of modern optoelectronics [15-16]. Particularly, they received a considerable

attention in thin films and solar cells' applications [17], light-emitting diodes [18], laser cooling [19]. Each specific application requires the identification and production of materials with the most suitable properties.

The band gap (Eg) and lattice constant (LC) are considered among the most important features that studies for specific purposes rely on. The band gap's energy is a critical property for photovoltaic semiconductors. Fabricating solar cells often requires the use of materials with band gap energy between 1.1 and 1.8 eV as solar absorbers [20]. Theoretical research on such compounds introduced a novel approach to accelerating the study and prediction of the band gap energy for novel ternary semiconductors [21–24]. Using machine-learning methods to develop effective computational tools to handle this specific challenge is a new but significant study area.

Many researchers have begun to investigate potential theories that may explain how the band gap and the basic chemico-physical properties of these chalcopyrite's compounds relate to one another. In 2001, Zeng et al [21].'s pilot research established the basis for our investigation. Using artificial neural networks, these scientists estimated the relationship between the band gap energy (and lattice constant) of chalcopyrite and their corresponding chemical properties. They demonstrated that the dependency could in fact be represented linearly, which encouraged the application of linear regression techniques in subsequent studies. Suh and Rajan (2004) [22] used PLS regression to estimate the underlying linear model using the same descriptors as in the Zeng et al. study. In a later work, Dey et al. (2014) [23] combined feature ranking with numerous prediction algorithms (OLS, LASSO, and SPLS) to create a more reliable quantitative model. An innovative approach based on feature selection and regression approaches was recently suggested by Khmaissia et al. (2018) [24] to predict the band gap energy of 156 chalcopyrite.

On the other hand, the lattice constant (LC) is a crucial parameter of crystalline materials that assists in identifying the material and disclosing its structural properties [25]. For instance, in thin film growth, a mismatch between both lattice parameters of the substrate and the film is necessary to achieve high film quality [26–29]. Machine learning methods have become a viable method for properly predicting lattice constants. These models are able to quickly and effectively capture the nonlinear interactions involved [30]. Predictive models that are capable of precisely estimating these lattice parameters are therefore in great demand. Systems to establish a relationship between the lattice constant and chemical properties have been presented by a number of researchers [31-37].

In 2007 [31], Jiang et al. suggested a model that makes use of empirical constants that are then optimized and 6-fold Shannon effective ionic radii. They included Goldschmidt's tolerance factor for perovskite structures [32]. Because the A ions in the initial perovskite structure had a coordination number of 12, Moreira and Dias [33] argued that the Goldschmidt tolerance factor could not be calculated using ionic radii with a coordination number of 6. Nevertheless, in comparison to the Jiang et al. model, their approach produced a higher error rate. An empirical model based on the 6-fold effective ionic radii was presented by Ubic [34] to estimate the lattice parameters of cubic and pseudo-cubic ABC3 perovskites. An improved mean squared error of 0.0010 Å is the desired result for this model. In 2019, Sidey [35] created a model based on the modified data from the Ubic [34] and Jiang et al. models using multiple linear regressions. Their model was tested based on the mean squared error, and it was determined to have a prediction accuracy of 0.0011 Å. In 2020, Owolabi [36] developed a hybrid model that combined the support vector regression algorithm (SVRA) with particle swarm optimization (PSO) technology. Their models, with a mean squared error (MSE) of 4.424×10^{-4}Å, performed better than those of other investigations. Sandra et al. had predicted the lattice constant of A_2BCO_6 double perovskite using the MLR and ANN methods [37].

The research work that will be presented in this thesis essentially focuses on two axes.

- ✓ Predicting the band gap and lattice parameter properties for the ABC_2 chalcopyrite materials and the simple and double perovskites types, respectively using different computational intelligence methods.
- ✓ Predicting the band gap energy and lattice constant for the Ba_2BiTaS_6 compound using the DFT approach.

In this thesis, our initial objective is to accelerate the discovery of material properties with lower computational costs by proposing appropriate new computational intelligence -based models with improved accuracy in predicting the properties of different compounds. Furthermore, we aim to optimize the proposed machine learning methods for more robust and accurate results by applying a new architecture and various optimization algorithms. Particularly, we intend to accelerate the discovery of the band gap for the chalcopyrite with the I-III-VI_2 and II-IV-V_2 formulas and the lattice parameters of the simple and double perovskites with ABX_3 and A_2BBO_6 formula, respectively, and boost their prediction accuracy. Moreover, we will employ all the properties of the atoms in the compounds under consideration to elucidate the importance of

the last element. After that, we will use the DFT method in the Wien2k environment, which is based on the FP-LAPW method, to look into the lattice constant and the energy band gap of the Ba_2BiTaS_6 double perovskite compounds.

Following this general introduction, which provides a brief presentation of the relevant areas of our research, this thesis manuscript is organized as follows. The first chapter will be devoted to a thorough overview of the chalcopyrite and perovskite material families. It basically includes their basic definitions, physic-chemical properties, and domain applications. It will also introduce the new perspective on material innovation.

The second chapter will present the theoretical underpinnings and fundamental principles of artificial intelligence and machine learning. It will introduce the artificial neural network method, including its process and various types. Additionally, we'll explore the core concepts of type-1 fuzzy logic, including type-1 fuzzy sets, their representations, set operations, and type-1 fuzzy systems. Furthermore, it will adequately outline how the analysis of optimization algorithms is performed; particularly the metaheuristic algorithms: PSO, ICA, and IWO.

The third chapter will describe the various methods used in this work. In particular, it will describe the DFT-based methods, the different approximations used for the exchange-correlation energy, the FP-LAPW method, in addition to the Wien2k calculation code used to simulate the lattice constant and the band gap energy of the proposed material.

The fourth chapter will consolidate the main obtained results and their interpretation. Finally, a comprehensive conclusion will summarize the key outcomes and outline the prospects for future research directions.

References

[1] Z.Zarouali & Y. Nougal, (2021). Etude des propriété structurale électronique et magnitique de demi-heusler licab pour application en spintronique.University of Tissemsilt .

[2] N.Alharbi, B.Varela & R. Hailstone, (2020). Alkali-activated slag characterization by scanning electron microscopy, X-ray microanalysis and nuclear magnetic resonance spectroscopy. Materials Characterization, 168, 110504.

[3] M. E.Eberhart & D. P. Clougherty(2004). Looking for design in materials design. Nature Materials, 3(10), 659–661.

[4] M.Jansen & J. C. Schön, (2006). "Design" in Chemical Synthesis—An Illusion? Angewandte Chemie International Edition, 45(21), 3406–3412. Portico.

[5] Philippe Depondt ,(1998) PHYSIQUE NUMERIQUE. Le calcul numérique sur ordinateur au service de la physique : une introduction.

[6] Bouafia Hamza, (2014). Etude théorique des propriétés structurales, mécaniques et électroniques des matériaux pérovskites de type KAX3 et NaAX3 (X= F, O) . Doctoral dissertation, University of mohamed boudiaf of Oran.

[**7]** M. R. DJELTI, (2017). Etude des Propriétés Spintroniques du Double Pérovskite Type $ABCO_6$ .Doctoral dissertation, University of Mostaganem.

[8] R. O. Jones, (2015). Density functional theory: Its origins, rise to prominence, and future. Reviews of Modern Physics, 87(3), 897–923.

[9] R. G.Parr & Y. Weitao, (1995). Density-Functional Theory of Atoms and Molecules. Oxford University Press.

[10] P.Hohenberg & W. Kohn, (1964). Inhomogeneous Electron Gas. Physical Review, 136(3B),B864–B871.

[11] W.Kohn & L. J. Sham, (1965). Self-Consistent Equations Including Exchange and Correlation Effects. Physical Review, 140(4A), A1133–A1138.

[12] P. G. Schultz, (2003). Commentary on combinatorial chemistry. Applied Catalysis A: General, 254(1), 3–4.

[13] A.Majid, A. Khan & T.-S. Choi, (2011). Predicting lattice constant of complex cubic perovskites using computational intelligence. Computational Materials Science, 50(6), 1879–1888.

[14] A. B. Maksov, (2018). Artificial Intelligence in Materials Science: Applications of Machine Learning to Extraction of Physically Meaningful Information from Atomic Resolution Microscopy Imaging.Doctoral Dissertation . University of Tennessee, Knoxville.

[15] P.Docampo & T. Bein, (2016). A Long-Term View on Perovskite Optoelectronics. Accounts of Chemical Research, 49(2), 339–346.

[16] M, Hadjab, M.Ibrir, S.Berrah et al (2018). Structural, electronic and optical properties for chalcopyrite semiconducting materials: ab-initio computational study. Optik, 169, 69–76.

[17] A. R.Uhl, Z. Yang, A. K.-Y.Jen & H. W. Hillhouse, (2017). Solution-processed chalcopyrite–perovskite tandem solar cells in bandgap-matched two- and four-terminal architectures. Journal of Materials Chemistry A, 5(7), 3214–3220.

[18] G.Motomura, K.Ogura, T. Kameyama et al ,(2020). Efficient quantum-dot light-emitting diodes using ZnS–AgInS2 solid-solution quantum dots in combination with organic charge-transport materials. Applied Physics Letters, 116(9). https://doi.org/10.1063/1.5143618

[19] SCHULTZ Christof, (2021) Laser patterning of chalcopyrite and perovskite based solar cells: Investigation of the laser-material interaction and laser-induced damages.Doctoral thesis.University of Berlin .

[20] I.Bouzateur, H. Bennacer, M. A.Ouali & al, (2023). A new ANN-PSO framework to chalcopyrite's energy band gaps prediction. Materials Today Communications, 34, 105311.

[21] Y.Zeng, S. J. Chua & P. Wu, (2002). On the prediction of ternary semiconductor properties by artificial intelligence methods. Chemistry of materials, 14(7), 2989-2998,

[22] C.Suh, A.Rajagopalan, X. Li & Rajan, K. (2003). Combinatorial materials design through database science. MRS Online Proceedings Library (OPL), 804.

[23] P.Dey, J.Bible, S.Datta et al . (2014). Informatics-aided bandgap engineering for solar materials. Computational Materials Science, 83, 185-195,

[24] F.Khmaissia, H. Frigui, M .Sunkara, & Menon, M. (2018). Accelerating band gap prediction for solar materials using feature selection and regression techniques. Computational Materials Science, 147, 304-315,

[25] I.Bouzateur, H.Bennacer, M. A.Ouali & al. (2023). A new ANN-PSO framework to chalcopyrite's energy band gaps prediction. Materials Today Communications, 34, 105311.

[26] I. O.Alade, I. A.Olumegbon & A.Bagudu, (2020). Lattice constant prediction of A2XY6 cubic crystals (A = K, Cs, Rb, TI; X = tetravalent cation; Y = F, Cl, Br, I) using computational intelligence approach. Journal of Applied Physics, 127(1).

[27] M. I.Kotelyanskii,I.M. Kotelyanskii, & V. B. Kravchenko, (2000). New buffer sublayers for heteroepitaxial III–V nitride films on sapphire substrates. Technical Physics Letters, 26(2), 163–164.

[28] J. Y.Guo, Y. W. Zhang& C. Lu, (2008). Effects of wetting and misfit strain on the pattern formation of heteroepitaxially grown thin films. Computational Materials Science, 44(1), 174–179.

[29] A.Khan & S. G. Javed,(2008). Predicting regularities in lattice constants of GdFeO3-type perovskites. Acta Crystallographica Section B Structural Science, 64(1), 120–122.

[30] R. L.Moreira & A. Dias, (2007). Comment on "Prediction of lattice constant in cubic perovskites." Journal of Physics and Chemistry of Solids, 68(8), 1617–1622.

[31] L. Q. Jiang, J. K.Guo, H. B.Liu & al. (2006). Prediction of lattice constant in cubic perovskites. Journal of Physics and Chemistry of Solids, 67(7), 1531–1536.

[32] T.Sato, S.Takagi, S. Deledda, B. & al. (2016). Extending the applicability of the Goldschmidt tolerance factor to arbitrary ionic compounds. Scientific Reports, 6(1).

[33] R. L.Moreira & A. Dias, (2007). Comment on "Prediction of lattice constant in cubic perovskites." Journal of Physics and Chemistry of Solids, 68(8), 1617–1622.

[34] R. Ubic, (2007). Revised Method for the Prediction of Lattice Constants in Cubic and Pseudocubic Perovskites. Journal of the American Ceramic Society, 90(10), 3326–3330. Portico.

[35] V. Sidey, (2019). A simplified empirical model for predicting the lattice parameters of the cubic/pseudocubic perovskites. Journal of Solid State Chemistry, 279, 120951.

[36] T. O. Owolabi, (2020). Extreme learning machine and swarm-based support vector regression methods for predicting crystal lattice parameters of pseudo-cubic/cubic perovskites. Journal of Applied Physics, 127(24).

[37] S.Dimitrovska, S.Aleksovska & I. Kuzmanovski, (2005). Prediction of the unit cell edge length of cubic A2BB′O6 perovskites by multiple linear regression and artificial neural networks. Open Chemistry, 3(1), 198–215.

CHAPTER I

THEORETICAL BACKGROUND ON CHALCOPYRITE AND PEROVESKITE

I.1. Introduction

Material physics and materials science are very important parts of the technological development that spans various industries, including the electronic and optoelectronics fields [1]. The emergence of a wide range of technological applications in industry necessitates the use of materials with specific properties that can actively enhance device operation through their unique behavior. As a result, interest in material research has grown in recent years. Finding new materials by combining two or more element in periodic table, which often results in a new substance with unique or improved qualities, is crucial to the very existence of many technologies and the ongoing evolution of their performance. The current decade has witnessed extensive research on numerous significant materials. Recent studies have strongly identified ternary compounds as the most common classes of materials with distinct properties and structures [2, 3].

Research on ternary chalcopyrite with II-IV-V_2 and I-III-VI_2 structures has become one of the most interesting compounds. They have become a potential material for broad technological applications (solar cells, light-emitting diodes, sensors, etc.) due to a variety of their chemical and physical properties, such as poor symmetry, high stability, a high optical coefficient, and an excellent band gap [4]. For instance, the most promising material in this group for use as an absorber in photovoltaic conversion is zinc-silicon arsenide ($ZnSiAs_2$) [5]. It is a ternary compound with band gap energy of 1.8 eV and a fairly high absorption coefficient. On the other hand, perovskite-structured compounds with the ABC_3 general formula are also a large family of compounds that make up one of the most important families of used materials in the optoelectronics industry [6]. This family comprises a wide variety of space-group subclasses. The optoelectronics field highly values the unique optical and electrical properties of the cubic structure; including carrier diffusion length, low trap-state density, high absorption coefficient, and strong photoluminescence [7, 8]. For example, halide perovskites have attracted a lot of interest recently for a variety of uses, such as laser cooling, water splitting, solar cells, and light-emitting diodes [9–12]. Researchers currently believe that chalcopyrite and perovskites-based thin-film solar cells will represent the next advancement in photovoltaic cell technology [6, 13]. As a result, these two material families are receiving a lot of interest from researchers in several fields (as shown in figures I.1.a and I.1.b).

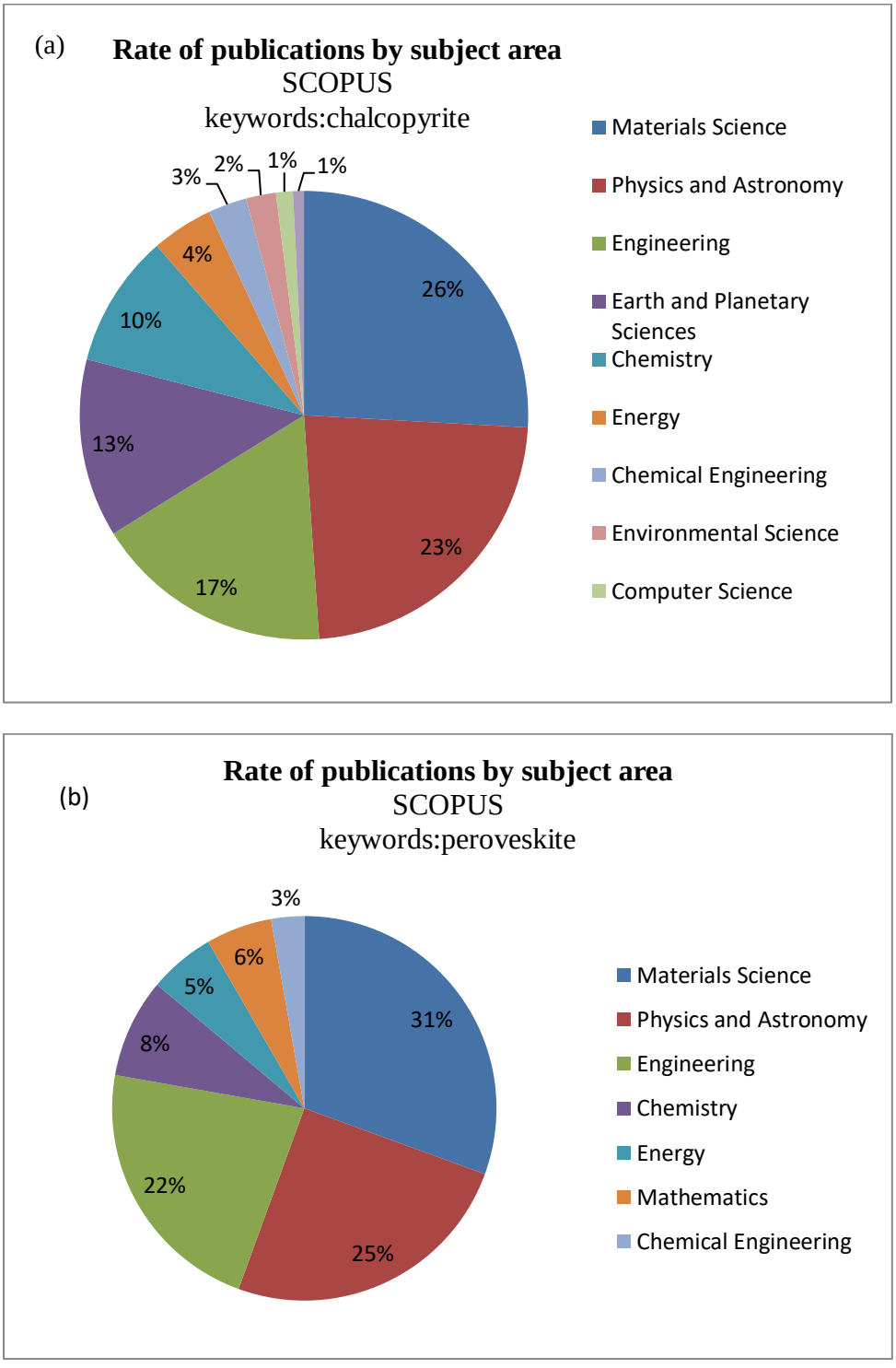

Figure.I.1 Number of publications by year on chalcopyrite (Fgure.I.1.a) and perovskite (Figure.I.1.b) by subject area. **Keyword** search "Chalcopyrite solar cell" in Title, Abstract and Keywords. Source **Scopus**.

In this chapter, we presented a thorough overview of the chalcopyrite and perovskite material families, including their basic definition, structural methodology, physical-chemical properties, and domain applications. Furthermore, we introduce a new perspective on material innovation.

I.2. Materials with chalcopyrite structure

I.2.1. Overview on the chalcopyrite compounds

A logical progression in the hunt for novel semiconductors was to look at ternary compounds that display tetrahedral coordination, which is an analogue of the diamond structure. In 1953, Hahn et al. [14] created compounds I–III–VI_2, about which Goodman and Douglas (1954) [15] talked regarding the prospect of semi conductivity. In 1957, Goodman [16] demonstrated that one can create new type II-IV-V_2 semiconducting compounds by substituting group III atoms in III-V compounds with ordered group II and IV atoms [17]. These I-III-VI_2 and II-IV-V_2 compounds were later referred to as chalcopyrite SCs [18]. Since then, chalcopyrite has garnered significant interest in various applications, particularly in optoelectronic devices such as solar cells [18], as illustrated in Figure (I.2).

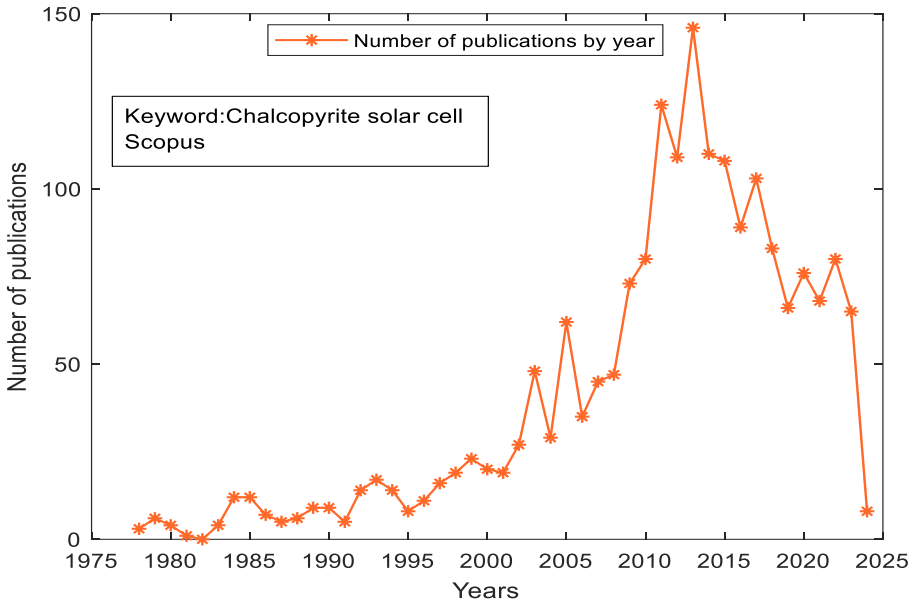

Figure.I.2 Number of publications by year on solar cell based on chalcopyrite. **Keyword** search "Chalcopyrite solar cell" in Title, Abstract and Keywords. Source **Scopus**.

As a result, there has been a steady increase in research on chalcopyrite materials with better performance and best property efficiency (Figure I.3). Shay and Wernick published the first paper on compounds with chalcopyrite structures in their book [19] at the beginning of 1970. Kazmerski et al. prepared the first thin film based on CuInS2 with the two conductivity types (P and N) in 1975 [20]. After that, research has focused a lot more on studying the properties of chalcopyrite, particularly due to their huge potential in photovoltaic applications. In the early years of 1980, Jaffe and Zunger were the first researchers to use density functional theory (DFT) to calculate the electronic structures and optical properties of many ternary chalcopyrite Cu-III-VI_2 materials. Their findings laid the groundwork for many more studies that have looked at how SC chalcopyrite can be used as absorbent thin films in solar cells [21–23]. In 2002, Zeng et al. used a new computational approach based on artificial intelligence to estimate the correlation between the electronic and structural properties of the ternary chalcopyrite and their chemical elementary features [24]. Today, research on this group of materials is still growing.

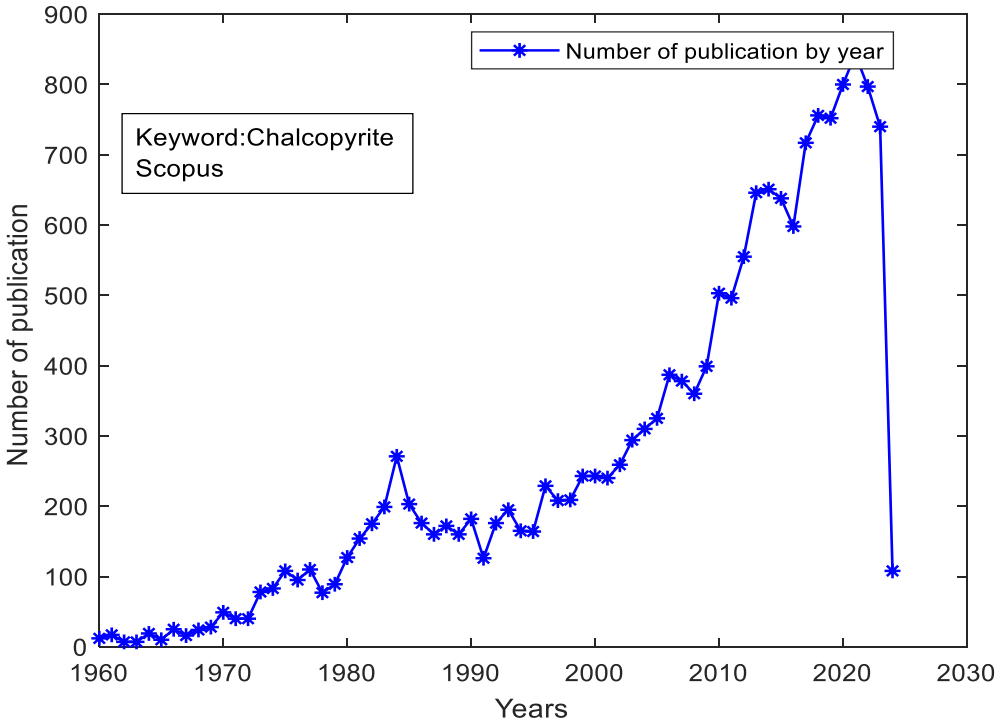

Figure.I.3 Number of publications by year on chalcopyrite. **Keyword** search "Chalcopyrite" in Title, Abstract and Keywords. Source **Scopus.**

Chalcopyrite is a mineral composed of copper iron sulfide with the chemical formula $CuFeS_2$, consisting of 34.5% copper, 30.5% iron, and 35% double sulfide, as reported by Henckel in 1725. It crystallizes as a tetragonal structure with cell dimensions of a = 5.289 Å and c = 10.423 Å and space group I2d [18, 25]. Figure I.4 illustrates the crystal structure of $CuFeS_2$.

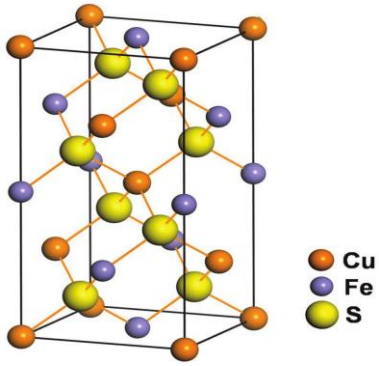

Figure.I.4. Crystal structure of $CuFeS_2$ [26].

Burdick and Ellis first reported the crystal structure of $CuFeS_2$ in 1917 [27]. In 1932, Pauling and Brockway provided the first accurate description of the $CuFeS_2$ structure by X-ray diffraction [28]. Kambara [29] conducted the first theoretical study on the crystal structure of $CuFeS_2$ in 1974, employing the Wolfsburg-Helmholtz method [30], a semi-empirical orbital molecular method. However, computational limitations restricted the system under study to a small final group of only 17 atoms, discouraging further study of the crystal structure [18].

I.2.2. Description of the crystal structure of chalcopyrite materials

The chalcopyrite structures (I-III-VI_2) and (II-IV-V_2) with the I2d space group (8 atoms via primitive unit cell) are still the subject of a lot of research. Where are natural extensions of the binary compounds with zinc blend F43m structure types $A^{II}B^{VI}$ and $A^{III}B^{V}$(2 atoms via unit cell), respectively. This relationship can be represented by a cationic cross-substitution diagram discovered in 1957 by Goodman [16]. According to Figure I.5, a chalcopyrite structure typically consists of two zinc-blend unit cells, which together form its tetragonal unit cell. Each unit cell has four group I/II atoms, four group III/IV atoms, and eight group VI/V atoms.

Comparing the ternary chalcopyrite to its binary equivalents, numerous intriguing structural anomalies exist [31]. Firstly, as illustrated in Figure I.5, we can obtain the chalcopyrite structure

by dividing the cubic cell of the zinc blende crystal in half along the z-axis. This is the c-axis of the chalcopyrite structure. Moreover, the zinc blende structures consist of a single cation, whereas ternary chalcopyrite has two cations with different chemical properties. If A and B are two distinct cations, the vertical direction over intervals of c2 yields the sequence ABABAB. On the other hand, if we travel horizontally at intervals of a, we find the sequence AAA... [22, 31]. Therefore, the arrangement of the two cations in the chalcopyrite crystal is periodic rather than random. Chalcopyrite, a tetragonal structure with coordination 4, connects every atom to four neighbouring atoms, forming a roughly regular tetrahedron. In order to produce the previously indicated tetrahedron, each anion (C) is associated with two cations (A) and two cations (B), and each cation is connected to four anions [18]. However, in zinc-blende binary (AC) compounds, every cation A has four nearest neighbors, anions C, and vice versa. In the space group (I2d), table I.1 illustrates the atomic positions of the chalcopyrite structure.

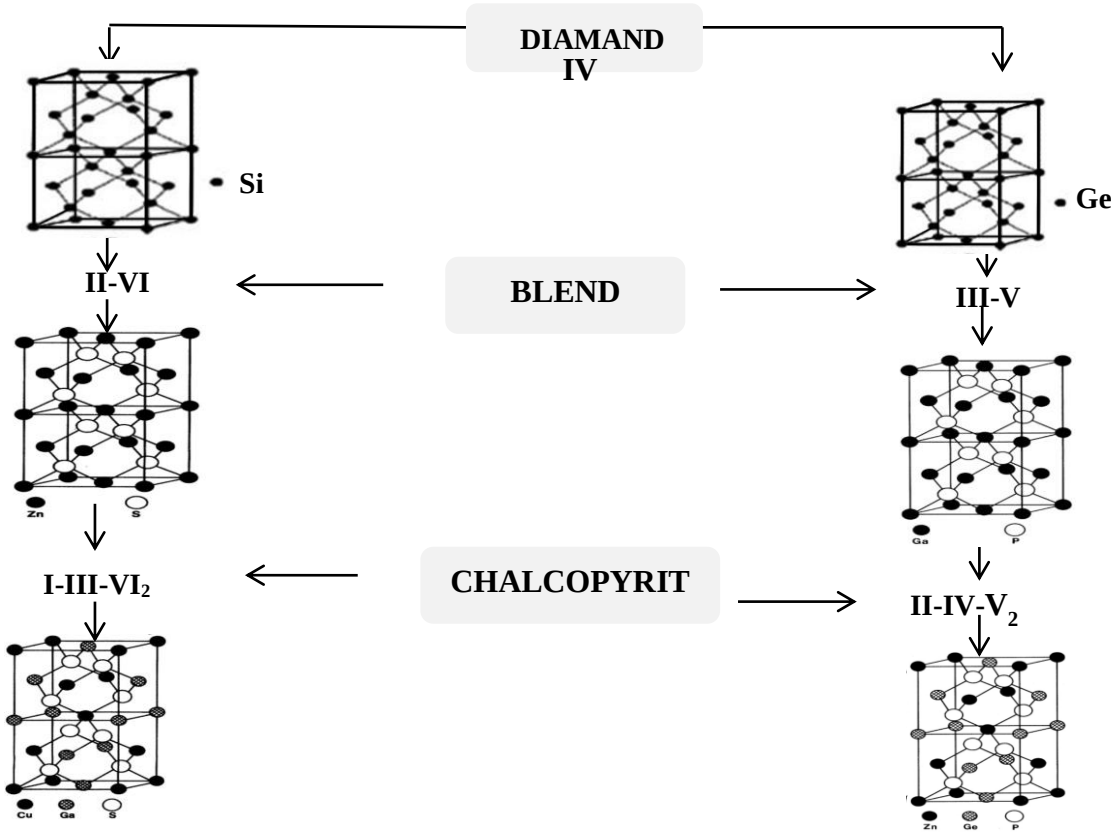

Figure.I.5 Diagram using the cross-substitution technique developed by Goodman [16].

Table.I.1 The coordinates and positions of the eight atoms in the chalcopyrite structure's tetragonal lattice as a function of a, c, and u are given by Zunger [23].

Atoms	Coordinates(x,y,z)
A_1	0, 0, 0
A_2	0, *a* /2, c/4
B_1	*a* /2, *a* /2,0
B_2	*a* /2, 0, c/4
C_1	*a* (1/4+u), *a* /4,c/8
C_2	*a* (3/4-u),3 *a* /4,c/8
C_3	*a* /4, *a* (3/4+u), 3c/8
C_4	3 *a* /4, *a* (3/4-u), 3c/8

The chalcopyrite crystal's primitive cell is four times bigger than the zinc blende crystal's unit cell when it is crushed along the z-axis. This means that zinc blende's Brillouin zone is four times bigger than chalcopyrite's (Figure I.6).

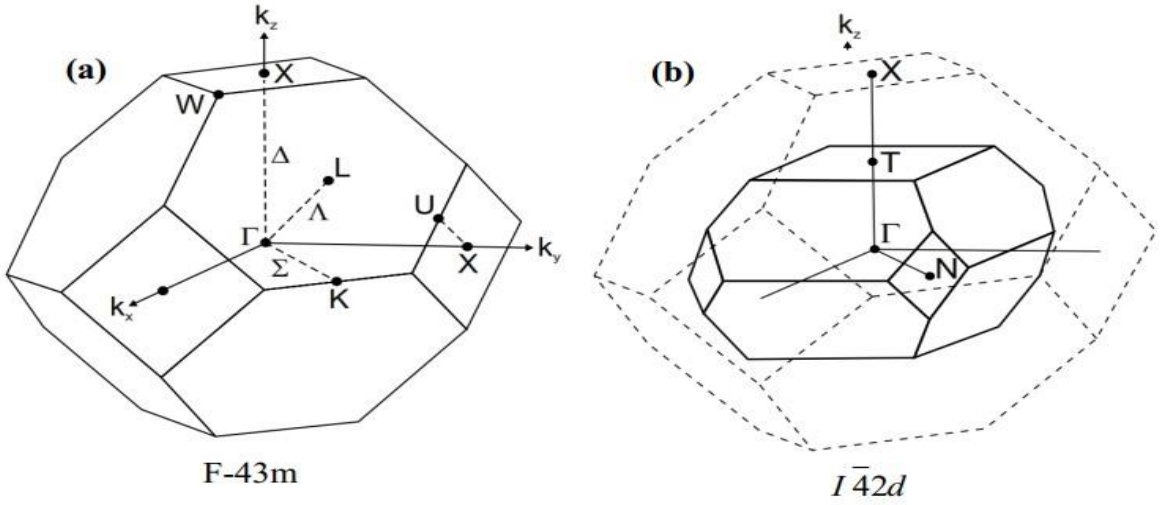

Figure.I.6 The graph representation of the first Brillouin zone for the two structures (a) zinc blende and (b) chalcopyrite [25].

It is assumed that the cations in the zinc blend analogue have been distributed randomly throughout the cation locations. Following that, the compound has the zinc blende structure,

where the anion occupies one of the two positions in the primitive cell and the average of the two cations occupies the other $\left(\frac{A^2+B^2}{2}\right)$. The Brillouin zone is only reduced by the order of the two cations and their different potentials [32].

The chalcopyrite quadratic structure for these types of crystals is often characterized by a ternary chalcopyrite, which is typically made up of a combination of two zinc blende structures and introduces a change in symmetry group from (F-43m) to ($I\bar{4}2d$). The lattice parameter c should be double that of parameter a [17], although chalcopyrite exhibit structural changes due to alternating cation: quadratic compression c/a and anionic displacement u, or the internal parameter [25, 33]. Only the distances between the initial neighbor (d_{AX} and d_{BX}) are affected by the internal parameter u. Furthermore, the lattice is deformed with a unit cell height that is slightly less than twice the base due to the radius differences between the two cations. Along the crystallographic c-axis, the entire crystal is crushed. This quadratic compression is defined by τ [25, 33] where:

$$u = \frac{1}{4} + \frac{d_{AX}^2 - d_{BX}^2}{a^2} \quad \text{(I.1)}$$

$$\tau = 2 - \frac{c}{a} \leq 0 \quad \text{(I.2)}$$

$$\eta = \frac{c}{2a} \quad \text{(I.3)}$$

In the case of an ideal chalcopyrite, η = 1, c/a = 2, τ =0 and u = 1/4 [17]. A tetrahedron of atoms X coordinates the two atoms A and B. the two most important distances, d_{AX} and d_{BX}, are connected to the cell length and the X atom by its free internal position According to the relation [18, 36]:

$$d_{AX} = \sqrt{a^2u^2 + \frac{4a^2+c^2}{64}} \quad \text{(I.4)}$$

$$d_{BX} = \sqrt{a^2(u^2 - \frac{1}{2})^2 + \frac{4a^2+c^2}{64}} \quad \text{(I.5)}$$

I.3. Chalcopyrite compounds properties and applications

Chalcopyrite compounds and their defects, as well as doped or substituted alloys, are the most interesting options from both an experimental and a theoretical point of view. This is due to their potential applications in solar cells, optoelectronic devices, electro-optics, and nonlinear

optical devices. They are a broad group of semiconducting materials with a variety of structural, optical, and electrical properties [17, 22, 23-34]. Such as:

- Their gap energy is within the ideal range for converting solar radiation into electrical energy at the earth's surface.
- Their lattice constant and electronic affinities are well coordinated with those of CdS, a large-gap semiconductor that has already demonstrated its use in solar cells.
- These compounds' low thermal conductivities are what first prompted research into them. These are the most likely materials to be used in solar cells [4], photovoltaic detectors, modulators, filters (like optical light eliminator filters [35]), light-emitting diodes [36], nonlinear optics [37], and optical frequency conversion in all solid-state-based tunable laser systems [38].

I.4. Materials with perovskite structure

I.4.1. Perovskite with ABX_3 formula

This mineral species was first described in 1839 by German geologist Gustav Rose, based on samples from the Ural Mountains in Russia, and named after the great Russian mineralogist, Count Lev Aleksevich Von Perovski (1792–1856) [39]. The mineral with the formula $CaTiO_3$, which is composed of calcium oxide and titanium, is the source of the name perovskites, which originally comprised a large family of crystalline materials [40].

In general, compounds with the formula ABX_3 are commonly referred as "perovskite". Which is typically cubic and belongs to the space group Pm3m (space group no. 221 in international crystallographic databases), with a lattice parameter on the order of (see figure I.7) [41]. The structure consists of a three-dimensional series of BX_6 octahedral connected by vertices divided into corners, where each B-X-B angle is 180° [42].

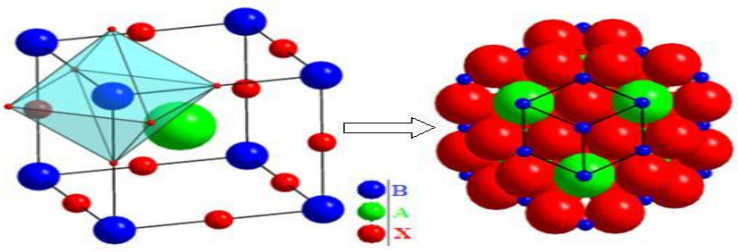

Figure.I.7 Description of ABX_3 perovskite [41].

Where:

- "A" can be an alkaline earth, or rare earth (lanthanide) cation (La, Sm, Ba, Gd, Ca, Pr,….).of large ionic radius with a coordinance equal to 12.
- "B" a transition metal cation (Cr, Mn, Fe, V, Ti, Al....) of smaller ionic radius than A. with a coordination of 6,
- "X" is typically oxygen, but can also occasionally be sulfur, nitrogen, halogen, or hydroxide. (see figure I.8)

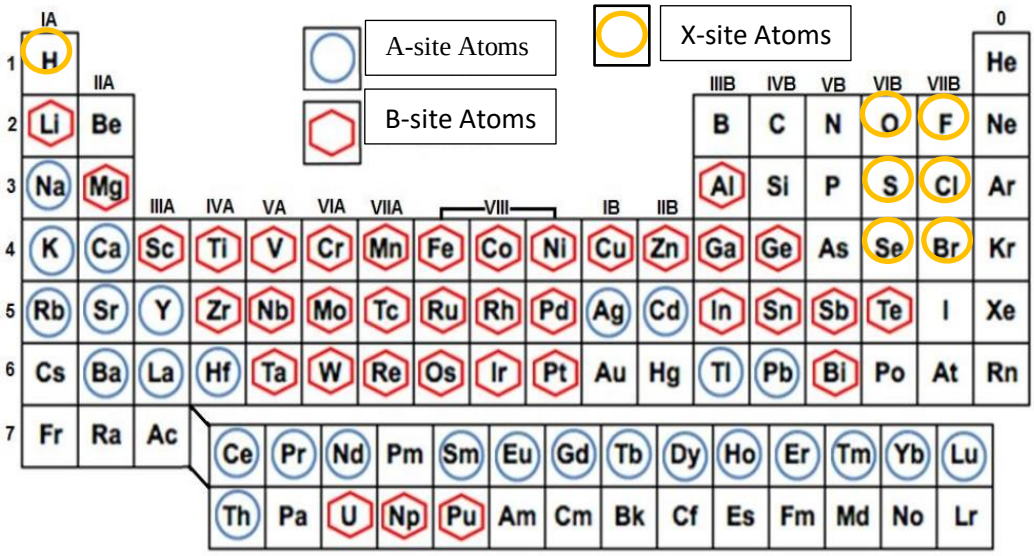

Figure.I.8 General representation of simple perovskites ABX_3[42].

Numerous possible compounds may be created by modifying these A and B cations in pairs, which changes the material's basic properties. Due to the possibility of observing a wide range of intriguing properties based on the chemical and electronic nature of the A and B atoms, including electronic, electrical, magnetic, optical, and more, this structure has attracted the interest of researchers over the past 20 years (Figure I.9). Based on the chemical element composition of the ABX_3 crystal structure, perovskite materials can be classified into two basic types [43]: inorganic oxide perovskites and halide perovskites. Furthermore, depending on the element X in the halide perovskite, there are several categories of materials that occur naturally on Earth, such as the chloride perovskite $ABCl_3$ and the fluoro perovskite ABF_3.

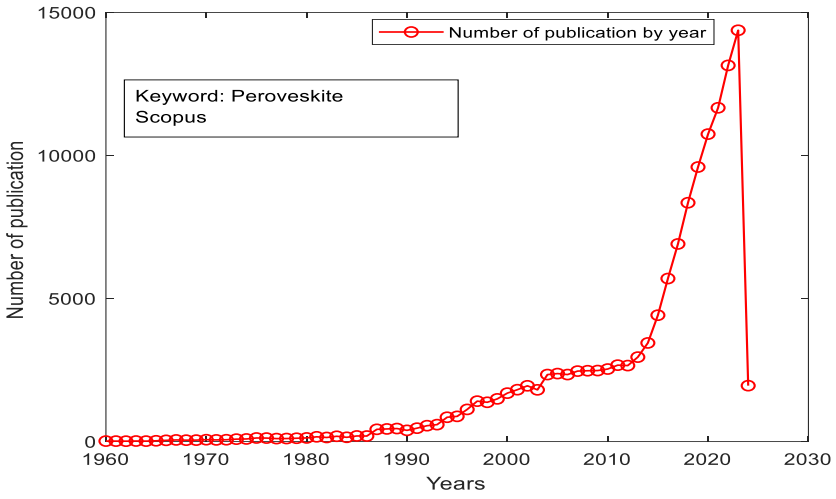

Figure.I.9 Number of publications by year on perovskites. **Keyword** search "Perovskites" in Title, Abstract and Keywords. Source **Scopus.**

I.4.2. The Structural Perovskite with ABO_3 formula

There are four types of perovskites oxide depending on the use of sites A and B [44]:

- The basic perovskites where just one sort of atom occupies sites A or B: $LaMnO_3$, $LaCoO_3$, $LaFeO_3$, and $CeNiO_3$.
- The perovskites complexes where one of the two sites, A or B, is occupied by two different types of atoms: Pb $Mg_{1/3}Nb_{2/3}O_3$, Pb $Sc_{1/2}Ta_{1/2}O_3$, etc.
- Superstructures in which the two sites A and B are occupied by various cation types, such as $(PbLa)(ZrTi)O_3$.
- Double perovskites: the mesh volume is twice that of the perovskite mesh. for example: Ca_2CoNbO_6

In 1926, Victor Goldschmidt first described the ABO_3 type structure of perovskites in his study on tolerance factors. Later in 1945, Helen Dick Megaw used X-ray diffraction on barium titanium ($BaTiO_3$) to publish it [43]. Its Bravais network is a simple cubic of centro-symmetric space group, which consists of an anionic sub lattice and two cationic sublattices made up of cations with various sizes and occupying specific positions [43].

The cubic structure is considered ideal when all of the atoms are in close proximity. Where the distance between B and O is a /2 and the distance between A and O is a /√2. As a result, the ionic radii of atoms A, B, and O have the relationship (r_A $+r_O$,) = √2 (r_B $+r_O$), a is the lattice

parameter. This implies that there is a structural deformation that converts cubic symmetry to rhombohedral, orthorhombic, or monoclinic symmetry [44].

Goldschmidt [45] proposed the tolerance factor t, which predicts the symmetry of the structure [46], for measuring the deviation from ideality. It assesses the constraints in the perovskite structure according to the relationship between the ionic radii of atoms A, B, and O, which are represented by r_A, r_B, and r_O, respectively by using equation I.6. The interatomic distances between atoms A and B are represented by the symbol r [47].

$$t = \frac{r_A + r_B}{\sqrt{2} r_B + r_A} \quad \text{(I.6)}$$

The unit cell of the ideal cubic structure can experience many distortions accompanied by a series of phase transitions, dependent on the size and composition of cations A and B as well as the tolerance factor (Table I.2) [42]. The specifics of the octahedral rotations can determine whether these structural distortions are orthogonal, rhombohedral, tetragonal, monoclinic, or triclinic in nature [42].

Table.I.2. The different symmetries adopted by the perovskite structure as a function of the tolerance factor $\boldsymbol{t}$.

Tolerance factor t	observed symmetry
1<t<1.06	Hexagonal
0.95<t<1	Cubic
0.90<t<0.95	Quadratic
0.80<t<0.90	Orthorhombic

In addition to the tolerance factor, there are other requirements for the ideal stability of the structure [44]:

- The volume, where the VA/VB polyhedron's volume ratio for cations A and B, respectively is 5. The degree of distortion of the perovskite structure may be determined with the use of this VA/VB ratio. The structure's distortion decreases with increasing VA/VB ratio.

- The coordination as a function of ionic radius, where cation B must have an ionic radius greater than 0.51 Å in order to retain coordinacy 6, and the ionic radius of A must be greater than 0.9 Å for coordinacy 12.
- The iconicity of anion-cation bonds: The average ionic character of the ABO_3 structure is quantified from the differences in electronegativity given by the Pauling scale [48].

$$\bar{\chi} = \frac{\chi_{A-O}+\chi_{B-O}}{2} \quad (I.7)$$

Where: χ_{A-O} , χ_{B-O} are the differences in electronegativity between cations A and B and the associated oxygens, respectively. The perovskite structure becomes more stable when the bonds mentioned above exhibit strong ionic behavior. Covalent perovskites are less stable than more ionic perovskites [49].

In general, the ideal ABO_3 perovskites are defined in the Pm-3m space group, which have a cubic lattice element and a lattice parameter of order a = 3.9A°. This is a three-dimensional structure where cation A, in the center of a 12-oxygen polyhedron, is linked together by square faces, while cation B, the smallest cation, is placed in an octahedral medium, the octahedral being linked together by the vertices. The structure is detailed in Table I.3, according to Miller and Love's choice of origin [50]. If A is at the origin, in position (0, 0, 0), B is at the center of the cube, in position (1/2, 1/2, 1/2), and the oxygens are located in the middle of each face, in position (0, 1/2, 1/2) (Figure I.10.a), and if the origin is shifted by one vector (1/2, 1/2,1/2), which leads A to occupy position (1/2, 1/2, 1/2), B position (0, 0, 0), and the oxygens are located in the middle of each edge, in position (1/2, 0, 0) (Figure I.10.b), X corresponds to oxygen.

Table.I.3. The atomics positions of a cubic structure [44].

A at the origine	A at the origine
A(0,0,0)	B(0,0,0)
B(0.5,0.5,0.5)	A(0.5,0.5,0.5)
O1(0,0.5,0.5)	O1(0.5,0,0)
O2(0.5,0,0.5)	O2(0,0.5,0)
O3(0.5,0.5,0)	O3(0,0,0.5)

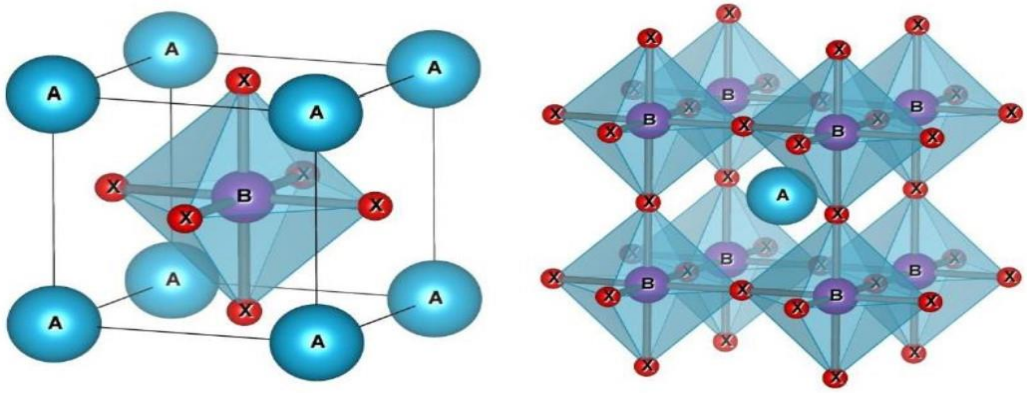

Figure.I.10 Different representations of the ideal crystal structure for the perovskite ABX_3 oxide:-(a) atom B at the origin -(b) atom B at the origin [44].

I.4.3.From single to double perovskite

In 1960, Ward and Longo [51] presented the first proposal for double-ordered perovskite oxides. Which is present in certain alloys and materials with $A_2BB'X_6$ compositions. The crystal structure of double perovskites is similar to that of single perovskites. They called it a double perovskite due to its unit parameter, which is twice that of a single perovskite. The double perovskite structure $AA'BB'O_6$ consists of a combination of two materials having the perovskite structures ABO_3 and $AB'O_3$ in the three crystallographic directions. It is derived from the perovskite ABO_3 when half of the octahedral BO_6 is replaced by another suitable B'O6 octahedral for the B' cations. Where A and A' can be alkaline earth metals or lanthanides, and B and B' are transition metals. Often, A and A' represent the same element. In the double perovskite structure, the X anions form an octahedron around each B and B' atom. At the same time, each B atom is surrounded by six B' atoms, and vice versa. These atoms are arranged in the cube as follows: the A cations are in the center, the B cations are at the vertices, and the B' anions are at the midpoints of the cube's edges. In this case, the tolerance factor for the double perovskite structure is defined as [52]:

$$t = \frac{r_A + r_B}{\sqrt{2} r_{BB'} + r_A} \quad \text{(I.8)}$$

Where $r_{BB'}$ represents the B and B' cations' average ionic radius.

The ideal double perovskite also shows cubic symmetry for (t≈1) with a space group ($Fm\bar{m}$), and the lattice parameter is double that of the simple cubic ABX_3, which is studied by Lavat and Baran [53, 54].

Two subfamilies of double perovskites are further distinguished: the first includes B-site-ordered compounds with the formula $A_2BB'O_6$. Where A usually denotes an alkaline earth metal, B and B' are transition metals, and X is a oxygen anion (figureI.11). This family, includes Sr_2LiOsO_6, Pr_2NiRuO_6, Er_2NiRuO_6, and others, is well-known and the objective of much study by both theorists and experimentalists. The second contains compounds with A-site orders that have the formula $AA'B_2O_6$. It is also possible to mix materials with various A and B atoms to create an $AA'BB'O_6$ structure, in which the B and B' atoms and the A and A' atoms are arranged in a cubic configuration [54]. In our studies we are only interested with the simple perovskites structure ABX_3 and the double perovskites structure with $A_2BB'O_6$ formula.

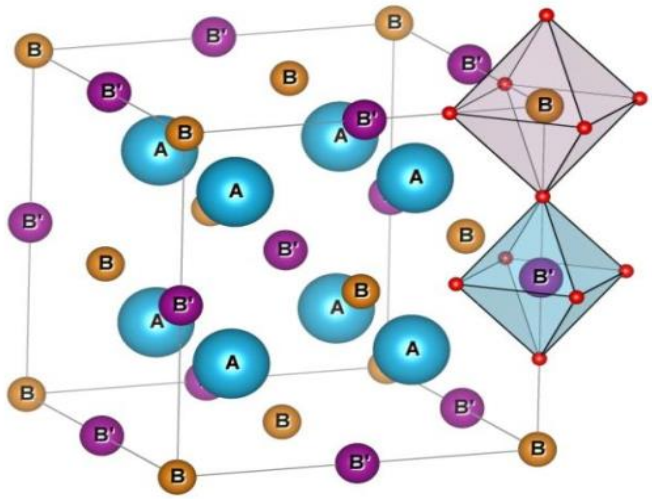

Figure.I.11 The elementary cell for the double perovskite $A_2BB'O_6$ ordered at B[40].

I.5.Perovskite properties and applications

Microelectronics uses perovskites extensively because of their high dielectric permittivity, high piezoelectric coefficient, semi-conductivity, ferroelectricity, thermoelectricity, catalytic activity, their optoelectronics properties and most importantly, their low cost and abundance [55].These properties have a wide range of technological applications, including high-

temperature oxygen probes, optical waveguides, dynamic selective access memories, surface acoustic wave devices, and frequency doubles [55–57]. Furthermore, these highly promising materials are crucial to modern electronics, where they are used for memory, capacitors, microwaves, and ultrafast electronics [40]. In addition, they have been widely used in optoelectronic devices, including solar cells [58], photodetectors [59, 60], and light-emitting diodes (LEDs) [61, 62]. Therefore, perovskites make excellent candidates for the development of new materials that meet the needs of emerging industries, especially those related to modern electronics. Recently, the research has focused on perovskites for solar cells, as shown in figure I.12.

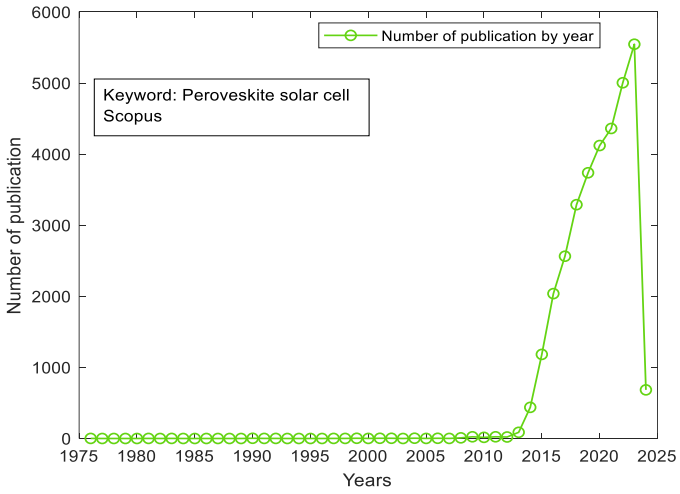

Figure.I.12 Number of publications by year on solar cell based on perovskite. **Keywords** search "perovskite solar cell" in Title, Abstract and Keywords. Source **Scopus.**

I.6.Materials discovery and Technological Progress

The rate of technological advancement has been in line with what is known as Moore's Law, which bears the name of Gordon E. Moore, who postulated in 1965 that the complexity of component costs would double annually until 1975 and then every two years after that [63]. But by 2025, it's anticipated that it will come to an end since conventional transistor scaling has hit its physical limits. [64] Intel's growth has already doubled in the last almost six years. [65] This makes finding unconventional means to maintain the current pace of technological advancement

necessary. Scaling down the creation and production of materials and gadgets to the atomic level is one approach.

The properties of a material may be understood as the result of the interplay of several factors, each with varying relative importance based on the final target. This includes a wide range of intrinsic and extrinsic factors that can be connected to the preparation circumstances, such as the atomic composition, the morphology and microstructure of the material, its physical state, and others [66]. The idea of creating novel tools that can simulate these interactions and assist in predicting the properties of a possible new compound is motivated by this basic observation. In order to refine the search space and narrow down the pool of potential materials, the developed models have to capitalize on the existing theoretical foundation.

A new infrastructure for material innovation was built by taking advantage of this exceptional potential. This infrastructure, as seen in Figure I.13, is predicated on the interoperability of three primary processes: theory, experiments, and computer simulations. Each of these operations involves distinct fields of study and skill [67].

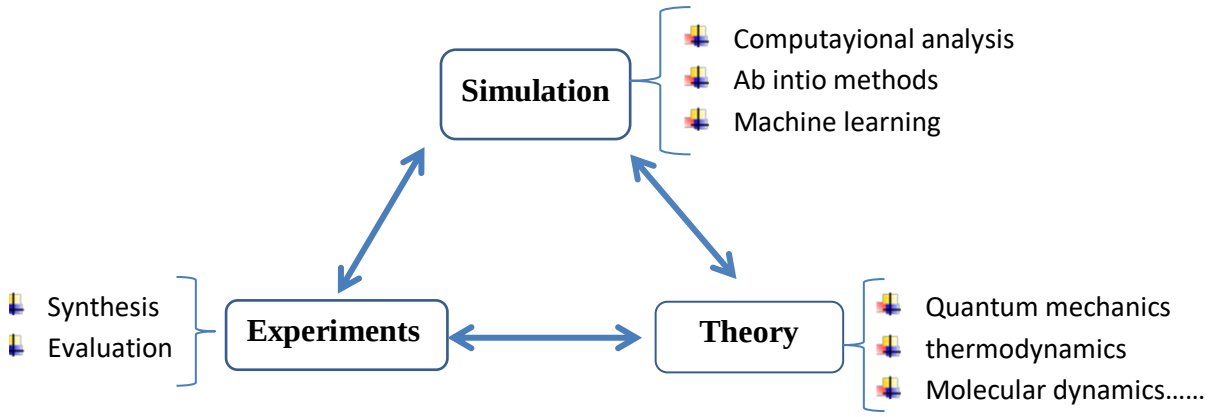

Figure.I.13 Infrastructure for Materials Discovery [67].

I.7 Conclusion

In this chapter, we have provided an overview of two types of compounds: chalcopyrite materials and perovskites with a particular focus on the simple ABX_3 and double perovskites

$A_2BB'O_6$ materials structures. We defined chalcopyrite based on its crystalline structures, and we also explored the properties of perovskites based on their potential applications. Next, we expanded our discussion to include perovskites and their various types, exploring their diverse fields of interest due to their intriguing properties.

Perovskites and chalcopyrite are among the most promising materials for a variety of applications, particularly optoelectronic devices, which are currently considered to be the next generation of photovoltaic cells. Hence, in many applications, thorough knowledge of these materials' optical and electronic properties is important. Despite the significant amount of research on these materials, this knowledge continues to evolve. In this research, we aim to forecast various attributes, including electronics and structural characteristics, through innovative techniques rooted in artificial intelligence and ab initio, as detailed in the upcoming chapters.

References

[1] M.A.Rahman, (2014). A review on semiconductors including applications and temperature effects in semiconductors. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 7(1), 50-70.

[2] Z.Guo, J.Li, R.Pan, J. Cheng & al (2020). All-inorganic copper-based ternary metal halides: promising materials toward optoelectronics. Nanoscale, 12(29), 15560–15576.

[3] M.W. Iqbal, M.Asghar, N. A. Noor & al (2021). Analysis of ternary $AlGaX_2$(X = As, Sb) compounds for opto-electronic and renewable energy devices using density functional theory. Physica Scripta, 96(12), 125706.

[4] L. L. Kazmerski (1983). Ternary-compound thin-film solar cells. Il Nuovo Cimento D, 2(6), 2013–2028.

[5] K.D. EDDINE, 2018 Etude des propriétés physiques des chalcopyrites ABC_2 (A=Zn, Cd; B=Ge, Si; C=P, As) .Doctoral dissertation, University of Sciences and Technologie of Oran.

[6] S. Sun, M.Lu, X.Gao, Z.Shi & al (2021). 0D Perovskites: Unique Properties, Synthesis, and Their Applications. Advanced Science, 8(24). Portico. https://doi.org/10.1002/advs.202102689

[7] W.J.Yin, T. Shi & Y. Yan (2014). Unusual defect physics in $CH_3NH_3PbI_3$ perovskite solar cell absorber. Applied Physics Letters, 104(6). https://doi.org/10.1063/1.4864778

[8] K. X.Steirer, P.Schulz, G.Teeter & al. (2016). Defect Tolerance in Methylammonium Lead Triiodide Perovskite. ACS Energy Letters, 1(2), 360–366.

[9] J.Burschka, N.Pellet, S.J.Moon & al (2013). Sequential deposition as a route to high-performance perovskite-sensitized solar cells. Nature, 499(7458), 316–319.
[10] G.Xing, N. Mathews, S. S. Lim & al (2014). Low-temperature solution-processed wavelength-tunable perovskites for lasing. Nature Materials, 13(5), 476–480.

[11] Z.K.Tan, R.S.Moghaddam, M.L.Lai & al . (2014). Bright light-emitting diodes based on organometal halide perovskite. Nature Nanotechnology, 9(9), 687–692.

[12] S.T.Ha, C.Shen, J. Zhang & , Q. Xiong. (2015). Laser cooling of organic–inorganic lead halide perovskites. Nature Photonics, 10(2), 115–121.

[13] S.Sharma & A. S. Verma,(2014). Structural, electronic, optical, elastic and thermal properties of $ZnXAs_2$ (X = Si and Ge) chalcopyrite semiconductors. The European Physical Journal B, 87(7).

[14] H.Hahn, G. Frank, W. Klingler, A. Meyer &G. Störger, (1953). Untersuchungen über ternäre Chalkogenide. V. Über einige ternäre Chalkogenide mit Chalkopyritstruktur. Zeitschrift Für Anorganische Und Allgemeine Chemie, 271(3–4), 153–170.

[15] C.H.L.Goodman &R.W. Douglas, (1954). New semiconducting compounds of diamond type structure. Physica, 20(7–12), 1107–1109.

[16] C.H.L.GOODMAN, (1957). A New Group of Compounds with Diamond type (Chalcopyrite) Structure. Nature, 179(4564), 828–829.

[17] Wernicke & al, Ternary Chalcopyrite Semiconductors: Growth, Electronic Properties, and Applications. (1975).

[18] R.MAHDJOUBI, (2019) Propriétés structurales,electroniques, optique et thermiques des ternaires $CuXTe_2$ (X= Al, Ga, In).. Doctoral dissertation, University of Sciences and Technologie of BADJI MOKHTAR .

[19] J. L.SHAY& J. H. WERNICK, (1975). INTRODUCTION. Ternary Chalcopyrite Semiconductors: Growth, Electronic Properties, and Applications, 1–2.

[20] L. L.Kazmerski, M. S.Ayyagari & G. A. Sanborn, (1975). $CuInS_2$ thin films: Preparation and properties. Journal of Applied Physics, 46(11), 4865–4869.

[21] J. E.Jaffe & A. Zunger, (1983). Electronic structure of the ternary chalcopyrite semiconductors $CuAlS_2$, $CuGaS_2$, $CuInS_2$, $CuAlSe_2$, $CuGaSe_2$, and $CuInSe_2$. Physical Review B, 28(10), 5822.

[22] J. E. Jaffe & A. Zunger, (1984). Theory of the band-gap anomaly in AB C 2 chalcopyrite semiconductors. Physical Review B, 29(4), 1882.

[23] A.Zunger & J. E. Jaffe, (1983). Structural Origin of Optical Bowing in Semiconductor Alloys. Physical Review Letters, 51(8), 662–665.

[24] Y.Zeng, S. J.Chua & P. Wu, (2002). On the Prediction of Ternary Semiconductor Properties by Artificial Intelligence Methods. Chemistry of Materials, 14(7), 2989–2998.

[25] H.BENNACER,(2015). Contribution au développement des nouvelles cellules solaires à base des semi-conducteurs composes. Doctoral dissertation, University of Djilali liabes sidi bel abbes.

[26] H. Xie, X. Su, G.Zheng& al. (2016). The Role of Zn in Chalcopyrite CuFeS2: Enhanced Thermoelectric Properties of $Cu_{1-x}Zn_xFeS_2$ with In Situ Nanoprecipitates. Advanced Energy Materials, 7(3).

[27] C.L.Burdick & J. H. Ellis, (1917). The Crystal Structure of Chalcopyrite Determined by X Rays. Proceedings of the National Academy of Sciences, 3(11), 644–649.
[28] L.Pauling &L.O. Brockway, (1932). The Crystal Structure of Chalcopyrite CuFeS2. Zeitschrift Für Kristallographie - Crystalline Materials, 82(1–6), 188–194.

[29] T.Kambara, (1974). Optical Properties of a Magnetic Semiconductor: Chalcopyrite CuFeS. II. Calculated Electronic Structures of $CuGaS_2$ and $CuFeS_2$. Journal of the Physical Society of Japan, 36(6), 1625–1635.

[30] C. J.Ballhausen & H. B. Gray,(1962). The Electronic Structure of the Vanadyl Ion. Inorganic Chemistry, 1(1), 111–122.

[31] S.MISHRA,(2012) Structural, electronic and optical Properties of chalcopyrite type Semiconductors. Doctoral dissertation, National Institue of Technology ,Rourkela .

[32] D. Lincot, (2020). Chapitre 7 : Les filières photovoltaïques en couches minces et leurs perspectives d'application à l'habitat. La Chimie et l'habitat, 127–150.

[33] Ouahrani Tarik, (2011). Calcul des properties structurales, thermiques et optiques des composes chalcopyrites par la méthode FP-(L)APW . Doctoral dissertation, University of Tlemcen.

[34] Parthé, E. (1964). Crystal chemistry of tetrahedral structures. CRC Press.

[35] H.Horinaka, S.Mononobe & N. Yamamoto, (1993). A criterion for Applying Chalcopyrite Semiconductors to Optical Line Elimination Filters. Japanese Journal of Applied Physics, 32(S3), 109.

[36] J. L.Shay, L. M.Schiavone, E.Buehler & J. H. Wernick, (1972). Spontaneous- and Stimulated-Emission Spectra of $CdSnP_2$. Journal of Applied Physics, 43(6), 2805–2810.

[37] B. F. Levine, (1973). Bond-Charge Calculation of Nonlinear Optical Susceptibilities for Various Crystal Structures. Physical Review B, 7(6), 2600–2626.

[38] F. K. Hopkins, (1995). Nonlinear materials extend the range of high-power lasers. Laser Focus World, 31(7), 87-93.

[39] E.C.C. Souza & R. Muccillo, (2010). Properties and applications of perovskite proton conductors. Materials Research, 13(3), 385–394. [40]S.HAID, (2019) .Etude Ab-Initio des propriétés structurales, électroniques et thermodynamiques des matériaux double Perovskite. Doctoral dissertation. University of Mostaganem-Abdelhamid Ibn Badis.

[41] M.S.Mahboub &A. Boudjada (2012). Synthèse, caractérisation par diffraction X et spectroscopie Raman des composés Ca1-xSrxFeO2. 5-δ (δ= 0, 0.5).

[42] M.Aissat Fares,(2019).Synthése ,caracterisation et activité catalyque des materiaux de structure peroveskite.2019. Doctoral dissertation.University of USTHB .

[43] N.D.CANICOBA, (2018).Development and characterization of perovskite based devices:field effect transistors and solar cells. 2018. Doctoral dissertation .University of Rennes1.

[44] W.Bensomane.Etude des Propriétés Optoélectroniques, Magnétiques et Thermodynamique des Pérovskites et des Doubles Pérovskites.2020. Doctoral dissertation.University of Abdelhamid Iben badis of Mostaganem.

[45] P.Gillet, F. Guyot, G. D.Price & al (1993). Phase changes and thermodynamic properties of CaTiO3. Spectroscopic data, vibrational modelling and some insights on the properties of MgSiO3 perovskite. Physics and Chemistry of Minerals, 20(3), 159–170.

[46] J.F.Marucco, (2004). La chimie des solides. EDP Sciences.

[47] R. D. Shannon, (1976). Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. Acta Crystallographica Section A, 32(5), 751–767.

[48] L.Pauling, (1931). The Nature Of The chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. Journal of the American Chemical Society, 53(4), 1367–1400.

[49] Lamrani Epouse Amaouz Naoura,(2011).Synthèse et caractérisation des matériaux diélectriques à structure pérovskite complexe de type $Ca_{1-x}A_xTi_{1-y}$ B_yO_3(A=Sr,B=Zr...). Doctoral dissertation, University of Mouloud Mammeri Tizi ouzou

[50] S. C. Miller and W. F. Love,Tables of Irreducible Representations of Space Groups andCorepresentations of Magnetic Space Groups, Pruett. (Boulder, Col.) (1967)]

[51] R. Ward & J. Longo, (1960).Magnetic phases of the perovskite type .Journal of the American Chemical Society, 82(22), 5958–5958.

[52] V. M. Goldschmidt, (1926). Die geset zeder krystallochemie.Naturwissenschaften14(21) 477-485.

[53] A.E.Lavat & E. J. Baran, (2003). IR-spectroscopic characterization of A2BB′O6 perovskites. Vibrational Spectroscopy, 32(2), 167–174.

[54] A.Rachedi.(2021) Etude de premier principe des propriétés magnétiques, optoélectroniques et thermodynamiques des doubles pérovskites Ba_2XWO_6 (X : Ni, Co, Mn). Doctoral dissertation.University of Ibn khaldoun Tiaret.

[55] M.A.TAG, (2017).Etude et modelisation de l'anistropie des peroveskites et des nanoparticules magnetique .Doctoral dissertation. University of BADJI MOKHTAR-ANNABA.

[56] V. E.Henrich & P. A. Cox, (1994). The surface science of metal oxides. Cambridge university press.

[57] E.Mete, R.Shaltaf & Ş. Ellialtıoğlu, (2003). Electronic and structural properties of a 4 d perovskite: cubic phase of $SrZrO_3$. Physical Review B, 68(3), 035119.

[58] Best Research-Cell efficiency chart, https://www.nrel.gov/pv/ cell-efficiency.html (accessed: March 2021).

[59] S.Li, Y.Xu, C.Li & al. (2020). Perovskite Single-Crystal Microwire-Array Photodetectors with Performance Stability beyond 1 Year. Advanced Materials, 32(28).

[60] L.Dou, Y.Yang, J.You & al. (2014). Solution-processed hybrid perovskite photodetectors with high detectivity. Nature Communications, 5(1).

[61] M.Lu, H.Wu, X.Zhang & al. (2018). Highly Flexible CsPbI3 Perovskite Nanocrystal Light-Emitting Diodes. ChemNanoMat, 5(3), 313–317.

[62] H.Wu, Y.Zhang, X.Zhang, M.Lu al . (2017). Fine-Tuned Multilayered Transparent Electrode for Highly Transparent Perovskite Light-Emitting Devices. Advanced Electronic Materials, 4(1). Portico.

[63] G. E. Moore, (1998). Cramming More Components Onto Integrated Circuits. Proceedings of the IEEE, 86(1), 82–85.

[64] M. M. Waldrop, (2016). The chips are down for Moore's law. Nature, 530(7589), 144–147.

[65] D.Clark, (2015). Intel rechisels the tablet on Moore's law. Wall Street Journal.

[66] R.Potyrailo, K.Rajan, K.Stoewe & al . (2011). Combinatorial and High-Throughput Screening of Materials Libraries: Review of State of the Art. ACS Combinatorial Science, 13(6), 579–633.

[67] F.KHMAISSIA,(2017). Data driven discovery of materials properties. Doctoral dissertation . University of Louisville.

CHAPTER II

AN OVERVIEW OF ARTIFICIAL INTELLIGENCE AND OPTIMIZATION ALGORITHMS

II.1. Introduction

Artificial intelligence (AI) has steadily grown to be a crucial driver for the fourth industrial revolution with the quick advancement of computer performance. Several computational scientists proposed the first idea of artificial intelligence (AI) at the Dartmouth conference in 1956 [1]. Their initial goal was to investigate the use of machines to mimic human learning and other aspects of intelligence [1].

Artificial intelligence technology is currently a trend that is spreading around the globe as human society develops. Particularly, it is considered an effective use of visual recognition, speech recognition, natural language comprehension, and man-machine games, which also speed up the fusion of several fields of knowledge [2]. In general, there are three steps for the implementation of AI: handwritten knowledge, statistical learning, and context adaptation, which is currently in the statistical learning stage [3]. Since 2012, the rapid growth of data volume brought on by superior monitoring equipment or devices, the advancement of computing power, the introduction of a new machine learning (ML) algorithm, and the development of deep learning (DL) have all continuously contributed to the growing popularity of AI [4]. The fundamental technologies, including deep learning [5], reinforcement learning [6], adversarial learning [7], etc., rely on statistical machine learning. It could significantly improve the effectiveness of research in fields such as information data mining, data classification, and new data prediction [2]. Despite their relationship, the terms "AI, ML, and DL" are actually distinct from each other. In general, ML can be seen as a way to actualize AI, capable of identifying inherent patterns in data and then offering a logical choice as a recommendation. DL is considered the main branch of ML, which constantly uses more complicated multi-layer neural network topologies and is inspired by how human brains process information. Despite DL requiring less human supervision than other ML techniques, it still necessitates large volumes of data to uncover the complex, varied, and inherent relationships hidden within the data [1]. The majority of machine learning algorithms work by constructing a surrogate model and then optimizing the objective function (or loss function) using the optimization algorithms (OA) to produce the best-performing, ideal model [1].

In fact, artificial intelligence (AI) based on this machine learning process has emerged as the most significant source of innovation in science and industry. However, the industrial revolution

and the development of human society both depend greatly on material innovation. In the near future, the development of novel material science and technology is accelerating in an interconnected and intelligent direction, which will further promote the fourth industrial revolution, also known as the green intelligence revolution [8]. One of the most important AI fields is machine learning, which is a powerful tool for material innovation thanks to its low computer costs, quick development cycles, and great data analysis and prediction capabilities when compared to other experimental and computational techniques. As far as we know, a few comprehensive research papers on machine learning and material application have attracted close attention in recent years, focusing on machine learning in physiochemical discovery [9–11]. It is particularly important to employ novel approaches to improve the accuracy of predictions in this field.

In this chapter, we gave a comprehensive overview and the basics of artificial intelligence, machine learning, and optimisation algorithms. Moreover, we introduced different approaches based on AI techniques, such as artificial neural networks, fuzzy logic methods, and different metaheuristic algorithms.

II.2. Artificial Intelligence

The idea of creating intelligent beings that can reason and perform tasks has long attracted humanity. According to Greek mythology, Hephaestus and Daedalus created metal statues of animated people and animals that could think. Early Judaism mentioned the Golems, mud-based artefacts capable of serving as servants. In the modern era, Karel Capek coined the term "robot" in 1920 refer to intelligent machines, and the idea has since been a recurrent theme in literature [12]. Alan Turing was the first scientist to postulate that a machine would be able to comprehend logic and carry out tasks, which gave rise to the Turing Test, a method for separating artificial intelligence from human intelligence [13]. Artificial intelligence research commenced in 1956 with significant enthusiasm at the Dartmouth Artificial Intelligence Conference [14], where the definition of AI was given as follows: "Every facet of learning or any other feature of intelligence can, in principle, be so clearly specified that a computer can be created to replicate it". Since then, numerous sciences in a variety of domains, including computer sciences and even philosophy, have continued to debate what it means to be an AI. And how could you say that a piece of equipment, a computer program, or anything else you have in your lab is an AI? [15].

Artificial intelligence is the foundation of computer science [15, 16]. It is a technological advancement that creates machines capable of performing tasks only human intellect can, such as teaching, reasoning, self-correction, and self-programming. However, the main goal of artificial intelligence is to enable computers to perform human-level tasks [17]. In several scientific and practical sectors, artificial intelligence is vital. Learning, pattern recognition, problem solving, reasoning, language comprehension, visual perception, and problem solving are some of the cognitive abilities that artificial intelligence can improve. In addition, AI studies are extremely technical, specialized, and significantly dependent on logic, information, planning, education, communication, imaging, and having the ability to move and handle objects [18].

There are many terms used in the AI industry. Figure II.1 illustrates that deep learning is a subset of machine learning, and machine learning is a subset of artificial intelligence, which is considered the most comprehensive collection [19]. This thesis primarily focuses on AI and ML. In summary, the difference between AI and ML is that the former refers to the broader concept of robots being capable of carrying out tasks in what we regard as "intelligent" ways. The idea that we should be able to provide machines access to data and let them learn on their own is at the heart of machine learning (ML) [20].

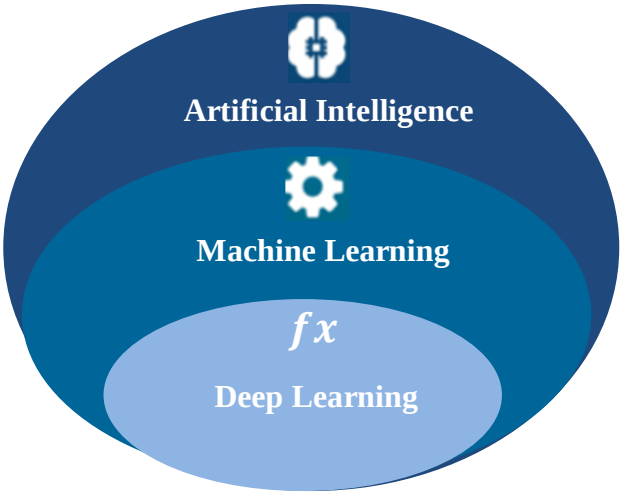

Figure.II.1 The relation between AI, ML and DL [18].

II.3.Machine Learning

To make it simpler to comprehend, machine learning is a subset of AI, or an application of AI. Professionals define machine learning as a class of algorithms designed to process input data, use statistical analysis to predict output data, and continuously update output data in response to

new input data [21]. A machine can categorize data into many groups based on the recognition of picture, color, size, text, and other types of elements that it contains. The machine may then provide you with the group of facts you desire, depending on any human need. Imagine how much time you could save in any sector by using machine learning techniques [18-22]. In general, machine learning (ML) is a subfield of artificial intelligence that focuses on using data and algorithms to mimic how humans learn through iterative training and improvement, without explicit human programming, while gradually increasing the accuracy of the model. This means that machine learning also enables the automatic generation of data, the training of algorithms, relationship detection, and pattern recognition. While it can be applied to a new and undiscovered data set to provide predictions and improve processes. Today, a variety of fields and applications use machine learning, and its use is steadily growing. For example, since the start of 2000, research in the materials science field using machine learning has received significant attention (see Figure II.2).

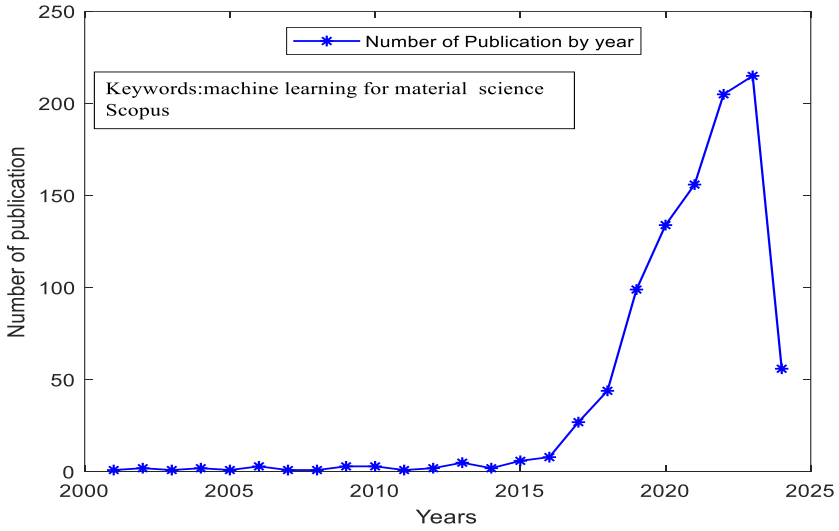

Figure.II.2 Number of publications by year of machine learning for materials. **Keyword** search "machine learning for materials science" in Title, Abstract and Keywords. Source **Scopus.**

Typical ML workflow can be summarized as follows [23]:

✓ Process for Making a Decision: Machine learning techniques are used to generate predictions or classifications of certain data that was labelled or unlabeled. These methods lead to an estimated pattern in the data, which helps in decision-making.

✓ Error Function: It evaluates a model's prediction and determines whether it produces true or false effects. Additionally, this function has the capacity to compare results in order to assess how well the model and machine learning methods are working.

✓ Model Optimization Process: If the recommended model effectively fits the training data set, we will adjust the weights to reduce the discrepancy between the suggested work and the estimated model. The technology will iteratively improve the implementation method, repeat this assessment, and enhance the weights until accuracy is attained.

In contrast to conventional software development, machine learning emphasizes independent learning from data and knowledge. ML systems, therefore, derive their own approach code from data and learn from it. These approaches adapt to their surroundings and develop according to the conditions they will encounter.

II.3.1. Types of machine learning

In formal terms, we can express the learning problem as follows [24]: Given a known set X, estimate or predict the unknown function $y = f(X)$. An element x from the set X is referred to as a feature (or property) vector, or simply an input. The set is termed a feature space. The capacity of the model to predict the output for unknown samples outside of the training data using the learned approximation function $\hat{y} = \hat{f}(X)$is known as model generalization. Based on the types of inputs and outputs handled, machine learning issues fall into a few categories (as shown in figure II.3) [25]:

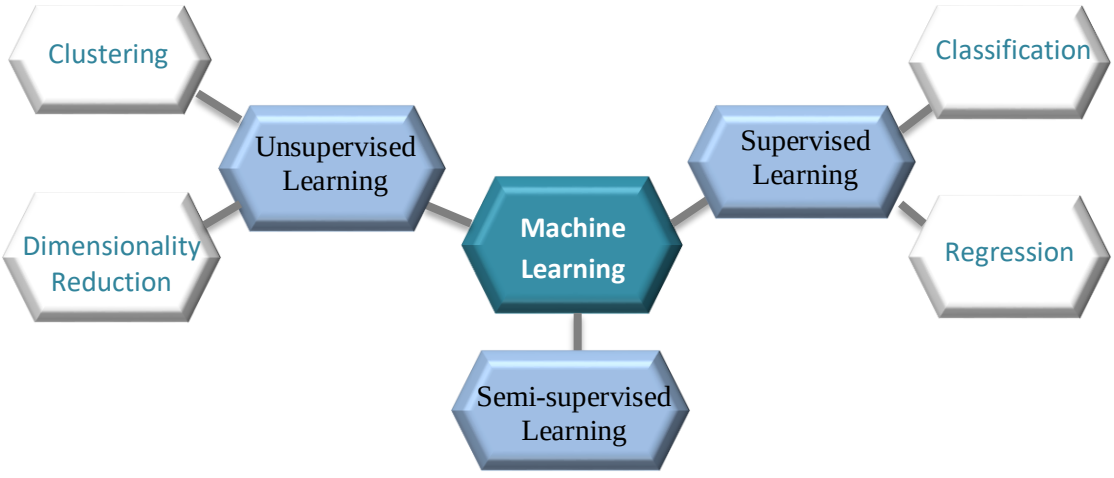

Figure.II.3 Machine Learning types [26].

❖ **Supervised learnng**: the aim of supervised learning is to discover the function that maps inputs to outputs given a collection of labelled data $(x_i, y_i) \in (X, f(X))$ with i = N instances, which is referred to as the training set (in contrast to an unknown test set). It is referred to as a classification issue if the output y_i type is a nominal or categorical finite set (such as a metal or an insulator) and it predicts the class label for unknown samples. Otherwise, it is referred to as a regression problem, which will forecast the output values for the unidentified samples, if the outputs are continuous real-valued scalars $y_i \in R$.

❖ **Unsupervised learning**: often referred to as descriptive learning, The objective of unsupervised learning is to identify structure in the data given just unlabeled inputs ($x_i \in X$) where outputs are unknown. The learning process, known as clustering if $f(X)$ is finite, divides data into a (known or unknown) number of clusters based on how similar their attributes are. Conversely, when $f(X)$ is in $[0, \infty)$, the learning process is referred to as density estimation, and it involves learning the marginal distribution of the features. Dimensionality reduction is another significant type of unsupervised learning, which reduces the number of input variables needed to describe the data and is helpful when $f(X)$ has a high dimensionality and a complicated data structure to identify patterns.

❖ **Semi-supervised learning :**Semi-supervised learning is another type of machine learning problem between supervised and unsupervised learning, which combines a large amount of unlabeled data with a small amount of labelled data; multi-task and transfer learning, which involve using information from related problems to improve the learning task (usually one with limited data available [27]); and reinforcement learning, which uses feedback on decisions rather than input or output in order to maximize a reward signal towards learning desired actions in an environment [24].

II.4. Regression analysis

The statistical method known as regression analysis is used to determine the correlations between a group of independent features known as "predictors" and a dependent variable. In a regression model, there are three basic variables: the features, which may be represented by a matrix X holding the set of descriptors, the dependent variable Y, and the unknown parameters, indicated by β, which can be either scalar or vector.

The model establishes a relationship between the dependent variable Y and the independent variable X, with the parameter β, expressed as $Y = f(X; \beta)$. This representation is typically formulated as $E(Y|X) = f(X; \beta)$, where $E(Y|X)$ denotes the expected value of Y given X.

Depending on the regression method employed, the unknown parameter, estimated as $\hat{\beta}$, is acquired through different means: it may involve deriving a closed-form expression, solving an estimating equation, or optimizing an objective function, often with specific constraints in place. Conceptually, β and $\widehat{\beta}$ represent the coefficients that elucidate the underlying relationship between the descriptors and the target. Subsequently, a regression model is constructed based on these coefficients [28]. In addition to fitting the training set well, the accuracy with which a regression model $f(X; \widehat{\beta})$ predicts a future response in the presence of new unlabeled test data is another sign of quality [29].

Additionally, to assess how robust the predicted regression models are, cross validation is frequently employed as a trustworthy average [28].

Let $Y = \{y_1, y_2, y_3, \dots . y_N\}$ in N dependent variable, $X = \{(x_{ij}), i = 1 \dots \dots N \; and \; j = 1 \dots P\}$ be the N p-dimensional feature vectors, and the model's parameter coefficients$\beta=\{\beta_1, \beta_2, \beta_3 \dots . \beta_p\}$. β can be calculated in several ways based on the regression approach that is employed.

II.5. Artificial Neural Network

II.5.1. Artificial Neurons

A fundamental component of each artificial neural network is the artificial neuron. Its structure and operations are based on the study of a biological neuron, which is the fundamental unit of biological neural networks, or systems, that comprise the brain, spinal cord, and peripheral ganglia [30]. Figure II.4 shows how the designs and functionality are similar between them. On the left side of the diagram (Figure.II.4.a), an artificial neuron with its inputs, weights, transfer function, bias, and outputs was represented, whereas a real neuron with its neurons, dendrites, and axons was presented on the right (Figure.II.4.b) [12].

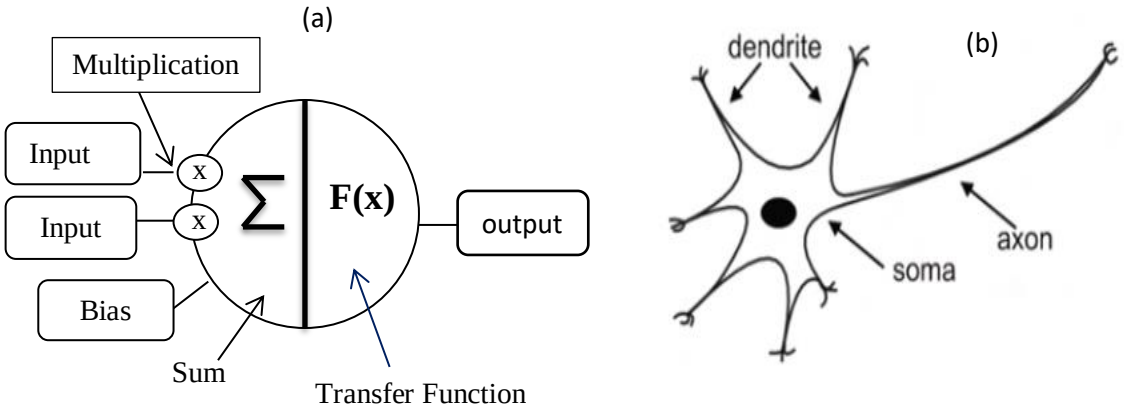

Figure.II.4 Design of artificial and biological neurons [30].

Generally, artificial neural networks (ANN) are machine learning algorithms that form the basis of deep learning [31]. McCulloch and Pitts created the ANNs in 1943 [32]. Both their name and structure are derived from brain neurons. Although ANNs are influenced by biological systems for information processing, they are not the same as living things. While ANNs are static and symbolic, the biological brains of most living things are dynamic and analogue, as noted in [33]. Based on its architecture (Figure II.4.a), the net function and the activation function are the two components that make up each neuron. However, the net function determines how the network inputs are integrated into the neuron [34]. The weighted linear combination used in this graphic is as follows:

$$u = w\,x + b \tag{II.1}$$

Where: x stands for the input linked to the neuron, w for the weight of the connection linking the input to the neuron, and b for the internal threshold of the neuron. The weighted sum of the neuron's inputs is represented by u.

The network input u and the neuron's output y are connected by a linear or nonlinear transformation known as the activation function:

$$y = f(u) \tag{II.2}$$

The sigmoid function, linear function, and hyperbolic tangent are some of the most often utilized transfer functions for time series prediction, despite the fact that many activation

functions have been reported in the literature. Table.II.1. illustrates different activation functions that are applied to ANN modeling.

Table.II.1. Different activation functions for modeling artificial neural networks [35].

Function	Formula
Log sigmoidal	$f(x) = \frac{1}{1+e^{-x}}$
Tan sigmoidal	$f(x) = \frac{1-e^{-x}}{1+e^{-x}}$
Hyperbolic tangent	$f(x) = \tanh x$
Pure linear	$f(x) = x$
Inverse tangent	$f(x) = \frac{2}{\pi}\tan^{-1}(x)$
Gaussian radial basis	$f(x) = \exp\left(-\frac{\|x-m\|^2}{\sigma^2}\right)$
Linear	$f(x) = ax + b$

II.5.2. Types of neural network

Neural networks fall into three main types: feed-forward neural networks, recurrent neural networks, and conventional neural networks.

II.5.2.1. Feed-forwardNeural networks

The Feed-forward Neural networks are the most commonly used type of neural network, in which data passes from the input layer to the output layer through the hidden layers and back again in one direction as well as the forward direction. There is not a cycle in the connections between the nodes [36]. Inputs are transformed by each hidden layer, producing a new representation at each level that follows. Neural networks are often arranged in layers; depending on how many layers feedforward neural networks can be classified as "single layer" or "multi-layer" [37].

- ✓ A single-layer feed-forward neural network structure is composed of two layers, including the input layer and the output layer. However, because no calculation is done in the input layer, it

is not taken into account. The neurons in the output layer compute the output signals after receiving input signals via the weights.

✓ A multi-layer feed-forward neural network structure contains a layer of "hidden neurons" between the input and output layers, as opposed to a single-layer network. According to Haykin [38], hidden neurons are supposed to serve as a useful intermediary between the network's output and its external input .Higher-order statistics can be extracted by the network if there are one or more hidden layers.

In general, the multilayer neural network is the most common type of FFNN used, where an ANN is made up of many neurons that are stacked one on top of another in multiple layers. A typical multilayer feed-forward neural network consists of three layers: the input layer, the hidden layers, and the output layer. The basic architecture of the ANN model is depicted in figure II.5.

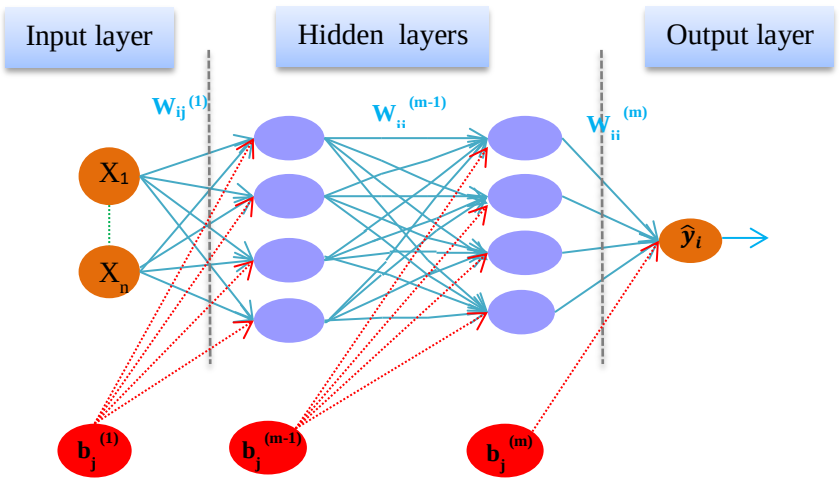

Figure.II.5 Multilayer neural network architecture.

Whereas the number of input parameters defines how many neurons are in the input layer, and the total number of output parameters determines how many neurons are in the output layer. The number of neurons in the hidden layers is often determined using Equation (II.3) [39]:

$$number\ of\ hidden\ neurons = \frac{number\ of\ inputs + output\ neurons}{2} + \sqrt{number\ of\ data\ points} \qquad \text{(II.3)}$$

This relies on the quantity of input and output neurons as well as training data points [39]. Additionally, determining the number of hidden neurons is often done by trial and error [40]. The input layer of a neural network comprises the specific attributes or traits that the network is taught to identify. The hidden layers are responsible for executing the necessary calculations by using nonlinear transformations on the input neurons. The output layer generates the outcome by accepting input from the last hidden layer. Fractional numbers that connect these layers to one another are called weights and biases; they are the main parameters of NN. Weights are changed during training to forecast accurate results, and weight values become constant when the error reaches a certain threshold value.

An ANN's mathematical structure can be generally described as:

$$y_i = \sum_{j=1}^{m} X_i W_{ij} + b_j \quad \text{(II.4)}$$

Where X_i are the inputs to the node j or output of earlier layers, W_{ij} are weights that indicate the strength of the connection between the ith and jth nodes, and bj is the bias. y_i is the net input to the node j in the hidden or output layer. However, The sum of weighted inputs must be passed through an activation function represented by $\boldsymbol{\varphi}(y_i)$ in order to get the output and pass it as input to the following layer.

II.5.2.2.Recurent neural network

Models of artificial neural networks (ANNs) with directed cycles formed by unit connections are called recurrent neural networks (RNNs), as shown in the figure. II.6. In particular, a directed cycle is a series that has the appearance of a certain order since the walk over the vertices and edges is entirely controlled by the set of edges employed. RNNs are frequently employed expressly for handwriting and speech recognition [41].

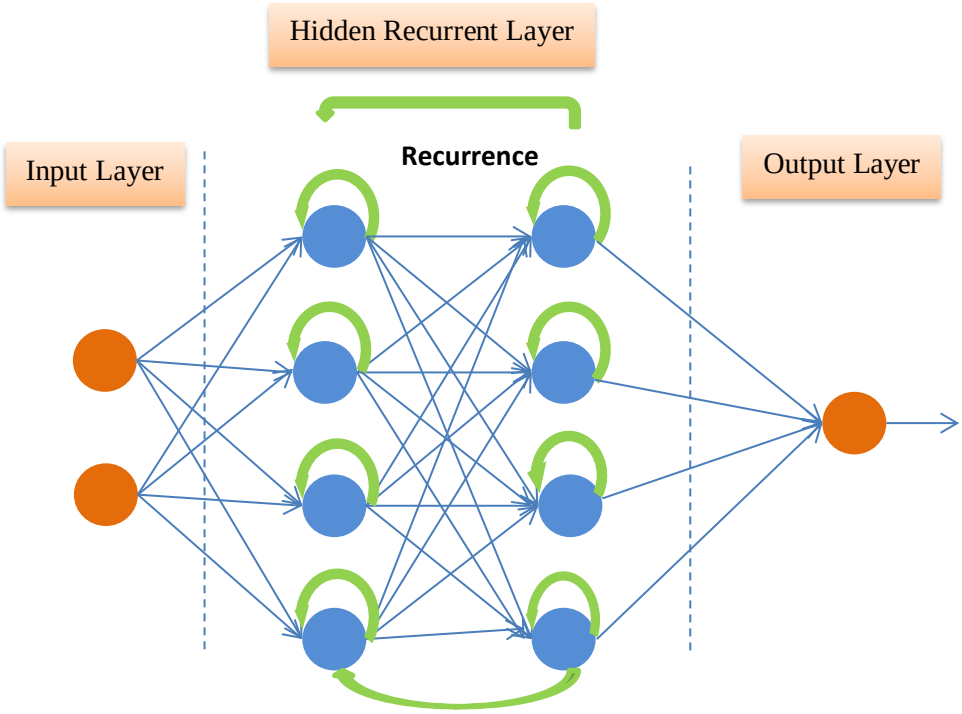

Figure.II.6 Recurrent neural network architecture [42].

II.5.2.3. Convolutional neural network

Convolutional Neural Networks are Deep Learning algorithms that can recognize distinct objects or features in a picture and distinguish between them by applying weights and biases that can be learned. A ConvNet requires a lot less pre-processing than other classification techniques. While filters are manually designed in more archaic techniques, ConvNets may learn these properties and filters with sufficient training [41].

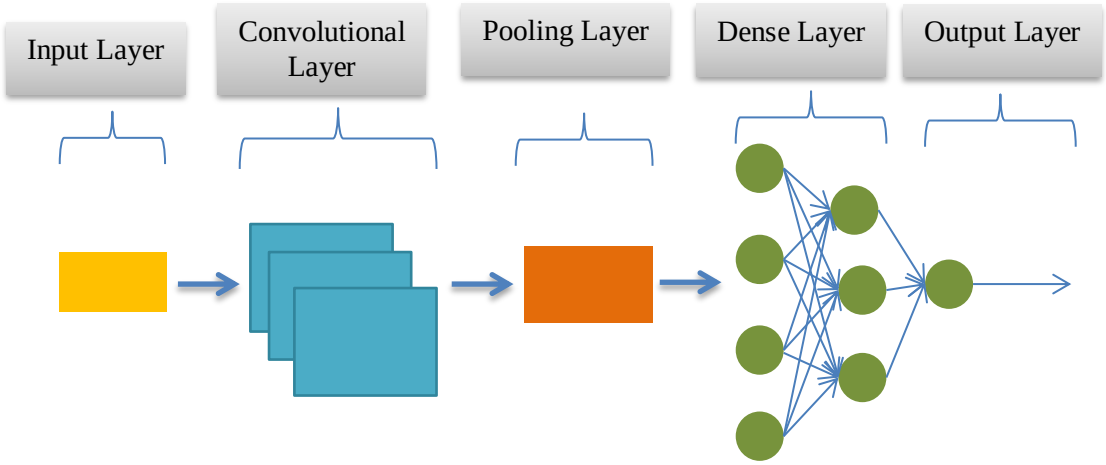

Figure.II.7 Convolutional Neural Networks architecture[43].

II.5.3. The ANN Process

After the neural network's structure has been established, the imported data input and output are used to train the network. The complete input-output parameter data sets are split into two groups: the training data set, which contains a higher percentage of data points and is used to train the neural network, and the validation data set, which contains the remaining data points and is used to validate the trained neural network. Neural networks import input-output parameters together with their training data points. This network is trained until it achieves the permitted error. Once determining the acceptable error, the trained network is validated by importing the validation data set's input parameter values and predicting the output that corresponds to the parameter values. In order to determine which neural network is best for the prediction, the trained neural network is compared to the corresponding actual values of the validation data set's output parameters. If the difference between the actual and predicted values (the objective function) reaches the smallest possible error, the trained neural network is the one that will be chosen to use. This process is known as supervised learning or training. During the training phase, the ideal weights should be chosen in order to predict the output with the lowest possible margin of error. The number of hidden layers, the number of neurons in the hidden layers, and the algorithm used in the hidden layer for appropriate weight calculation all affect how well these networks function. An artificial neural network's weight determination procedure is regarded as an optimization problem. Different algorithms can be used to achieve this. Gradient descent is the method most frequently used to optimize the weights and biases in an ANN. Nonetheless, the

problem of being trapped in local minima is a recognized drawback of this method [44]. In this context, several researchers have proposed other optimization algorithms to overcome this issue [45, 46].

II.6.Fuzzy logic

II.6.1.An Overview of Fuzzy Logic

We use fuzzy logic and fuzzy ideas so frequently in our daily lives that no one ever notices. For example, 'not very satisfied' or 'quite satisfied' are frequently used as vague or ambiguous answers in specific survey questions, or what specific level of satisfaction or dissatisfaction is expressed with a certain service or product?. Only humans are capable of creating and carrying out these evasive answers. Could a machine provide the same straightforward answers to the survey questions as a person can? It is not even conceivable. Computers can only comprehend binary numbers, such as '0' or '1', as well as high or low. All machines can process this data, also known as crisp or classic data. So, is it possible to provide machines with human assistance to manage such ambiguous data? If so, how are robots and computers able to process that ambiguous data? Yes, this is the response to the first query. On the other hand, some fuzzy logic strategies and familiarity with the fuzzy inference system are required in order to respond to the second question [47].

Professor Zadeh of the University of California, Berkeley, developed the concept of fuzzy logic in 1965 [48]. Since the 1980s, an increasing number of fuzzy implementations, including those in banks, hospitals, libraries, automated control, car manufacture, industrial manufacturing, and academic education, have been documented [47]. In today's world, fuzzy logic approaches are widely used in every facet of life.

II.6.2. Fuzzy sets

Fuzzy sets are a notion that was developed to deal with specific situations [49]. The basis of fuzzy set theory is the idea of partial membership, which states that each element gradually or partially becomes a part of the defined fuzzy sets. Instead of being "crisp," the contours of each fuzzy set are "fuzzy" or "gradual," as illustrated in figure II.8:

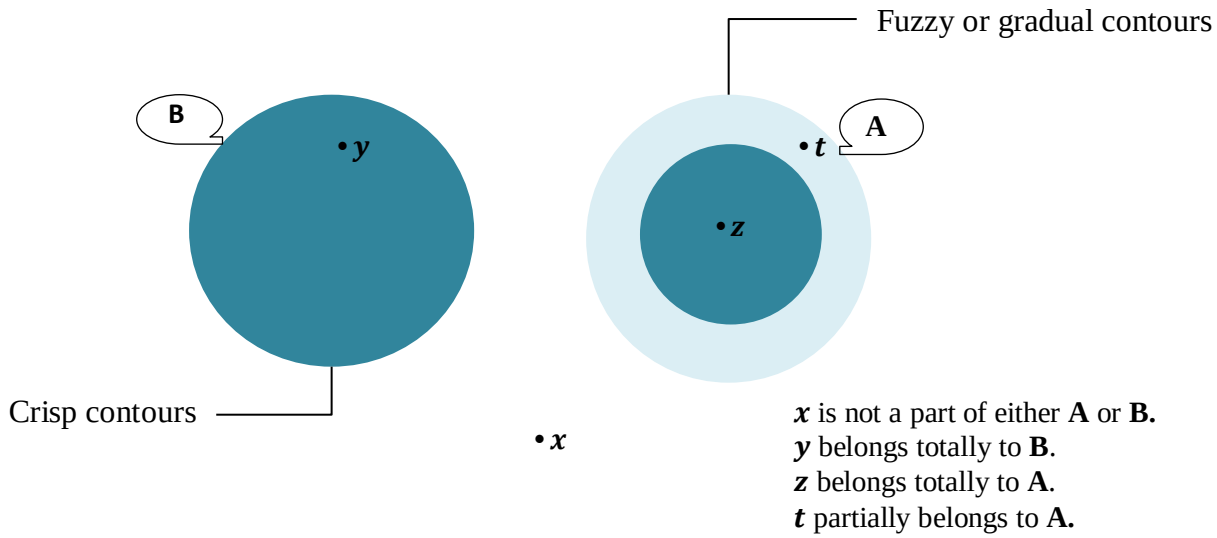

Figure.II.8 A graphic representation of a fuzzy set and a classical set [49].

In a crisp set, the membership status of an element in the set is determined by a characteristic function, denoted by if for membership and if for non-membership. Fuzzy set theory expands on this notion by introducing the concept of partial membership. A fuzzy set, defined over a universe of discourse, is identified by a membership function, which assigns values within the range of [0-1] [50]. Commonsense word labels such as heavy, tiny, huge, quick, slow, medium, high, tall, and so forth are represented by fuzzy sets, where it is possible for one element to belong to many fuzzy sets simultaneously. A collection of ordered pairs can be used to represent a fuzzy set. A generic element (x) and its grade of membership function make up each pair [48-53].That is;

$$A = \{(x, u_A(x)) | x \in U\} \quad \text{(II.5)}$$

Where x is called a support value if $u_A(x) > 0$.

Depending on the type of membership function, different types of fuzzy sets will be obtained. A membership function can be defined as a curve that determines the mapping of each point in the input space to a membership value, also known as the degree of membership, which ranges from 0 to 1. Membership functions may be divided into two categories: Gaussian-shaped functions and functions composed of straight lines. Figure.II.9 displays the membership functions that are most commonly used [54].

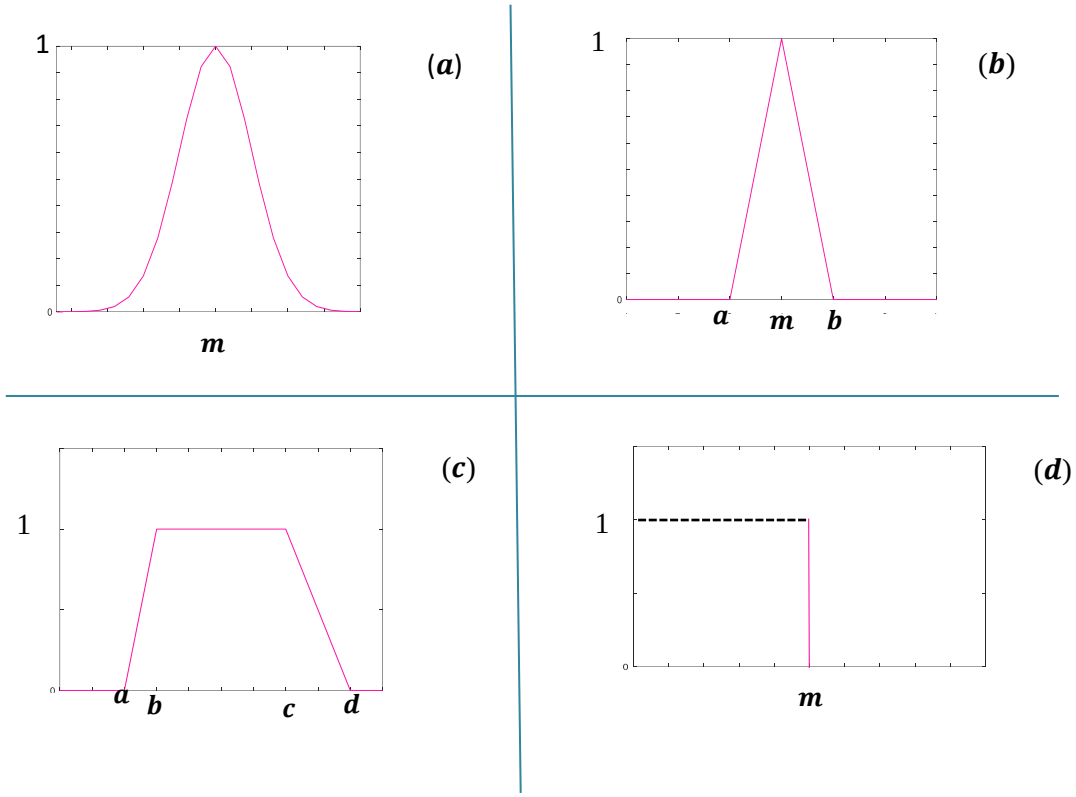

Figure.II.9 The membership functions: (a) Gaussian,(b) triangular,(c) trapezoidal and (d) singleton.

✓ **The Gaussian membership function** :In general, the Gaussian membership function is displayed as Gaussian (x:m,sd), where m and sd refer to the mean and standard deviation, respectively (see figure II.9(a)).

✓ **The triangular membership functions**: this curve is formed by a combination of two linear functions. And characterized by three parameters that involve angles a, b, and m from $u_A(x)$. These angles are denoted by a lower limit (a), an upper limit (b), and a midpoint (m). with the constraint that a < m < b. Due to its high computational efficiency value, this curve is frequently utilized in implementation (see figure II.9 (b)).

✓ **The trapezoidal membership functions** curve shares similarities with the triangle curve. The triangular curve has a single point, or value, where the degree of membership equals 1, whereas the trapezoid curve consists of multiple values or points with a membership value of 1(see figure II.9 (c))..

✓ **The singleton membership functions** this membership function takes the value of 0, in all of the discourse with the exception of the point m, where it takes the value of 1 (see figure II.9(d)).

II.6.3.Basic Operations on Fuzzy Sets

Let A and B be two fuzzy sets with their membership functions $u_A(x)$and $u_B(x)$), respectively, on the universe of discourse X, As in classical set theory,the union, intersection, and complement of fuzzy sets are defined. These relationships are translated using the operators "AND," "OR," and "NOT". New membership functions related to these operators are created as follow [49, 55]:

- x belongs to A or $B \leftrightarrow x \in A \cup B \leftrightarrow u_{A\cup B}(x)$
- x belongs to A and $B \leftrightarrow x \in A \cap B \leftrightarrow u_{A\cap B}(x)$
- x belongs to belongs to the complement of $A \leftrightarrow x \in \bar{A} \leftrightarrow u_{\bar{A}}(x)$

The operator “AND” is defined by a triangular norm (t − norm) D:

$$D = [0,1] \times [0,1] \rightarrow [0,1]$$

$$(x, y) \rightarrow z = xDy \tag{II.6}$$

In the same way, the operator “OR” is defined by a triangular co-norm (t − norm) D:

$$S = [0,1] \times [0,1] \rightarrow [0,1]$$

$$(x, y) \rightarrow z = xSy \tag{II.7}$$

To delineate fuzzy logic operators, we must identify the appropriate operators that maintain the outcomes of applying the AND, OR, and NOT operators. The most commonly used operators that provide the solution are Min, Max, Sum, and the complement. They can be defined as follows:

❖ The operator "AND " can established by :

$$u_{A\cap B}(x) = min\big(u_A(x), u_B(x)\big)$$
$$u_{A\cap B}(x) = u_A(x) + u_B(x) \tag{II.8}$$

❖ The operator OR can established by :

$$u_{A\cap B}(x) = max\big(u_A(x), u_B(x)\big)$$

$$u_{A\cap B}(x) = u_A(x) + u_B(x) \quad \text{(II.9)}$$

❖ The operator "NOT " can established by :

$$u_{\bar{A}}(x) = 1 - u_A(x) \quad \text{(II.10)}$$

II.6.4.Linguistic variables

A linguistic variable is one whose values are words or phrases in either an artificial or real language [49]. An example of a linguistic variable is age, which takes on language values instead of numerical ones. Examples of such values include young, not young, very young, fairly young, old, not very old, not very young, and others. A linguistic variable can be x in the universe discourse U can be established by $T(x) = \{T_x^1, T_x^2, T_x^3, \dots \dots T_x^k\}$ and $u(x) = \{u_x^1, u_x^2, u_x^3, \dots \dots u_x^k\}$ .Where $T(x)$ is the set of linguistic values of the variable x ,also called a linguistic term or linguistic label, with each T_x^k fuzzy number with membership function u_x^k in the universe discourse U[50].

II.7.Fuzzy Inference system

The aim of a fuzzy inference system (FIS) is to transform input data into output data by evaluating a set of rules. The expert's expertise typically defines the set of rules, while the inputs originate from the fuzzification process. Figure 6 shows a general model of a fuzzy inference system (FIS). Crisp inputs are converted into crisp outputs via the FLS [56, 57]. The fuzzifier, inference engine, rule base, and defuzzifier are three stages of the FIS, which are illustrated in Figure II.10.

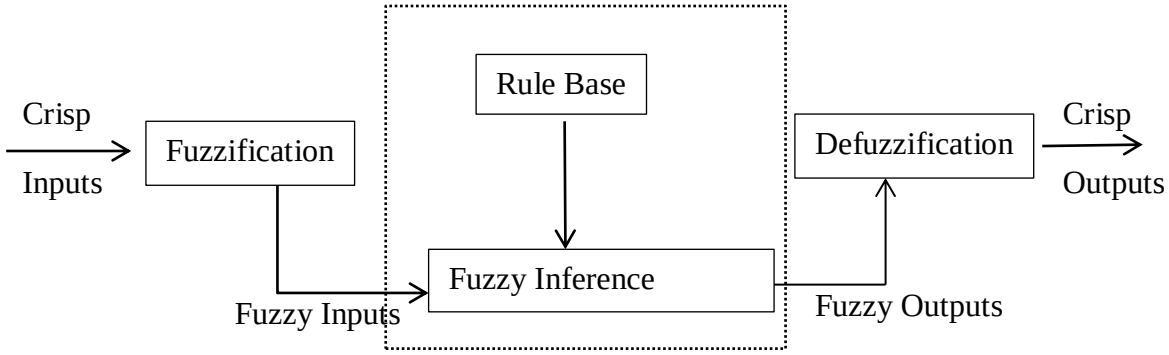

Figure.II.10 Fuzzy interference system diagram[58].

- The first step is fuzzification, which involves characterizing the linguistic variables used in the system. This entails transforming the real inputs into a fuzzy part, which is defined on a representation space linked to the input. This representation space is normally a fuzzy subset. During the fuzzification stage, each input and output variable is linked to fuzzy subsets.
- The second step is the inference engine, which is a mechanism for condensing system information through a set of rules defined for any problem representation. Each rule delivers a partial conclusion, which is then aggregated with the other rules to provide a conclusion (aggregation).
- The third step is defuzzification, which is the reverse of fuzzification and converts the fuzzy inference output into a non-fuzzy value as the SIF's final response.

II.7.1. Takagi-Sugeno-Kang fuzzy Inferences system

The Takagi-Sugeno-Kang fuzzy rule-based model may represent complicated, highly non-linear processes using a minimal number of rules. Compared to the use of the Mamdani fuzzy model type, a much smaller number of rules may be required [59, 60].Additionally, a drawback of the TSK fuzzy model is its less comprehensible presentation. The basic idea of the TSK model is that an arbitrarily complex system is comprised of a number of subsystems that are mutually interconnected. If K areas in the state space under examination are identified as separate sub-systems, then the behaviour of the system in these regions may be explained using more straightforward functional relationships [61].

A Type-1 Takagi-Sugeno-Kang (TSK) fuzzy system is described by a set of IF-THEN rules representing the input-output relationships of a system. The first-order TSK fuzzy system is the most commonly used.

It has a base of N rules, each having k antecedents, where the l[th] rule is expressed as follows[49]:

$$R^l: IF\ x_1\ is\ F_1^l\ and x_2\ is\ F_2^l and, \dots \dots x_k\ is\ F_k^l$$

$$THEN\ y^l = c_0^l + c_1^l x_1 + c_2^l x_2 + \cdots + c_k^l x_k \qquad \text{(II.11)}$$

for l = 1,2,…..N c_j^l represents the parameters of the consequence, for j = 0,1,….. k; x_k is the input of the fuzzy system; y^l is the output of the lth rule; and F_k^l is the fuzzy set, for j = 0,1,… k. The final output of the incomplete first-order TSK model is deduced as follows [62]:

$$r = \sum_{l=1}^{N} f^l y^l \quad \text{(II.12)}$$

where f^l are the degrees of activation of the rules defined as follows :

$$f^l = T_{j=1}^{p} \mu_{F_j^l}(x_j) \quad \text{(II.13)}$$

Where T denotes the t-norm. When the membership functions are Gaussian (cf., Figure II.9(a)):

$$\mu_{F_j^l}(x_j) = \exp\left[-\frac{1}{2}\left(\frac{x_j - m_j^l}{\delta_j^l}\right)^2\right] \quad \text{(II.14)}$$

and the t-norm is used, Equation (II.12)can be expressed as follows:

$$r = \sum_{l=1}^{N} y^l \prod_{j=1}^{k} \exp\left[-\frac{1}{2}\left(\frac{x_j - m_j^l}{\delta_j^l}\right)^2\right] \quad \text{(II.15)}$$

where m_j^l and δ_j^l are the centers (means) and standard deviations of the l[th] fuzzy set, respectively.

II.8. Optimization

II.8.1.Optimization problems

In numerous fields, such as computer science, engineering, business, and others [63], decision-making focuses on optimization problems in order to find feasible solutions for a collection of parameters that maximize or minimize objective functions under given restrictions. The feasible solutions with objective function values that are superior to those of any other feasible solution are the optimal ones [64]. Generally, we can formulate optimization problems by combining both maximization and minimization problems. We can use the negative of the objective function to transform any maximizing problem into a minimization problem, and vice versa. They can be defined as follows [63, 65]:

$$\min_x fi(x) \leftrightarrow \max_x [-fi(x)] \quad \text{(II.16)}$$

$$\max_x fi(x) \leftrightarrow \min_x [-fi(x)] \quad \text{(II.17)}$$

Where fi(x) for (i = 1, 2... M), are objective functions, or simply cost functions, when M = 1, there is a single objective function. The components x_i of x are called design or decision variables. The area covered by the selection variables is called the design space, also known as the search space.

II.8.1.1.Types of optimization problems

Determining the ideal solution to an optimization problem can be a difficult undertaking, as it depends on selecting and using the right algorithms. The type of problem, the available algorithms, processing resources, and time limitations can all influence the algorithm that is selected [66]. We list the following instances of optimization problems [64, 67]:

- Combinatorial optimization problems**:** There are many optimization problems in which the independent variables are restricted to a set of discrete values. These problems are called combinatorial optimization problems.
- General unconstrained problems: where a nonlinear function is defined on an unconstrained set of real values.
- General problems with constraints: where a nonlinear function is defined over a limited set of real values. In general, optimization problems are constrained optimization problems.

II.9.Optimization methods

II.9.1.Gradient Descent

The gradient descent (GD) is one of the most basic methods for unconstrained optimization that may be defined as [68]:

$$x(t+1) = x(t) - \delta \nabla f(x(t)) \quad \text{(II.18)}$$

Where δ is an arbitrary ratio, sometimes known as the step size, and $x(t)$ is a candidate solution at time step t. Additionally, $\nabla f(x(t))$ represents the gradient of the objective function f at the point represented by $x(t)$.

The procedure functions based on the principle that the value of f(x) rises more quickly in the direction of the gradient, ∇f(x), for any given value of x. On the other hand, the direction of $-\nabla$f(x) yields the fastest reduction in the value of f(x); hence, gradient descent is also known as the steepest descent technique occasionally.

It follows that for rising values of t, iterating equation (II.18) with small enough values for δ will result in $x(t) \approx x_B^*$ given an initial location x(0) ∈ B ⊆ A.

As an unconstrained optimization technique, gradient descent (see algorithm II.1) can be employed if the gradient of f is known or can be estimated. Due to its simplicity, the technique is widely used in optimization, for example, in neural network training [69,70].

Algorithm II.1. The Gradient Descent pseudo code

1:Initialise $\boldsymbol{x(0) \in A}$

2:Repeat

$$\boldsymbol{x(t+1) = x(t) - \delta \nabla f(x(t))}$$

4:Until stopping condition is true

II.9.2.Evolutionary Algorithms

The last 20 years have seen an increase in interest in algorithms founded on the idea of natural evolution. The EAs—also known as meta-heuristic techniques—are population-based random search methods that are directed by a few heuristics [71]. All EAs typically operate in this way: first, a population of individuals is established, each of whom represents a potential solution to the issue at hand. A fitness function is used to assess each solution's quality. Every time an EA iterates, a selection procedure is used to create a new population. To guarantee that the fittest individuals will comprise the future population, the selection process is biassed towards them. Both unary transformation (mutation) and higher-order transformation (crossover) are used to change individuals. Until convergence is achieved, this process is repeated. It is anticipated that the best solution discovered will be almost optimal [64]. Algorithm II.2 shows the standard pseudo-code for an EA [72]. As a result, the two most commonly used operators in EAs are mutation, which creates a new individual by making a minor random change in an existing one, and recombination (crossover), which creates new people by combining pieces from two or more existing individuals [64].

Algorithm II.2. General pseudo code of Evalutionary Algorithm

1:Set the population to start.

2:Repeat

To create a new population, apply selection to the existing population.

Change the population's individuals by using evolutionary operators.

Determine each individual in the population's degree of fitness.

3: Until A few convergence requirements are achieved.

In the 1970s, the first works on artificial evolution focused on genetic algorithms (GA), evolution strategies (ES), and evolutionary programming (EP). These three types of algorithms were inspired by neo-Darwinism's common principles, which included using a population of individuals, evaluating individuals based on a function, selecting the best individuals, and generating a new population through crossover and mutation operations. Then, in the 1990s, genetic programming (GP) emerged, introducing tree representations [73, 49].

II.9.3.Swarm algorithms

Some researchers separate swarm intelligence from evolutionary computing. Swarm intelligence draws inspiration from the collective behavior observed in social animals and insects [49]. Some researchers emphasize that the first algorithms of swarm intelligence, ant colony optimization and particle swarm optimization, should not be considered evolutionary algorithms. However, other authors consider swarm intelligence to be a subset of evolutionary computing. For instance, one of the researchers refers to particle swarm optimization as an evolutionary algorithm [74].

In a swarm, numerous agents operate, with each agent exhibiting simple, localized, and stochastic behavior. While there's no centralized structure governing the rules of agent behavior in a single swarm, interactions between agents give rise to emergent global behaviors and intelligence. The individual actions of each agent and the interactions within the swarm can be likened to exploitation and exploration, respectively. These aspects are integrated into the agents' movements within a straightforward framework [75].

Several algorithms have been developed in the field of swarm intelligence, such as ant colony algorithms, the original ant system (Ant System), particle swarms, bee-inspired algorithms, artificial fish swarm algorithms, bat algorithms, firefly swarm algorithms, and bacterial foraging optimization algorithms. In this thesis, we will only focus on particle swarms, Imperialist competitiveness, and the Invasive Weed Optimization algorithms.

II.9.3.1.The Particle swarm optimization

Particle Swarm Optimization (PSO) has been applied extensively to optimization problems. It is a metaheuristic algorithm inspired by the collective movements observed in certain social animals, such as fish and migratory birds, which tend to imitate the successful behaviours they observe in their surroundings while adding their own personal variations [76, 77]. Kennedy and Eberhart developed PSO in 1995 [78], drawing its origins from the work of C. Reynolds [79] and Heppner and Grenander [80], who created mathematical models to replicate fish schools and group bird flights. In PSO, each particle moves within the population's search space and represents a potential solution. In order to find the best solution, both the particles and their neighbors must communicate with one another and modify their locations accordingly. Each particle modifies its position based on its personal experience as well as the collective experience of the swarm through several rounds of iteration; the algorithm finds the best solution.

There are three components that affect a particle's movement (as illustrated in figure II.11) [49]. The first component is physical, moving in the direction it is currently traveling; the second is cognitive, moving towards the best site it has already passed; and the third is social, depending on the experience of its neighbors to move towards the best site they have reached. The neighborhood can be defined sociometrically by taking into account the individual's position within the swarm, or spatially by calculating the Euclidean distance between the locations of two particles

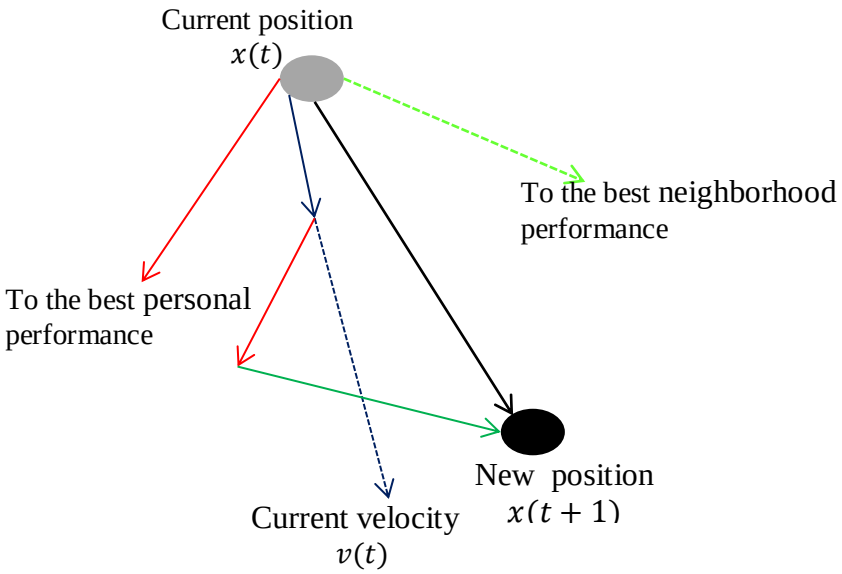

Figure.II.11 Movement of particle [49].

Additionally, each particle i in the d-dimensional search space has five properties: the position x_{id}(t), the velocity $v_{id}(t)$, objective function $J^{x_t d(t)}$ and the personal and the objective function of the best position this particle was experienced $P_{id(t)}, J^{P_{id}(t)}$, respectively. Initially, a random float number generator constrained by the lower and upper bounds of the d-dimensional search space is used to build a velocity vector $v_{id}(t)$. For each particle i , each particle's velocity is adjusted iteratively based on its neighborhood's best position as well as its own personal best position. Finally, the velocity causes each particle to shift positions. Equations (II.19) and (II.20) are used to calculate the updated velocity and position of particle:

$$v_{id}(t+1) = w.v_{id}(t) + c_1.r_1.\left(P_{id(t)} - x_{id}(t)\right) + c_2.r_2.\left(P_{gd(t)} - x_{id}(t)\right) \quad \text{(II.19)}$$

$$x_{id}(t+1) = v_{id}(t+1) + x_{id}(t); \quad \text{(II.20)}$$

Where $r_1$and $r_1$and are uniformly random numbers in the interval [0, 1], and $c_1$and $c_2$represent the personal and global learning coefficients, respectively. $P_{gd}(t)$ is the best location that a particle has attained in the neighborhood of the i[th] particle, as well as w its inertia weight. A new objective function is calculated once the particle positions and velocities are updated, and the procedure is continued until the specified stopping condition is satisfied [81]. The pseudo-code of PSO is provided in Algorithm II.3.

Algorithm II.3: Particle Swarm Optimization

1 Initialize **parameters & populations**
- Initialize the positions and velocities Randomly for each particle

2 Calculate the fitness value of particle
3 Calculate the personal best and global best
4 while Termination Criteria **do**
for each particle **do**
-Update velocity by Eq. (II.19)
-Update particle position by Eq. (II.20)
- Evaluate fitness.
- Update personal best.
- Update global best.
End for
return global best

II.9.3.2.Imperialist Competitive Algorithm (ICA)

The Invasive Imperialist Competitive Algorithm (ICA) is a revolutionary swarm intelligence technique inspired by imperialist rivalry. It was developed by Atashpaz-Gargari [82]. Imperialism relied significantly on mercantilism, an economic ideology that encouraged governments to extend their power and control outside of their boundaries [83]. ICA is a population-based algorithm where every individual splits into two groups: imperialists and colonies, according to their power, where the term "country" is given to each member of the population who represents a potential solution [84]. The main ICA phases can be summarised as follows:

- Generating initial empires: At first, in the N-dimensional search space, ICA begins with an initial population selected randomly, known as countries, as shown in Eqs. (II.21) and (II.22), which are then used to determine each fitness value.

$$\mathbf{p} = [p_1, p_2, p_3, \dots, p_N] \in \mathbb{R}^N ; \quad \text{(II.21)}$$

$$f(\mathbf{p}) = f(p_1, p_2, p_3, \dots, p_N) \in \mathbb{R} ; \quad \text{(II.22)}$$

Where f represents the objective function that has to be either minimized or maximized.

After that, the most powerful countries with the highest fitness values are selected as imperialists, and the remaining countries will be colonies. The power of the imperialists determines how colonies are divided among empires. As shown in the equation.

$$p_n = \frac{|NC_n|}{\left|\sum_{i=1}^{N_{imp}} NC_i\right|}; \quad \text{(II.23)}$$

The total number of imperialists is denoted by N_{imp}, and the normalized cost of the ith imperialist is represented byNC_i, which is defined as :

$$NC_n = C_n - \max_{i=1,2,\dots N_{imp}} \{C_i\} \quad \text{(II.24)}$$

Where the imperialism cost is denoted byC_i. Depending on the power of the imperialists, each empire has a variable number of colonies; the initial number of colonies is found in Eq(17):

$$NOC_n = \lfloor p_n . N_{col} \rfloor; \quad \text{(II.25)}$$

The overall number of colonies is denoted by N_{col}.

✓ The assimilation process: After an empire is formed, its colonies advance toward the corresponding imperialists in an effort to increase the empire's power. Assimilation is the term for this process. A colony becomes a new imperialist, and the former imperialist becomes a colony if it exceeds the imperialist in position. All colonies get closer to their imperialists using Eq(II.25):

$$\mathbf{x}(t+1) = \mathbf{x}(t) + U(0, \beta \times d) \times \mathbf{V} \quad \text{(II.25)}$$

Where $\mathbf{V} = \boldsymbol{X}(t) - \mathbf{x}(t)$ is the direction vector from the colon $\mathbf{x}$ (t) to its imperialist $\boldsymbol{X}$ (t), d is the distance between the colony and the imperialist, and β is the assimilation coefficient.

✓ The revolution phase: After assimilation, comprehensive algorithmic exploration moves on to the revolution phase. The following equation's defined θ value shows that a colony's movement deviates from its imperialist's due to the revolution:

$$\theta \sim U(-\gamma, \gamma); \quad \text{(II.26)}$$

Where γ is a randomly distributed value between [-γ, γ], and the parameter θ meets the deviation from the imperialist direction of the colony.

✓ The intra-competition phase: The Imperialist Competitive Algorithm (ICA) passes to an intra-competition phase after the revolution stage, where an empire starts to compete with itself. Their position vectors are switched if a colony's fitness value is higher than its imperialist's. The colony therefore gains the imperialist's place, and the imperialist turns into a colony of this new imperialist. The intra-competition phase of the ICA is the name given to this procedure.

The empires compete with one another using the normalized total costs given in Eq. (II.22)

$$NTC_n = TC_n - \max_{i=1,2,\ldots,N_{imp}} \{TC_i\}; \quad \text{(II.27)}$$

The following formula is used to determine the overall power of the n^th empire:

$$TC_n = f(\mathbf{X}_n) + \frac{\xi}{NOC_n} . \sum_{i=1}^{NOC_n} f(\mathbf{x}_i). \quad \text{(II.28)}$$

By capturing the colonies of other empires, the empires want to increase the area under their colonial control. Equation (II.29) is used to determine the probability that the imperialist $\mathbf{X}_n$ will acquire colonies.

$$p_{\mathbf{X}_n} = \left| \frac{NTC_n}{\sum_{i=1}^{N_{imp}} NTC_i} \right|. \quad \text{(II.29)}$$

A powerful imperialist has a decent chance of emerging victorious and wresting the colony from the weaker empire. Empires without colonies are destined to collapse. The term "elimination of the weakest empires" describes this stage. Apart from the most powerful empire, all of the others eventually collapse. In this instance, each colony will have the same location and expense [85]. As a result, there won't be any qualitative distinction between colonists and imperialists, and the algorithm finally ended when the halting conditions or the maximum number of iterations were achieved. The pseudo-code of ICA is shown in Algorithm II.4.

Algorithm II.4. Imperialist Competitive Algorithm

1 procedure ICA "Initialization of the algorithm"

2 Use Eq. (II.21) to specify the random positions in the space.

3 **Choose the empires and their colonies.**

4 iteration ← 1

5 repeat "Assimilation Phase"

6 For the colonies that are growing up around the imperialist, use Eq. (II.25) to shift it in the direction of the applicable imperialist.

7 Move the colonies of an empire toward the imperialist using Eq.(II.26) . "Revolution Phase"

8 Calculate the total cost of each empire "Intra Competition Phase"

9 Exchange the better colony and the imperialist, if the colony has better cost value than imperialist >Inter-Competition Phase

10 The empires compete according to Eq. (II.27), which gives them their normalized costs.

11 **Determine the weakest colony in the weakest empire and give it to the best empire**

12 **Eliminate the weak empires**

13 **iteration ← iteration +1**

14 Until The maximum number of iteration is reached or the stopping criteria is reached.

End procedure

II.9.3.3.The Invasive Weed Optimization Algorithm

The Invasive Weed Optimization Algorithm (IWO) is a population-based metaheuristics algorithm that was inspired by invasive weeds and identifies favorable settings for growth and reproduction [86]. It was developed by Mehrabian and Lucas in 2006 [87]. This method mimics the adaptability, resilience, and unpredictability of weeds that are invading an area. It is an effective algorithm based on agriculture that is utilized to resolve many optimization issues in the real world [88]. In order to efficiently converge to ideal solutions, the IWO algorithm makes use of two fundamental features: the production of seeds and weeds in a weed colony [89]. The IWO algorithm employs numerous critical processes to identify and enhance the solution [90], and it's crucial to elucidate these processes: After the "population initialization" process, in which many seeds are dispersed partially around the search space,. Every weed has the ability to generate seed, which is utilized as a testing ground for optimization. The relative fitness of each seed ''the highest and lowest fitness levels' 'affects the colony's ability to generate seeds. The most fit weed in the colony is permitted to generate the greatest number of seeds. The colony's seed output increases linearly from a low to a maximal seed number. Weeds are produced in the colony using the following formula:

$$weed_n = \frac{f - f_{min}}{f_{max} - f_{min}}(S_{max} - S_{min}) + S_{\min}\ ; \tag{II.30}$$

where f represents the weed's fitness in its current location. The fitness of the colony varies from f_{min}, the lowest, to f_{max}, the highest. The minimum and maximum number of seeds that can be produced in the colony are denoted by S_{min} and S_{max}.

The third stage (Spatial dispersion), seeds are created at random and distributed throughout the search space dimensions. By a different variation (variable) and a normal distribution at a rate of zero. In this phase, the seeds are dispersed randomly around the mother plant. However, at each step, the random function's standard deviation will decrease from a predetermined starting value ($\sigma_{initial}$) to a final value (σ_{final}) as follows:

$$\sigma_{iter} = \frac{(iter_{max} - iter)^n}{iter_{max}{}^n}\left(\sigma_{initial} - \sigma_{final}\right) + \sigma_{final}; \tag{II.31}$$

where n is the nonlinear modulation index, σ_{iter} is the current standard deviation, and $iter_{max}$ is the maximum number of iterations.

We use the following formula to determine where the new seeds are situated.

$$X_{seed} = X_{parent} + sd = X_{parent} + randn(0,1) * \sigma_{iter} \quad \text{(II.32)}$$

Where X_{seeds} is the location of the seeds. X_{parent} is the parents site, and $randn$ is the generate random numbers by normal distribution (0,1) and record.

Finally, for the optimization problem, the freshly generated seeds and their parent plants are both part of the solution set. There is a defined maximum number of weeds that can be present in the colony or search area. If the weed population grows above this level, a competitive exclusion strategy is employed to remove the weeds with lower fitness values, leaving only fitter plants in the colony. The most seeds may be produced by weeds with higher fitness ratings. Once we complete the procedure in the maximum number of iterations, the minimal colony cost function of the grasses is saved [91, 92]. Algorithm II.5 displays ICA's pseudo-code.

Algorithm II.5.: Invasive weed optimization IWO

1 Initialize population of weeds, set parameters;
2 for 1 to Max_iteration **do**
- Calculate the population's fitness levels, from best to worst.
- Determine the standard deviation (std) for each iteration.

End for
3 for each weed w in the population W **do**
- Determine the number of seeds for w based on its level of fitness.
- Choose the seeds from the possible solutions near the parent weed w in a neighborhood that follows a normal distribution with a mean of 0 and a standard deviation of std.
- Add seeds produced to the population W

4 if (|W|>Max_SizePopulation) **do**
- Distribute the population W in accordance with their level of fitness.
- W=Select Better (weed, seed ,Max_SizePopulation)

End if
iteration= iteration+1;
End for

II.10.Conclusion

In this chapter, the background theory of artificial intelligence techniques is briefly reviewed. Additionally, the chapter dedicates itself to presenting the basic techniques and tools we will use later, including artificial neural networks, fuzzy logic type 1, and optimization

algorithms such as PSO, ICA, and IWO. We have explicitly presented the theoretical basis of these methods to set the stage for the algorithms proposed in the fourth chapter.

References

[1] W. Zhang, X. Gu, L. Tang, & al. (2022). Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge. Gondwana Research, 109, 1–17.

[2] C.Gao, X.Min, M.Fang & al .(2021). Innovative Materials Science via Machine Learning. Advanced Functional Materials, 32(1).

[3] J.Launchbury, (2017). A DARPA perspective on artificial intelligence. Retrieved November, 11(2019), 3.

[4] A.Dikshit, B.Pradhan & A. M. Alamri.(2021). Pathways and challenges of the application of artificial intelligence to geohazards modelling. Gondwana Research, 100, 290–301.

[5] Y.LeCun, Y. Bengio & G.Hinton. (2015). Deep learning. Nature, 521(7553), 436–444.

[6] R. S. Sutton & A. G. Barto, (2018). Reinforcement learning: An introduction. MIT press.

[7] S.Lloyd & C. Weedbrook, (2018). Quantum Generative Adversarial Learning. Physical Review Letters, 121(4).

[8] A. Roy, (2020). The Fourth Industrial Revolution: Klaus Schwab, New York, Currency Books, 2017, 192 pp., Journal of International Consumer Marketing, 32(3), 268–270.

[9] R.Vasudevan, G.Pilania, & P. V. Balachandran. (2021). Machine learning for materials design and discovery. Journal of Applied Physics, 129(7).

[10] J.Wei, X.Chu, X. Sun & al. (2019). Machine learning in materials science. InfoMat, 1(3), 338–358.

[11] L. E.Vivanco-Benavides, C. L Martínez-González, C.Mercado-Zúñiga & C. Torres-Torres,(2022). Machine learning and materials informatics approaches in the analysis of physical properties of carbon nanotubes: A review. Computational Materials Science, 201, 110939.

[12] M. J. Apter & P. McCorduck, (1982). Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence. Leonardo, 15(3), 242.

[13] A. B. Maksov, (2018). Artificial Intelligence in Materials Science: Applications of Machine Learning to Extraction of Physically Meaningful Information from Atomic Resolution Microscopy Imaging.disseration thesis . University of Tennessee, Knoxville.

[14] G. W. Evans, (2017). Artificial intelligence: where we came from, where we are now, and where we are going.

[15] H. Nguyen, (2019). Artificial intelligence and its impact on workforce. Thesis dissertation. Centria University of Applied Sciences

[16] N. Miailhe & C. Hodes, (2017). The third age of artificial intelligence. Field Actions Science Reports. The journal of field actions, (Special Issue 17), 6-11.

[17] C. M. Signorelli, (2018). Can Computers Become Conscious and Overcome Humans? Frontiers in Robotics and AI, 5.

[18] K. Aggarwal, M. M. Mijwil, A. H. Al-Mistarehi & al .(2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. Iraqi Journal for Computer Science and Mathematics, 3(1), 115-123.

[19] R. Lalitha (2021). AI vs. Machine Learning vs. Deep Learning. Artificial Intelligence (AI), 73–82.

[20] LIU, Boyan.2022 Artificial Intelligence and Machine Learning Capabilities and Application Programming Interfaces at Amazon, Google, and Microsoft. Doctorate dissertation . Massachusetts Institute of Technology.

[21] D. M. Powers & C. C. Turk, (2012). Machine learning of natural language. Springer Science & Business Media.

[22] J. Xiao, K. A. Ehinger, J. Hays & A. Oliva (2016). Sun database: Exploring a large collection of scene categories. International Journal of Computer Vision, 119, 3-22.

[23] T. O. Ayodele. (2010). Types of machine learning algorithms. New advances in machine learning, 3(19-48), 5-1.

[24] G. R. Schleder, A. C. M. Padilha, A. C. M. Acosta, & A. Fazzio. (2019). From DFT to machine learning: recent approaches to materials science–a review. Journal of Physics: Materials, 2(3), 032001.

[25] S. W.Knox , (2018). Machine Learning. Wiley Series in Probability and Statistics.

[26] B.BRENNINKMEIJER , A. DE VRIES, E. MARCHIORI et al. 2019. On the generation and evaluation of tabular data using GANs. Doctoral disertation., Radboud University.

[27] M. L. Hutchinson, E. Antono, B. M. Gibbons, S.Paradiso & B. Meredig. (2017). Overcoming data scarcity with transfer learning. arXiv preprint arXiv:1711.05099.

[28] F.KHMAISSIA,(2017). Data driven discovery of materials properties. Doctoral dissertation . University of Louisville.

[29] J.Mendes-Moreira, C. Soares, A. M. Jorge & J. F. D. Sousa .(2012). Ensemble approaches for regression: A survey. Acm computing surveys (csur), 45(1), 1-40.

[30] N. Kalchbrenner, E. Grefenstette & P. Blunsom .(2014). A Convolutional Neural Network for Modelling Sentences. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).

[31] V.S.S.KANDARPA. 2022. Tomographic image reconstruction with direct neural network approaches. Doctoral dissertation . Université de Bretagne occidentale-Brest.

[32] W. S. McCulloch & W. Pitts.(1943). A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, 5(4), 115–133.

[33] S. R. Dubey, S. Chakraborty, S. K. Roy & B. B. Chaudhuri. (2020). diffGrad: An Optimization Method for Convolutional Neural Networks. IEEE Transactions on Neural Networks and Learning Systems, 31(11), 4500–4511.

[34]M. ZAKHROUF , 2021Development of a Neuronal System for Modeling Streamflows of Some Watersheds in Algeria. Doctoral dissertation ,Aboubakr Belkaïd University of Tlemcen

[35] A. Linares-Rodriguez, J. A. Ruiz-Arias, D. Pozo-Vazquez & J. Tovar-Pescador. (2013). An artificial neural network ensemble model for estimating global solar radiation from Meteosat satellite images. Energy, 61, 636–645.

[36] H. Kabbay. (2022). Artificial Neural Network Concepts and Examples.Master thesis . University of Missouri-St.

[37] M. H. SAZLI .(2006). A brief review of feed-forward neural networks. Communications, Faculty Of Science, University of Ankara, 11–17.

[38] S. Haykin.(1998). Neural networks: a comprehensive foundation. Prentice Hall PTR.

[39] M. Mohanraj, S. Jayaraj & C. Muraleedharan. (2009). Performance prediction of a direct expansion solar assisted heat pump using artificial neural networks. Applied Energy, 86(9), 1442–1449.

[40] S. A. Kalogirou & M. Bojic. (2000). Artificial neural networks for the prediction of the energy consumption of a passive solar building. Energy, 25(5), 479-491.

[41] N.BENHALIMA, Nessrine. Classification of banknote fitness by artificial neural networks. 2022. Doctoral dissertation . Université Ibn Khaldoun-Tiaret-.

[42] https://www.excurl.com/2023/02/artificial-intelligence-neural-networks.html

[43] https://medium.com/swlh/an-overview-on-convolutional-neural-networks-ea48e76fb186

[44] I.Goodfellow , Y. Bengio & A. Courville. (2016). Deep learning. MIT press.

[45] A. Askarzadeh & A. Rezazadeh. (2013). Artificial neural network training using a new efficient optimization algorithm. Applied Soft Computing, 13(2), 1206–1213.

[46] M. G. M .Abdolrasol, S. M. S.Hussain, T. S. Ustun & A. Milad. (2021). Artificial Neural Networks Based Optimization Techniques: A Review. Electronics, 10(21), 2689.

[47] Y. Bai & D. Wang (n.d.). Fundamentals of Fuzzy Logic Control — Fuzzy Sets, Fuzzy Rules and Defuzzifications. Advanced Fuzzy Logic Technologies in Industrial Applications, 17–36.

[48] L. A. Zadeh. (1965). Fuzzy sets. Information and Control, 8(3), 338–353.

[49] M.A.ASSAM. 2018. Modélisation, Débruitage, Extraction des caractéristiques et Classification des signaux électrocardiogrammes ECG. Doctoral dissertation . University of Batna 2.

[50] Arun D. KULKARNI. (2001)Fuzzy logic fundamentals. Computer vision and fuzzy-neural systems (Prentice-Hall, Englewood Cliffs, NJ, 2001), , 61-101.

[51] J. R. Tong (1995). La logique floue. Hermes, Paris, 91.

[52] J. S. R. Jang ,C. T. Sun & E. Mizutani. (1997). Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. IEEE Transactions on automatic control, 42(10), 1482-1484.

[53] J.R.Timothy. (2009).Fuzzy logic with engineering applications. John Wiley & Sons.

[54] A. G. U. S. Setiawan, E. R. Arumi & P. R. I. S. T. I. Sukmasetya. (2020). Fuzzy membership functions analysis for usability evaluation of online credit hour form. Journal of Engineering Science and Technology, 15(5), 3189-3203.

[55] M. Mizumoto & K. Tanaka. (1981). Fuzzy sets and their operations. Information and Control, 48(1), 30–48.

[56] R. Janarthanan, R. Balamurali, A. Annapoorani & V. Vimala. (2021). Prediction of rainfall using fuzzy logic. Materials Today: Proceedings, 37, 959–963.

[57] D. Chowdhury & S. Hovda. (2023). A hybrid fuzzy logic/genetic algorithm model based on experimental data for estimation of cuttings concentration during drilling. Geoenergy Science and Engineering, 231, 212387.

[58] KUMAR, Pravesh et VERMA, Karun Guide. (2009). Use of Fuzzy Set and Neural Network to Extract Fingerprint Minutiae Points and Location. Doctoral dissertation .

[59] C. C. Lee. (1990). Fuzzy logic in control systems: fuzzy logic controller. II. IEEE Transactions on Systems, Man, and Cybernetics, 20(2), 419–435. https://doi.org/10.1109/21.52552

[60] F. L.Chung & J. C. Duan. (2000). On multistage fuzzy neural network modeling. IEEE Transactions on fuzzy systems, 8(2), 125-142.

[61] D. Kukolj. (2002). Design of adaptive Takagi–Sugeno–Kang fuzzy models. Applied Soft Computing, 2(2), 89–103.

[62] K.Tanaka & M. Sugeno. (1998). Introduction to Fuzzy Modeling. Fuzzy Systems, 63–89.

[63] J. W. Chinneck (2006). Practical optimization: a gentle introduction. Systems and Computer Engineering), Carleton University, Ottawa. http://www. sce. carleton. ca/faculty/chinneck/po. html, 11.

[64] M.GH.OMRAN . (2006). Particle swarm optimization methods for pattern recognition and image processing. Doctoral dissertation . University of Pretoria.

[65] X. S Yang. (2010). Engineering Optimization: An Introduction with Metaheuristic Applications, John Wiley and Sons, Hoboken, NJ, USA.

[66] A. H. Gandomi, X.-S.Yang, S. Talatahari & A. H. Alavi. (2013). Metaheuristic Algorithms in Modeling and Optimization. Metaheuristic Applications in Structures and Infrastructures, 1–24.

[67] P. Gray, W. Hart, L. Painton, C. Phillips & al. (1997).A Survey of Global Optimization Methods, Sandia National Laboratories.

[68] D. BARLA-SZABÓ, & al. (2010). A study of gradient based particle swarm optimisers. Doctoral dissertation . University of Pretoria.

[69] F. Van den Bergh. (2002).An Analysis of Particle Swarm Optimizers. Doctoral dissertation , University of Pretoria, Pretoria, South Arica.

[70] A.P.Engelbrecht.(2007). Computational Intelligence.

[71] D.Corne& M. A. Lones. (2018). Evolutionary Algorithms. Handbook of Heuristics, 1–22.

[72] O.Bozorg-Haddad, M. Solgi & H. A. Loáiciga. (2017). Meta-heuristic and evolutionary algorithms for engineering optimization. John Wiley & Sons.

[73] H.Azzag, F.Picarougne, C.Guinot & G.Venturini. (2004). Un survol des algorithmes biomimétiques pour la classification. Classification et fouille de données.

[74] S.PALMER.(2019) Evolutionary algorithms and computational methods for derivatives pricing. Doctoral dissertation. UCL (University College London).

[75] T. Nakane, N.Bold, H. Sun, X.Lu & al. (2020). Application of evolutionary and swarm optimization in computer vision: a literature survey. IPSJ Transactions on Computer Vision and Applications, 12(1). https://doi.org/10.1186/s41074-020-00065-9

[76] C.Blum & X.Li .(2008). Swarm intelligence in optimization. In Swarm intelligence: introduction and applications (pp. 43-85). Berlin, Heidelberg: Springer Berlin Heidelberg.

[77] M.Abdel-Basset, L.Abdel-Fatah& A. K. Sangaiah. (2018). Metaheuristic algorithms: A comprehensive review. Computational intelligence for multimedia big data on the cloud with engineering applications, 185-231.

[78] J.Kennedy& R. Eberhart, (n.d.). Particle swarm optimization. Proceedings of ICNN'95 - International Conference on Neural Networks.

[79] C. W. Reynolds. (1987). Flocks, herds and schools: A distributed behavioral model. ACM SIGGRAPH Computer Graphics, 21(4), 25–34. https://doi.org/10.1145/37402.37406
[80] Frank Heppner & Ulf Grenander. (1990).A stochastic nonlinear model for coordinated bird flocks. The ubiquity of chaos, pages 233–238.

[81] H.Zamani & M.-H. Nadimi-Shahraki. (2016). Swarm Intelligence Approach for Breast Cancer Diagnosis. International Journal of Computer Applications, 151(1), 40–44.

[82] E.Atashpaz-Gargari & C. Lucas. (2007). Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition. IEEE Congress on Evolutionary Computation.

[83] T.YOU, Y.HU, P.LI & Y. TANG. (2019). An improved imperialist competitive algorithm for global optimization. TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES, 27(5), 3567–3581.

[84] Z.Beheshti & S. M. H. Shamsuddin. (2013). A review of population-based meta-heuristic algorithms. Int. j. adv. soft comput. appl, 5(1), 1-35.

[85] K.Günel, İ.Gör & K. Tekeli. (2020). ICA-RD: The Regional Domination Policy for Imperialist Competitive Algorithm from Imperialism to Internationalism. Arabian Journal for Science and Engineering, 45(12), 10529-10589.

[86] L. Huang, P. G.Asteris, M. Koopialipoor & al. (2019). Invasive Weed Optimization Technique-Based ANN to the Prediction of Rock Tensile Strength. Applied Sciences, 9(24), 5372.

[87] A. R.Mehrabian & C. Lucas. (2006). A novel numerical optimization algorithm inspired from weed colonization. Ecological Informatics, 1(4), 355–366.

[88] A.Rezaei Pouya, M.Solimanpur & M. Jahangoshai Rezaee. (2016). Solving multi-objective portfolio optimization problem using invasive weed optimization. Swarm and Evolutionary Computation, 28, 42–57.

[89] N. Razmjooy & M. Ramezani. (2014). An improved quantum evolutionary algorithm based on invasive weed optimization. Indian J Sci Res, 4(2), 413-422.

[90] W.Abdulelah Qasim & B.Ahmed Mitras. (2020). A Hybrid Algorithm Based on Invasive Weed Optimization Algorithm and Grey Wolf Optimization Algorithm. International Journal of Artificial Intelligence & Applications, 11(1), 31–44.

[91] M. R.Panda, S.Dutta & S. Pradhan. (2017). Hybridizing Invasive Weed Optimization with Firefly Algorithm for Multi-Robot Motion Planning. Arabian Journal for Science and Engineering, 43(8), 4029–4039.

[92] O.Takahashi & R. J. Schilling. (1989). Motion planning in a plane using generalized Voronoi diagrams. IEEE Transactions on Robotics and Automation, 5(2), 143–150.

CHAPTER III
OVERVIEW OF DFT APPROACH

III.1. Introduction

Over the past thirty years, theoretical advances and continuous advancements in processing power have made it possible to analyze different physicochemical materials properties using their tiny components in a variety of ways. They utilized techniques from statistical physics and quantum mechanics. The Density Functional Theory (DFT) is considered the most widely used method for calculating the structural, elastic, electrical, optical, and thermodynamic properties of materials, both in quantum chemistry and condensed material physics. In 1927, Llewellyn Thomas [1] and Enrico Fermi [2] conducted research aimed at introducing electronic density into the study of materials physical properties. However, the current approach on which DFT is based only emerged in the mid-1960s, following the work of Hohenberg, Kohn, and Sham [3]. Figure III.1 shows an increase in the number of publications specifically related to the study of various properties of materials using the DFT method over time.

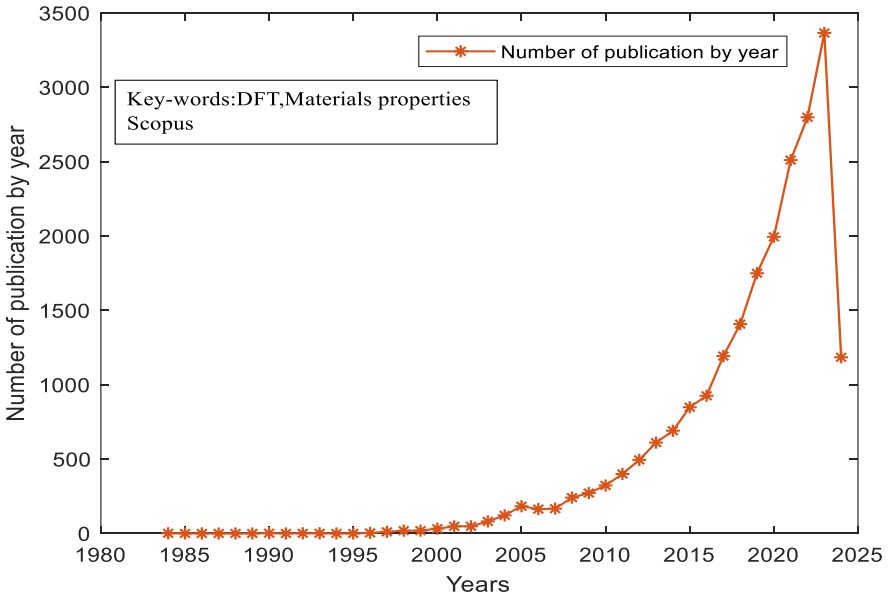

Figure.III.1 Number of publications by year on materials properties using DFT methods. **Keyword** search "DFT in materials properties" in Title, Abstract and Keywords. Source **Scopus.**

The main goal of density functional theory (DFT) is to computationally solve the Schrödinger equation without relying on experimental data and quantitatively determine the wave function and energy of a many-electron system [4]. The basic concept of DFT involves replacing the complex multi-electron wave function, which relies on 3N variables (where N represents the

total number of particles in the system), with the electronic density [5].This density depends on only three spatial coordinates, or, if considering both spin values (up and down), a maximum of six coordinates to describe the system in its magnetic state. This simplification enables a more intuitive and practical understanding of the wave function, both in terms of concept and mathematics [6].

In this chapter, we will briefly cover the DFT foundations and the several approximation levels used in its implementation.

III.2. Schrödinger Equation

It is necessary to characterize the physical properties of a condensed material, such as whether it's structural, electrical, magnetic, optical, and so on. This necessitates understanding the behavior of particles in a physical system that approaches reality and its interatomic interactions [7]. A wave equation developed by Schrödinger can describe how a particle behaves in a force or potential environment [8]. We can find the Schrödinger equation, the fundamental equation of quantum mechanics, as follow:

$$H\psi = E\psi \tag{III.1}$$

Where E is the total energy of the system, ψ is the wave function that fully describes the state of a particle at any given time t, and H is the Hamiltonian operator of the study system that is defined by the summation of operators for the kinetic moments of the nucleus and electrons, $T_N$and T_e, as well as the potential energies representing the Coulomb interactions between electron-electron V_{ee}, electron-nucleus V_{eN}, and nucleus-nucleus V_{NN}.As a result, there are two components to the Hamiltonian operator: a kinetic contribution and a potential contribution ,and it can be written as [9]:

$$H = T_e + T_N + V_{ee} + V_{eN} + V_{NN} \tag{III.2}$$

This operator can be defined, respectively, as follows:

$$T_e = \frac{-\hbar^2}{2m}\sum_i \nabla_i^2 = \frac{-\hbar^2}{2m}\sum_i \Delta_i \tag{III.3}$$

$$T_n = -\frac{\hbar^2}{2}\sum_k^n \frac{1}{M_k}\nabla_k^2 = -\frac{\hbar^2}{2}\sum_k^N \frac{1}{M_k}\Delta_k \tag{III.4}$$

$$V_{ee} = \frac{1}{2}\sum_i^N \sum_j^N \frac{e^2}{4\pi\varepsilon_0 r_{ij}} \tag{III.5}$$

$$V_{NN} = \frac{1}{2}\sum_k^N \sum_i^N \frac{z_k z_l e^2}{4\pi\varepsilon_0 R_{kl}} \quad \text{(III.6)}$$

$$V_{eN} = \frac{1}{2}\sum_i^N \sum_k^N \frac{z_k e^2}{4\pi\varepsilon_0 r_{ik}} \quad \text{(III.7)}$$

Hence the Schrödinger equation can be written as follow:

$$\left(\frac{-\hbar^2}{2m}\sum_i \Delta_i + -\frac{\hbar^2}{2}\sum_k^N \frac{1}{M_k}\Delta_k + \frac{1}{2}\sum_i^N \sum_i^N \frac{e^2}{4\pi\varepsilon_0 r_{ij}} + \frac{1}{2}\sum_k^N \sum_l^N \frac{z_k z_l e^2}{4\pi\varepsilon_0 R_{kl}} + \frac{1}{2}\sum_i^N \sum_k^N \frac{z_k e^2}{4\pi\varepsilon_0 r_{ik}}\right)\psi = E\psi \quad \text{(III.8)}$$

Where: the electron mass is m, r_{ij} denotes the distance between electrons I and J, M_k: stands for the nose mass, z_k and z_l are the atomic numbers for ions k and l. R_{kl} is the distance between their centers.

The Schrödinger equation can be solved and provides only an analytical solution in the case of N=1 such as for hydrogen atoms with a single electron, but it is too difficult to solve in a physical system that seems to be real (N>1).Thus, how can we solve the multi-body problem? Furthermore, how can one get the lowest energy, or fundamental energy, from the Schrödinger equation?

It is difficult to respond to this issue without making several approximations. Therefore, in all other cases, it will be necessary to make a number of approximations about the nature of the wave function, as modern quantum mechanics has no method to solve problems concerning a large number of particles. This is not only due to technical difficulties in calculating order but also to the impossibility of doing this, since in 1 cm^3 of a crystalline solid there are about 5×10^{22} atoms, which means a significant number of variables on the order of 10^{24}. The Born-Oppenheimer approximation, Hartree and Hartree-Fock approximation, and Density Functional Theory are the three main approximation approaches proposed to facilitate the resolution of the N-electron problem.

- **Born Oppenheimer approximation** [8]: In general, quantum calculations are based on the ability to separately treat the electrons and nuclei of a real system. This forms the Born-Oppenheimer adiabatic approximation. The Born-Oppenheimer approximation considers the nucleus to be fixed due to its large mass, thus neglecting its kinetic energy (T_N) and potential energy (V_{NN}). However, it ignores the interactions between the electrons themselves, which necessitates the second approximation.

- **Hartree [10] and Hartree-Fock [11] approximations**: These approximations propose the use of a free electron that has no interaction with the other particles. They searched for a numerical solution based on the Slater determinant [12]. However, they are also insufficient for solving the Schrödinger equation due to the large number of configurations.
- **Density Functional Theory (DFT) [13]** which explained that the energy of an electronic system can be expressed in terms of its density ρ(r) as a basic variable, and to write any other parameters in terms of this variable. The DFT is the most accurate theory because it considers particle interactions. It is valid in high-dimensional systems as well as solids, compared to the Hartree-Fock assumption, which is only applicable in quantum chemistry.

III.3. Density Functional Theory (DFT)

Over the past 50 years, density functional theory has been the most effective and widely used theory in the theoretical calculation of the physical properties of condensed materials [14], due to its superior accuracy and practicality compared to other approximations. Hence, it can handle the correlation of a system comprising a large number of electrons [15–17]. As indicated by its name, DFT consists of using the electronic density ρ(r), which is defined as the number of particles per unit volume, as a variable of the problem rather than the wave function. The use of electron density as a fundamental variable to describe the system's properties has existed since the early approaches to the electronic structure of matter, but it only yielded tangible results through the demonstration of the two theorems known as the Hohenberg and Kohn theorems [18].

III.3.1. Hohenberg and Kohn theorems

The method employed by Walter Kohn and Pierre Hohenberg involves reformulating Thomas and Fermi's density functional theory into an exact many-body system theory [9]. The formulation is based on two fundamental theorems presented and proved by Hohenberg and Kohn in their work in 1964 [18], which can be applied to any system of interacting particles evolving within an external potential [9].

- The first theorem

The electron density $\rho(r)$ in the ground state can determine the external potential V_{ext} that interacts with a system of N particles. This theorem means that it is only necessary to know the electron density to determine all the wave functions. It is based solely on the principle of minimizing the energy at the fundamental level.

- The second theorem

For a given external potential and a fixed number of electrons, the ground state of the system is the global minimum of the functional, while the density that minimises this functional is the fundamental density. Any multi-particle system's total energy functional has a minimum that corresponds to the ground state. The density of particles in the ground state satisfies [9]:

$$E(\rho_0) = Min\, E(\rho) \tag{III.9}$$

Regrettably, the absence of an analytical expression for these two terms led Kohn and Sham in 1965 [19] to reframe the problem of N interacting particles as a non-interacting one. In clearer terms, they assume that the electrons are free and have the same electron density.

The Kohn and Sham equations become [9]:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{ion}(\vec{r}) + V_H(\vec{r}) + V_{XC}(\vec{r})\right]\psi_i(\vec{r}) = \varepsilon_i\psi_i(\vec{r}) \tag{III.10}$$

$\psi_i(\vec{r})$: The wave function of electron i.

$V_{ion}(\vec{r})$: Represent the ionic potential.

$V_H(\vec{r})$: Represent the Hartree term given by:

$$V_H(\vec{r}) = \int \frac{\rho(\vec{r}_1)(\vec{r}_2)}{|\overrightarrow{r_1} - \overrightarrow{r_2}|} d\overrightarrow{r_1} d\overrightarrow{r_2} \tag{III.11}$$

Regards to density, the exchange-correlation potential is obtained from the derivative of the exchange-correlation energy E_{xc}.

$$V_{xc}(\vec{r}) = \frac{\partial E_{xc}[\rho(\vec{r})]}{\partial \rho(\vec{r})} \tag{III.12}$$

Thus, the Kohn-Sham equations may be expressed as follows:

$$H\psi_i(\vec{r}) = \left\{-\frac{\hbar^2}{2m}\nabla^2 + V_{eff}(\vec{r})\right\}\psi_i(\vec{r}) = \varepsilon_i\psi_i(\vec{r}) \tag{III.13}$$

Where the effective potential produced by all of the nuclei and other electrons applies to every electron. This potential is provided by:

$$V_{eff}(\vec{r}) = V_{ext}(\vec{r}) + \int \frac{1}{|\vec{r_i}-\vec{r_j}|}\rho(\vec{r})d\vec{r_j} + V_{XC}(\vec{r}) \quad \text{(III.14)}$$

The interrelated Kohn-Sham equations can be summarised as shown in figure III.2.

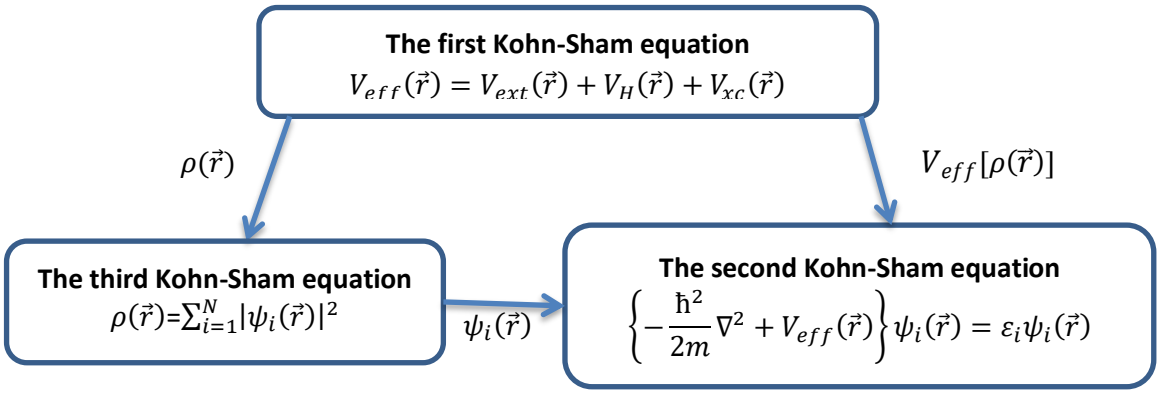

Figure.III.2 The interrelated Kohn-Sham equations [9].

These equations are resolved using the self-consistent field (see figure III.3.), which is a technique employed for solving the Kohn-Sham equations and deriving the ground-state electron density of a system. This method starts with an initial estimation of the electron density, computes the effective potential using this density, and solves the Kohn-Sham equations to obtain a new set of orbitals and eigenvalues. The electron density is subsequently adjusted according to the new orbitals, and the process iterates until convergence is reached. This iterative SCF approach can be applied for various electron counts, allowing exploration of the system's behavior across different scenarios.

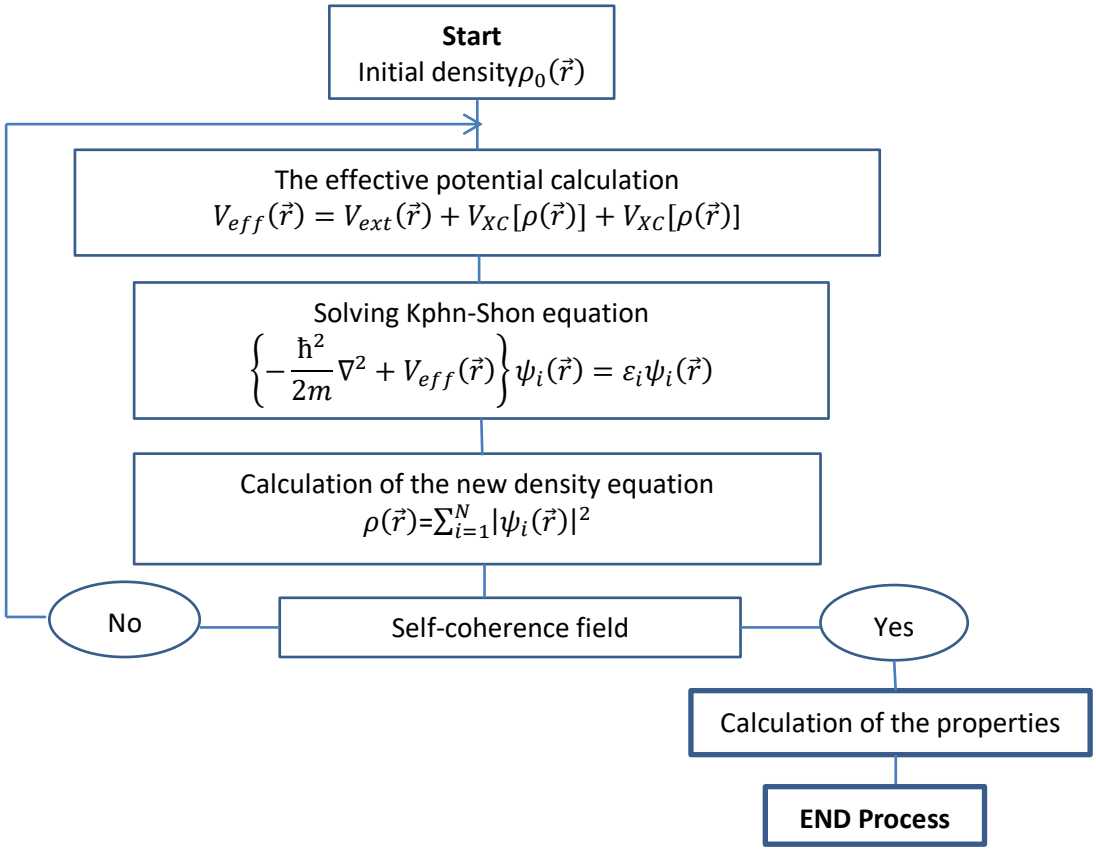

Figure.III.3 Self-consistent iteration process used to solve the Kohn equations Sham equations[20].

This solution only allows us to find the electron density and energy, but it remains unable to precisely address the Kohn-Sham equations if it cannot compute the exchange-correlation potential. Therefore, researchers have proposed numerous approximations for this purpose.

III.3.2. The different types of exchange-correlation functional

III.3.2.1.The Local Density Approximation (LDA)

The "LDA" (Local Density Approximation), a major breakthrough in the early days of DFT and the foundation of all current approaches, replaces the exchange-correlation energy with that of the well-known homogeneous gas to describe the exchange-correlation term [21].

In other words, we assume that in a small spatial region, the charge distribution of a non-homogeneous gas has a density similar to that of a homogeneous gas distribution. The LDA is an approximation based on the assumption that the exchange-correlation terms depend only on the local density value.

$$E_{xc}^{LDA}(\rho) = \int \rho(r)\varepsilon_{xc}[\rho(r)]d^3r \quad \text{(III.15)}$$

Where $\varepsilon_{xc}[\rho(r)]$ represents the exchange-correlation energy per particle of a homogeneous electron gas with density ρ.

The exchange-correlation potential $V_{xc}^{LDA}(r)$may be calculated variationally from the exchange-correlation energy $\varepsilon_{xc}[\rho(r)]$using the following equation:

$$V_{xc}^{LDA}(r) = \frac{\delta(\rho(r)\varepsilon_{xc}[\rho(r)])}{\delta\rho(r)} \quad \text{(III.16)}$$

LDA assumes that the exchange-correlation energy is purely local and is composed of two terms: the exchange energy and the correlation energy.

$$\varepsilon_{xc}[\rho(r)] = \varepsilon_x[\rho(r)] + \varepsilon_c[\rho(r)] \quad \text{(III.17)}$$

In accordance with the Dirac exchange functional [22], we can write the exchange-related term as follows:

$$\varepsilon_x[\rho(r)] = -\frac{3}{4}\left(\frac{3\rho(r)}{\pi}\right)^{\frac{1}{3}} \quad \text{(III.18)}$$

Numerous publications have been made about the parametric modeling of ε_{xc}^{LDA}, such as those by Vosko, Wilk, and Nusair [23], Perdew and Zungar [24].

In the case of magnetism, we refer to the Local Spin Density Approximation (LSDA) [24], which considers the population of spin-up and spin-down levels to describe the physical properties of a material in its ground state.

$$E_{xc}^{LSDA}[\rho\uparrow, \rho\downarrow] = \int \rho(r)\varepsilon_{xc}[\rho(r)\uparrow, \rho\downarrow(r)]d^3r \quad \text{(III.19)}$$

When the density fluctuates slowly in a system, the local density approximation (LDA) yields satisfactory results. It performs poorly in more heterogeneous density systems. This has led to numerous advancements to improve the outcomes. For this reason, the electron density gradient was introduced, leading to the generalized gradient approximation (GGA), where the exchange-correlation energy is a function of both the electron density and its gradient [9].

III.3.2.2.The generalized gradient approximation (GGA)

In reality, the electron density of molecular and atomic systems differs from that of the homogeneous electron gas and varies spatially. This leads us to the second approximation, known as "GGA." The generalized gradient approximation (GGA), sometimes referred to as non-local methods, has been developed to account for this variation in electron density, which is manifested by a gradient. As a result, the exchange-correlation energy is dependent not only on the electron density, but also on the magnitude of the electron density gradient [25–26].

It is generally given as follows:

$$E_{xc}^{GGA}[\rho] = \int f[\rho(\vec{r}), \nabla\rho(\vec{r})]\, d^3r \qquad \text{(III.20)}$$

Where $\nabla\rho(r)$ is the gradient of the electron density.

Perdew and Wang ("GGA-PW") [25], Perdew-Burke-Ernzerhof ("GGA-PBE") [26], and Perdew et al. and Tao et al. have developed several versions of the GGA approximation, including the "Meta-GGA" [27]. The difference between these lies in the choice of the derivatives of the XC energy with respect to density gradients higher than one, as well as the details of the parameterization. The GGA provides an improvement over the LDA, as it underestimates the lattice parameter and yields good results. However, it still fails to accurately predict the value of the band gap (the gap is consistently underestimated) [28–29].

III.4. The modified Becke-Johnson potential (mBJ)

This latter category corresponds to the use of a semi-local approximation with an exact exchange atomic potential and a screening term. Recently, F. Tran and P. Blaha developed a local potential called the modified Becke-Johnson potential (mBJ) [30], which is arguably the first local potential to provide energy bands that bridge experimental gaps (Figure III.3). The potential of the mBJ-LDA approach, as proposed by F. Tran et al [30, 31], is a parameterized version of the optimized Becke and Johnson (BJ) potential [32], aiming to reproduce the form of exact exchange for atoms, with the addition of LDA correlation energy [9]. The modified Becke-Johnson exchange-correlation $E_{xc}^{mBJ}(r)$ is proposed as follows [33]:

$$E_{xc}^{mBJ}(r) = c\left(E_x^{BR}(r) + (3 - \frac{2}{c})\frac{1}{\pi}\sqrt{\frac{1}{12}}\sqrt{\frac{2\tau_\sigma(r)}{\rho_{\sigma(r)}}}\right) + E_{xc}^{LDA}(r) \qquad \text{(III.21)}$$

Where: c is a real fractional number, $E_x^{BR}(r)$ symbolizes the Becke-Roussel exchange potential [34], $\tau_\sigma(r)$ represents the kinetic energy density, and $\rho_{\sigma(r)}$is the electron density [33].

The value of c is calculated self-consistently using equation (III.22).

$$c = \gamma + \beta\left(\frac{1}{\Omega}\int_{\Omega}\frac{|\nabla\rho(r')|}{\rho(r')}d^3r'\right)^{1/2} \qquad \text{(III.22)}$$

Where: γ(a constant with a value of -0.012) and β(=1.023$\sqrt{a_0}$) are two independent parameters, where a_0=5.29177×10^{-11}m is the Bohr radius.

Therefore, the mBJ-LDA potential is based on an approximate Slater potential (as a local approximation of Hartree-Fock) and a density filtering term, its gradient, as well as the Laplacian and kinetic energy density as input [33]. The strengths of the various terms are determined by an average of γ and β, thus no specific input is required. From this point on, mBJ-LDA's formalism is not a semi-empirical method but a semi-local potential procedure [9, 30].

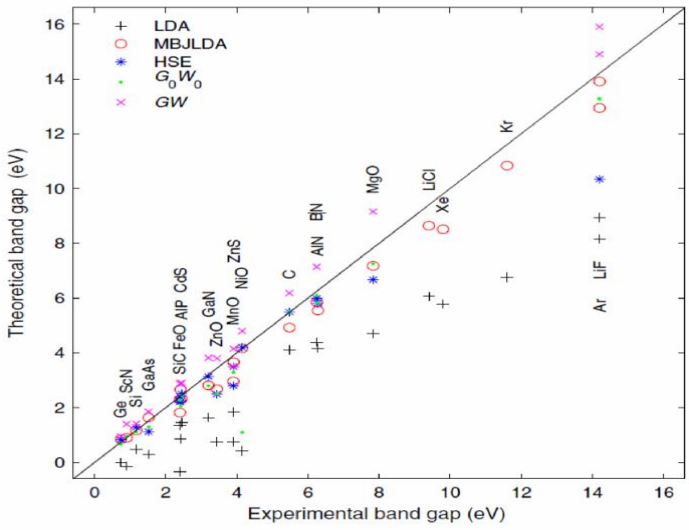

Figure.III.4 The band gap Eg (theoretical Vs experimental), for several semiconductors, materials and insulators using various methods [9, 34].

III.5.The Full-Potential Linearized Augmented Plane-Wave (FP-LAPW)

In fact, the DFT formalism has led to the creation of several computation techniques. Two main considerations inform the decisions made in an attempt to simplify the solution to the Kohn-Sham equations [35]:

- Selecting the wave function basis for projecting the Kohn-Sham states with a single electron. These wave function bases are divided into the following three categories:
  - ✓ Linear Combination of Atomic Orbitals (LCAO).
  - ✓ Plane Waves (PW).
  - ✓ Linearized Augmented Plane Waves (LAPW).
- Selecting the type of effectual potential generated by an infinite number of ions, or external potential, where three forms of potential may be identified:
  - ✓ The JELLUM model and the pseudopotential method.
  - ✓ The method of all electrons.

III.5.1.The Augmented Plane Wave (APW) method

According to Slater's proposal in 1937 [36], the solution to the Schrödinger equation is a plane wave for a constant potential. For a spherical potential, the solution is a radial function. To achieve this, Slater described the crystalline potential using the "Muffin-tin" approximation. This approximation involves dividing the crystalline system into two non-overlapping regions: an atomic (spherical) region and a second interstitial region. (See Fig.III.5).

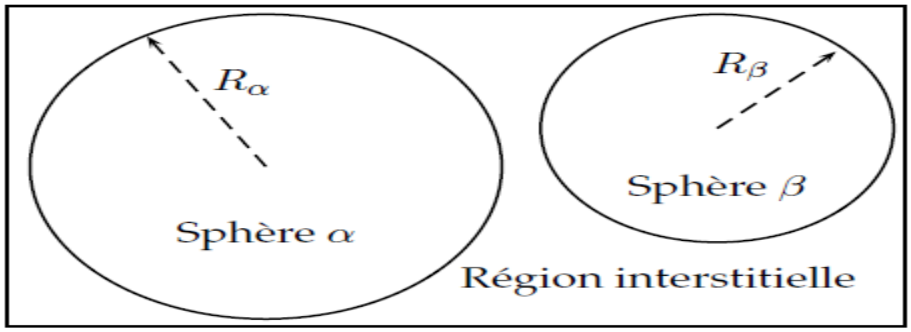

Figure.III.5 Distribution of the unit cell, an interstitial region, and spherical regions: α and β spheres with Muffin-tin radii Rα and Rβ, respectively [29].

The two spherical and interstitial regions are defined by the wave functions φ_S and φ_I:

$$\varphi_S(r) = \sum_{lm} A_{lm}\, u_l^{\alpha}(r, E_l) Y_{lm} \quad \text{(III.23)}$$

and

$$\varphi_I(r) = \frac{1}{\sqrt{\Omega}} \sum_G C_G e^{i(G+\vec{k})r} \quad \text{(III.24)}$$

Where r is the position in polar coordinates inside the sphere, C_G and $A_{lm}$are expansion coefficients, Ω is the volume of the unit cell, $\vec{k}$ is the wave vector in the irreducible Brillouin zone (IBZ), G is a vector in reciprocal space, and u_l is the numerical solution of the radial part of the Schrödinger equation with energy E_l, and V is the spherical component of the potential.

$$\left[-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + V(r) - E_l\right] r u_l(r) = 0 \quad \text{(III.25)}$$

This equation defines the radial function orthogonal to any eigenstate of the same Hamiltonian and which vanishes at the sphere boundaries. The overlap of these functions is constructed from:

$$(E_2 - E_1) r u_1 u_2 = u_2 \frac{d^2 u_1}{dr^2} - u_1 \frac{d^2 u_2}{dr^2} \quad \text{(III.26)}$$

Where u_1 and $u_2$are the radial solutions at different energies E_1 and E_1, respectively.

Slater employed this approximation called the "Muffin-tin" approximation with a condition that must be imposed (the representations defined in expressions (III.23 and III.24) are continuous at the boundaries of the spheres). The coefficients $A_{lm}$are determined in terms of the $C_G$in the expansion of the spherical harmonics of plane waves. The augmented plane waves are the individual functions labeled by G.

$$u_{lm} = \frac{4\pi i^l}{\frac{1}{\Omega^2} u_l(R)} \sum_G C_G J_l(|k+g|)\, Y_{lm}^*\left(\vec{k} + \vec{G}\right) \quad \text{(III.27)}$$

The issue with this method is that for each eigenvalue, one needs to find a set of different energy-dependent basis functions of APW. This creates a unique non-linear value problem in programming.

The main drawback of this approach (the asymptotic problem) is that the APW is quite complex and only works with basic systems with few inherent values. This issue leads us to the LAPW method, which exhibits accuracy and flexibility in calculating the band structure [29].

III.5.2. Linear Augmented Plane Wave method (LAPW)

In order to address the issues of the APW method, Anderson [37] proposed a more effective method in 1975, which is known as the "LAPW" method, or the linear method of elevated planes. This method's base functions and their derivatives are continuous and adjusted to the radial function and its derivative in terms of parameter fixations. This may address the APW method's asymptotic problem.

The term "linearization" in the LAPW method refers to the linear energy of each radial function within each atomic sphere, implying the use of a linear combination of the radial functions and their energy derivatives. Here is the definition of the LAPW method:

$$\varphi(r) = \begin{cases} \sum_{lm}\left[A_{lm}u_l(r) + B_{lm}\dot{u_l}(r)\right]Y_{lm} & pour\ r \in S \\ \frac{1}{\sqrt{\Omega}}\sum_G C_G e^{i(G+\overrightarrow{K})\vec{r}} & pour\ r \in I \end{cases} \qquad \text{(III.28)}$$

Where the coefficients A_{lm} and B_{lm} are determined in such a way as to meet the continuity requirements between zones (S) and (I).

In general, the LAPW enables the determination of:

- ✓ Radial functions and their derivatives
- ✓ Coefficients A_{lm} and B_{lm} satisfying boundary conditions. This also allows the determination of the cutoff angular momentum ($cutoff$) $lmax$, and the cutoff ($cutoff$) for the plane waves used $Gmax$.

In contrast to the APW method, the LAPW method obtains all of the own energy with a single diagonalization and continues all of the wave functions to the MT sphere [38].

III.6. WIEN2k code's description and uses

Blaha, Schwartz, and Luitz [39,40] from the Institute of Materials Chemistry at the Vienna University of Technology (Austria) developed the computational code Wien2k. Since its initial distribution in 1990, this code has undergone continuous revisions and updates. Subsequent versions are named according to the year of their release (WIEN93, WIEN95, WIEN97, etc.). The LUNIX operating system powers this Fortran-written computer program, known as the simulation code. It is based on density functional theory and the full-potential linearized augmented plane wave (FP-LAPW) method. A C-SHEL script links several independent programs, enabling self-consistent calculations. For each calculation, we will outline the essential procedures, which are as follows:

Insert the important parameters: In this step, we fill an input file named case.struct (signification: material name. Struct) with the following atomic parameters:

- The lattice type.
- The space group.
- The lattice parameters (a, b, and c in Bohr or Å).
- The angles (α, β, and γ).
- The positions of the atoms inside the cell (x, y, and z).
- The muffin-tin radii (Rmt), given in atomic units (Bohr radius).

Initialization: After generating this file "case.struct," we perform initialization using the command "init_lapw" to trigger several programs to execute successively. These programs are as follows [35]:

- ✓ NN: This program provides distances between nearest neighbors, aiding in determining the atomic radius of the sphere.
- ✓ LSTART: A program that generates atomic densities and determines how different orbitals are treated in the band structure calculation, such as core states with or without local orbitals.
- ✓ SYMMETRY: It generates symmetry operations of the space group, determines the point group of individual atomic sites, generates lm expansion for lattice harmonics, and determines local rotation matrices.
- ✓ KGEN: It generates a k-point mesh in the Brillouin zone.
- ✓ DSTART: It generates an initial density for the SCF cycle by superimposing atomic densities generated in LSTART.

Then a self-consistent cycle is initialized and repeated until the convergence criterion is met. This cycle consists of the following steps:

- LAPW0: Computes the potential as the sum of the Coulomb potential Vc and the exchange-correlation potential V_{xc}, where it utilizes the total electron density as input.
- ORB: This program is to be executed only in DFT+U (GGA+U or LSDA+U). It calculates the orbital-dependent potential.
- LAPW1: Computes the matrix coefficients of the Hamiltonian in the LAPW wave basis and finds the eigenvalues and eigenvectors by diagonalization.

- LAPW2: Determines the Fermi level, the expansions of valence electron densities consisting of electron densities within each MT sphere (expressed by spherical harmonics) and in the interstitial region (expressed by a Fourier series).
- LCORE: Computes core states inside the MT spheres, retaining only the spherical part of the potential.
- MIXER: Computes the new electron density by mixing core electron densities, semi-core states, and valence states to generate the input density for the next iteration.

These main calculation steps using the Wien2k code are illustrated in Figure.III.6.

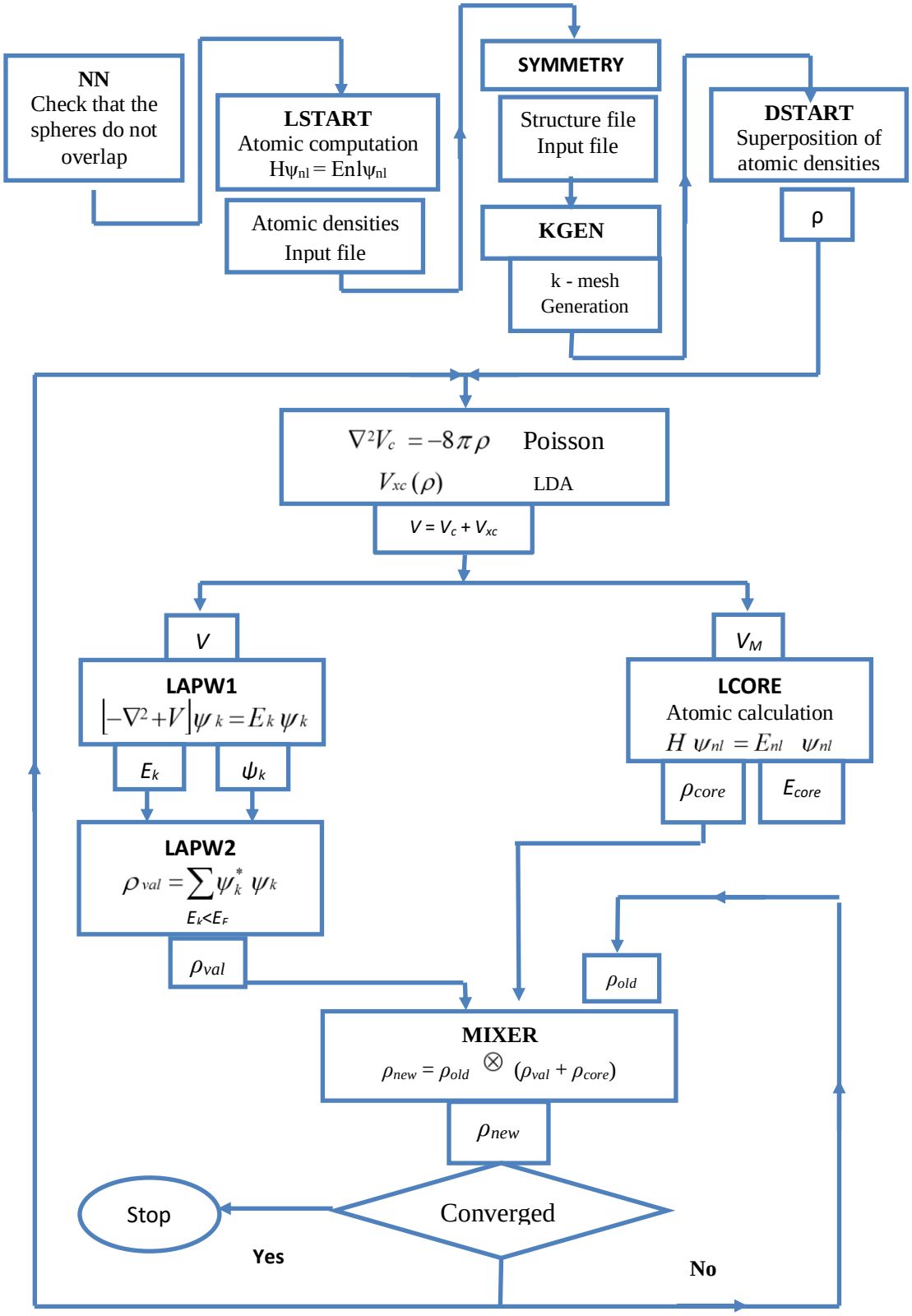

Figure.III.6 FP-LAPW method flowchart (Wien2k) [35].

III.7.Conclusion

We have dedicated this chapter to presenting the Density Functional Theory (DFT) and the Linearized Augmented Plane Wave (FP-LAPW) method. DFT uses two important approximations, namely the Local Density Approximation (LDA) and the Generalized Gradient Approximation (GGA). We wrap up this chapter by showcasing the Wien2k code, exclusively installed on the Linux operating system, and its implementation of the FP-LAPW method.

References

[1] L. H. Thomas, (1927). The calculation of atomic fields. Mathematical Proceedings of the Cambridge Philosophical Society, 23(5), 542–548.

[2] E. Fermi, (1927). Statistical method to determine some properties of atoms. Rend. Accad. Naz. Lincei, 6(602-607), 5.

[3] W.Kohn & L. J. Sham, (1965). Self-Consistent Equations Including Exchange and Correlation Effects. Physical Review, 140(4A), A1133–A1138.

[4] N.Argaman & G. Makov, (2000). Density functional theory: An introduction. American Journal of Physics, 68(1), 69–79.

[5] E. Bernes, (2021). Simulation of core-electron spectroscopies of gas-phase molecules and adsorbates by DFT and TDDFT methods.

[6] C.Fiolhais, F.Nogueira , & M. Marques, A. (Eds.). (2003). A primer in density functional theory (Vol. 620). Springer Science & Business Media.

[7] E. Schrödinger, (1948). 2400 Jahre Quantentheorie. Annalen Der Physik, 438(1), 43–48. Portico.

[8] M.Born, & M.Oppenheimer, M. (1927). Zur Quantentheorie der Molekeln. Annalen Der Physik, 389(20), 457–484.

[9] H. BENNACER, (2015). Contribution au développement des nouvelles cellules solaires à base des semi-conducteurs composes (Doctoral dissertation). University of Sidi Bel Abbesse.

[10] D. R. Hartree, (1928). The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. Mathematical Proceedings of the Cambridge Philosophical Society, 24(1), 89–110.

[11] V. Fock, (1930). Selfconsistent field? mit Austausch f ür Natrium. Zeitschrift Für Physik, 62(11–12), 795–805.

[12] J. C. Slater, (1951). A Simplification of the Hartree-Fock Method. Physical Review, 81(3), 385–390.

[13] H.Kronmüller, & S. Parkin. (2007). Handbook of Magnetism and Advanced Magnetic Materials.

[14] N.Argaman & G. Makov (2000). Density functional theory: An introduction. American Journal of Physics, 68(1), 69–79.

[15] Lakhdar DJOUDI, (2013). Etudes des propriétés structurales, electroniques et optiques de l'alliage Semi-conducteur BAlGaN: Etude ab-initio (Doctoral dissertation). University of Sidi Bel Abbesse.

[16] E.Engel& R. M. Dreizler, (2011). Density Functional Theory. In Theoretical and Mathematical Physics. Springer Berlin Heidelberg.

[17] S. Cottenier, (2002). Density Functional Theory and the family of (L) APW-methods: a step-by-step introduction. Instituut voor Kern-en Stralingsfysica, KU Leuven, Belgium, 4(0), 41.

[18] P.Hohenberg & W. Kohn, (1964). Inhomogeneous Electron Gas. Physical Review, 136(3B), B864–B871.

[19] W.Kohn & L. J. Sham, (1965). Self-Consistent Equations Including Exchange and Correlation Effects. Physical Review, 140(4A), A1133–A1138.

[20] D. Varsano, (2006). First principles response functions in low dimensional systems (Doctoral dissertation, UPV/EHU).

[21] D. M.Ceperley & B. J. Alder, (1980). Ground State of the Electron Gas by a Stochastic Method. Physical Review Letters, 45(7), 566–569.

[22] R. M.Dreizler & E. K. U. Gross, (1990). Density Functional Theory of Relativistic Systems. Density Functional Theory, 245–271.

[23] S. H.Vosko, L.Wilk, & M. Nusair, (1980). Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. Canadian Journal of Physics, 58(8), 1200–1211.

[24] J. P.Perdew, E. R. McMullen & A. Zunger, (1981). Density-functional theory of the correlation energy in atoms and ions: A simple analytic model and a challenge. Physical Review A, 23(6), 2785–2789.

[25] J. P. Perdew, (1985). Accurate Density Functional for the Energy: Real-Space Cutoff of the Gradient Expansion for the Exchange Hole. Physical Review Letters, 55(21), 2370–2370.

[26] J. P.Perdew, K.Burke, & M. Ernzerhof, (1996). Generalized Gradient Approximation Made Simple. Physical Review Letters, 77(18), 3865–3868.

[27] J. P.Perdew, S. Kurth, A.Zupan, & P. Blaha, (1999). Accurate Density Functional with Correct Formal Properties: A Step Beyond the Generalized Gradient Approximation. Physical Review Letters, 82(12), 2544–2547.

[28] M. B. Farouk , (2015). Polytypisme dans les composés II-VI: Etude de premier principe (Doctoral dissertation, University of ORAN).

[29] Tarik OUAHRANI,(2011).calcul des properities structurales ,hermiqur rt optiques des composes chalcopyrites par la methode FP-(L)APW. Doctorate dissertation. University of Tlemcen.

[30] F.Tran & P. Blaha, (2009). Accurate Band Gaps of Semiconductors and Insulators with a Semilocal Exchange-Correlation Potential. Physical Review Letters, 102(22).

[31] D.Koller, F.Tran & P. Blaha, (2011). Merits and limits of the modified Becke-Johnson exchange potential. Physical Review B, 83(19).

[32] Becke, A. D., & Johnson, E. R. (2006). A simple effective potential for exchange. The Journal of Chemical Physics, 124(22).

[33] T. Das, (2012). Theoretical study of the electronic and optical properties of photocatalytic inorganic materials .Doctoral dissertation, university of Nantes, (2012).

[34] A. D.Becke, & M. R. Roussel, (1989). Exchange holes in inhomogeneous systems: A coordinate-space model. Physical Review A, 39(8), 3761–3767.

[35] S.Haid. (2019). Etude Ab – Initio des Propriétés Structurales, Electroniques et thermodynamique des Matériaux Doubles Pérovskite .Doctorat dissertation ,University of Mostaganem.

[36] J. C. Slater, (1937). Damped Electron Waves in Crystals. Physical Review, 51(10), 840–846.

[37] O. K. Andersen, (1975). Linear methods in band theory. Physical Review B, 12(8), 3060–3083.

[38] Wafa GHAZAL . (2011). Propriétés structurales, électroniques, magnéto-optiques et stabilité structurale des semi-métaux magnétiques. Doctoral dissertation . University of Tlemcen.

[39] P.Blaha et al. (1997).Improved and updated Unix version of the original copyright WIEN code, Technical University. Vienna, (1997).

[40] P. Blaha, K.Schwarz, G.K. Madsen, D.Kvasnicka, & J. Luitz, (2001). wien2k. An augmented plane wave+ local orbitals program for calculating crystal properties, 60(1).

CHAPTER IV
RESULTS AND DISCUSSION

IV.1. Introduction

In materials science, conventional experimental and computational methods often require significant time and resource investments, with experimental conditions imposing constraints on the research. At times, these traditional approaches may fail to deliver the desired outcomes. Hence, there's a pressing need to innovate and devise new methodologies that can expedite experimental advancements while minimizing the wastage of time and resources.

Finding materials with the right property range is tough, even after defining constraints. Scientists mainly use complex ab initio techniques like Standard Density Functional Theory (DFT), which are computationally expensive. As a result, scientists have only conducted extensive studies on a limited number of compounds. However, advancements in instrumentation and electronics have empowered experiments to generate vast amounts of data. The next logical step from this point is the emergence of innovative tools capable of generating more data and extracting knowledge from it. In this context, statistical learning has introduced Artificial intelligence techniques, which are presently guiding research towards using new methods based on AI with the aim of discovering novel materials. In this chapter, we describe our novel approaches based on artificial intelligence techniques and ab initio methods for materials science property prediction, such as electronic and structural properties.

The primary objectives of this chapter can be summarized as follows:

- ✓ Boost prediction accuracy by proposing various approaches and discovering new materials.
- ✓ Determine the optimal approach that best describes the given predicted properties.

Firstly, in the basics of machine learning and computational intelligence methods, several proposed approaches will be applied to two applications:

-The first one consists of band gap prediction for chalcopyrite compounds.

- The second one aims to predict the lattice constants for simple and double-perovskite materials.

Finally, by using density functional theory (DFT), we determined the fundamental properties of the double perovskite Ba_2BiTaS_6 compound. These properties include both the lattice constant and the band gap energy aspects.

IV.2. Chalcopyrite's Energy Band Gaps Prediction

IV.2.1.Prespective

Predicting the electronic properties of a compound before it is manufactured is a crucial step in the creation of novel solar materials. The band gap of any solar material is a crucial property. A material's band gap may be roughly described as the energy required to change its conductive properties. In order to guarantee greater absorption of light with higher energy, the band gap energy of the solar cell material should be selected as low as possible (a higher wavelength). This will excite more electrons, creating more current. Generally, people construct solar cells using materials whose band gap energy ranges from 1.1 to 1.8 eV [1]. Thus, it would be ideal to have a prediction model that can accurately estimate this band gap energy. In this part of our research, we propose a novel method for band gap energy prediction of chalcopyrite compounds based on artificial neural networks (ANN) and the particle swarm optimization technique (PSO). The suggested structure will expedite the better-resolution detection of new chalcopyrite in photovoltaic solar cells. In addition, our method can resolve the local minima problem of an artificial neural network while maintaining fitting quality.

IV.2.2.Contribution

The electronic band gap energy is an essential photo-electronic parameter in the energy applications of engineering materials, particularly in solar cells and photo-catalysis domains. Hence, it would be ideal to have a prediction model that can accurately estimate this band gap energy. A new approach for predicting band gap energy is suggested in this thesis. The proposed structure is based on artificial neural networks (ANN) and the particle swarm optimization algorithm (PSO); this structure can solve the artificial neural network's local minima issue while preserving the fitting quality. Our technique will hasten the identification of novel chalcopyrite in photovoltaic solar cells with improved resolution. The suggested model combines two sub-systems in a parallel configuration. A conventional prediction system with a low resolution for the training data being considered makes up the first ANN sub-system. A second ANN sub-system, called the error model, is introduced to the primary system to address the resolution quality issue, representing uncertainty in the primary model. The particle swarm optimization algorithm is used to identify the parameters of the proposed neural system.

IV.2.3.Data description

In this first application, we concentrate on the ABC_2 chalcopyrite, composed of elements from the periodic table's $I - III - VI_2$ and $II - IV - V_2$ groups. These comprise a total of 302 compounds, each containing three elements and each possessing a unique set of chemical properties, which are gathered from previous studies conducted by Y. Zeng et al. [2], Khmaissia et al. [3], and P. Villars et al. [4–7]. Including the Mendeleev numbers (MN1 and MN2), pseudo-potential radii (PR) (atomic unit au), electron number in the valence band (VL), electron negativity (EN) (eV½), atomic numbers (AN), and melting point (MP) (k), which are the total of seven features that can be used to correlate the band gap energy (Eg) of the chemical ABC_2 as follows:

$$Eg = f(EN_X, AN_X, MP_X, VL_X, PR_X, MN1_X, MN2_X) \quad \text{(IV.1)}$$

X represents the five atoms. These constituted a total of twenty-eight descriptors for each compound. Table IV.1 presents the element input values that comprise the compounds under consideration, which are reported in [2-7]. In this table, Grp refers to the group number within the periodic elements.

Predicting the band gap remains a highly difficult process that depends on a number of the compound's properties. Our goal in studying this data set was to offer a novel method based on particle swarm optimization methods and artificial neural networks to predict the band gap of chalcopyrite.

Table.IV.1 The elements of our data set, with their respective values [2-7].

I-III-IV_2 Compounds									II-IV-V_2 Compounds								
Grp	Elm	EN	AN	MP	PR	VL	MN1	MN2	**Grp**	Elm	EN	AN	MP	PR	VL	MN1	MN2
	Cu	1.08	29	1358	2.04	11	66	68		Zn	1.44	30	692.7	1.88	12	69	71
I	Ag	1.07	47	1235	2.375	11	67	67	**II**	Cd	1.4	48	594.3	2.215	12	70	70
	Au	1.19	79	1338	2.66	11	68	66		Hg	1.49	80	234.3	2.41	12	71	69
										Be	1.45	4	1562	1.08	2	7	12
										Mg	1.31	12	922	2.03	2	8	11
										Ca	1.17	20	1112	3	2	9	10
										Sr	1.13	38	1042	3.21	2	10	9
										Ba	1.08	56	1002	3.402	2	11	8
										Ra	0.9	88	973	3.53	2	12	7
III	B	1.9	5	2365	0.795	3	72	76	**IV**	C	2.37	6	3800	0.64	4	77	81
	Al	1.64	13	933.5	1.675	3	73	75		Si	1.98	14	1687	1.42	4	78	80
	Ga	1.7	31	302.9	1.695	3	74	74		Ge	1.99	32	1211	1.56	4	79	79
	In	1.63	49	429.8	2.05	3	75	73		Sn	1.88	50	505.1	1.88	4	80	78
	Tl	1.69	81	577	2.235	3	76	72		Pb	1.92	82	600.7	2.09	4	81	77
VI	O	3.32	8	54.36	0.465	6	87	91	**V**	N	2.85	7	63.15	0.54	5	82	86
	S	2.65	16	388.4	1.1	6	88	90		P	2.32	15	317.2	1.24	5	83	85
	Se	2.54	34	494	1.285	6	89	89		As	2.27	33	1089	1.415	5	84	84
	Te	2.38	52	722.7	1.67	6	90	88		Sb	2.14	51	903.9	1.765	5	85	83
										Bi	2.14	83	544.6	1.997	5	86	82

IV.2.4.Proposed approach

The primary goal is to create a new prediction system based on ANN-PSO that will speed up and improve the accuracy of the discovery for different material properties. For this work, we used the proposed approach to predict the band gap of chalcopyrite compounds (c.f.chaptre1). Finding the free parameters of the created ANN structure is the second step's goal. This investigation uses the particle swarm optimization algorithm (PSO) method .which is widely employed in diverse settings and disciplines[8,9], to adaptively alter the parameters of the suggested ANN framework .As well as , It works effectively for solving complex optimization problems(c.f.chaptre II). Where the best solution is obtained in parallel with the treatment of its population. The type of neural network used in this study is a feedforward neural network discussed in the section II.5.2.

The proposed framework is based on two sub-models. The primary model initially uses descriptors, including electronic and crystal structure properties of chalcopyrite materials, which are based on ab initio calculations and experimentation. Next, the error between the real band gap and the output of the primary model is modeled to create the error model. This latter model

represents uncertainties in the primary model that can be easily removed by a simple subtraction of the error model output from the primary model output. Details of this method will be presented in the following phases:

***Phase* 1**: Data pre-processing;
***Phase* 2**: Primary prediction system identification;
***Phase* 3**: Error Process identification;
***Phase* 4**: Final prediction system design.

IV.2.4.1. Adjusting the ANN system's parameters

The computational algorithms of artificial neural networks (ANNs) mimic "neurons-based" biological systems. This is described as a network of linked "neurons" that can calculate values based on inputs. The two most important steps are the structure and parameter identification of an artificial neural network prediction system. Therefore, the structure identification step is an essential tool because it tackles the problem of building the ANN prediction system's framework with input-output data [10].

The particle swarm optimization approach can be used to adjust the following free parameters of artificial neural network systems:

- ✓ **Weights**: these are the parameters that change the input data in a neural network's hidden layers.
- ✓ **Biases**: Constants are added to the product of features and weights; they are applied in order to offset the results.

IV.2.4.2.The objective function

The term objective function, or cost function, is used in mathematical optimization to designate a function that serves as a criterion for determining the best solution to an optimization problem. The objective function used is the Mean Square Error (MSE) criterion, which combines actual and predicted values as follows:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n)^2 \; ; \qquad \text{(IV.2)}$$

Where y_n is the actual output, $\hat{y}_n$ is the predicted output and N is the length of data.

IV.2.4.3.Data pre-processing

The preprocessing phase can enhance the success and effectiveness of any machine learning training process. In the current study, all data were normalized using equation IV.3 before the networks were trained, so that all of their values fell into the same range between 0 and 1.

$$Normalized\ data = \frac{Original\ data - \text{Min}\ of the\ original\ dat}{standard\ deviation}. \quad \text{(IV.3)}$$

IV.2.4.4.Primary prediction system identification

This step uses the input-output data provided in the training phase to identify the main ANN-prediction system (Fig. IV.1). The main prediction system's architecture is essentially an online adaptation of the feedforward neural network that is being used. The primary ANN-prediction system's parameters (the parameters optimization block in Figure IV.1) are adjusted using the particle swarm optimization method until the error (e_1) between the experimental band gap values (y) and the primary ANN-prediction system's output ($\hat{y}_p$) achieves the minimum value.

$$e_1 = y - \hat{y}_p; \quad \text{(IV.4)}$$

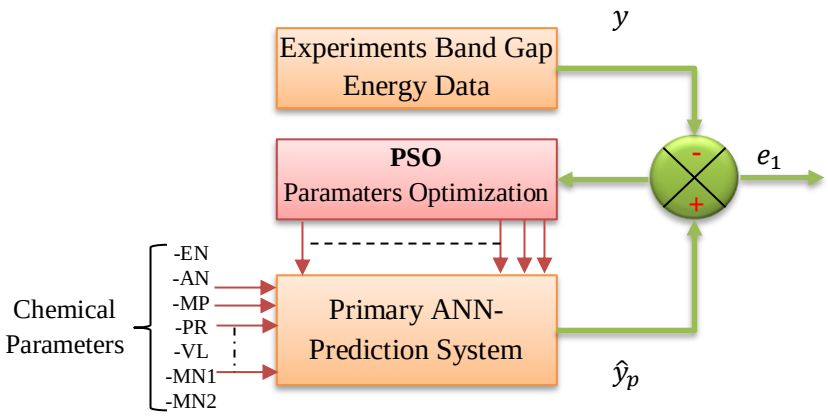

Figure.IV.1 Primary ANN-prediction system.

IV.2.4.5. Error process identification

The main contribution of this study is the introduction of a new prediction module called the error model, which provides the idea of prediction. Let's define the error process e_1 (Eq. IV.4), which stands for the main ANN-prediction system's uncertainty. For the purpose of modeling e_1, a second ANN system (an ANN error model) can be created using this error.

It is reasonable to use an ANN-Auto regressive model to create the model for the error e_1, as it is a time series that seeks to predict the current output based on the past outputs. As a result, we suggest applying the subsequent ANN-AR model to identify e_1:

$$\hat{e}_{1k} = b_0 + b_1\hat{e}_{1k-1} + b_2\hat{e}_{1k-2} + \cdots + b_n\hat{e}_{1k-n} + \mathcal{E}_k \quad \text{(IV.5)}$$

Where $\hat{e}_{1k}$ is the estimated value of e_{1k} ,($\hat{e}_{1k-1}$, $\hat{e}_{1k-2,\ldots}\hat{e}_{1k-n}$) are the previous estimated values of the error process, $\mathcal{E}_k$ is the residual and n is the number of regressors.

Figure IV. 2 shows the structure of this phase; the boxes indicating the delay correspond to the error's previous time series value ($\hat{e}_1$). Now, the goal is to use the particle swarm optimization approach to update the ANN-Error model's parameters online until the error between the model's output $\hat{e}_1$, and error e_1 achieves the smallest value possible.

$$e_2 = e_1 - \hat{e}_1 \quad \text{(IV.6)}$$

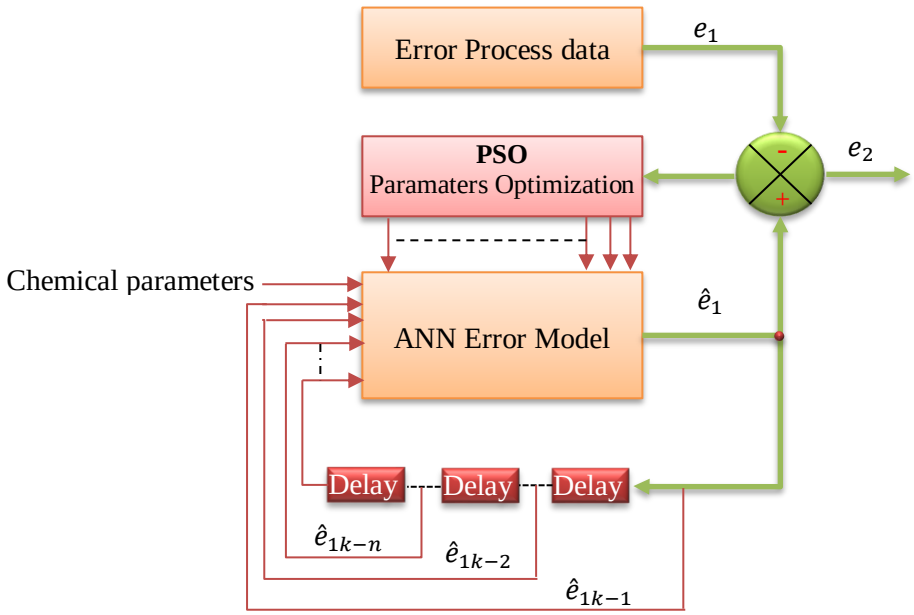

Figure.IV.2 ANN-error model.

IV.2.4.6.Final prediction system

Finally, we should connect the ANN-Error model and the main ANN-prediction system in parallel in order to create the final prediction model (Figure.IV. 3). By using this system, we will be able to minimize the mistake that the basic ANN prediction system produced and achieve

a novel system that is more accurate

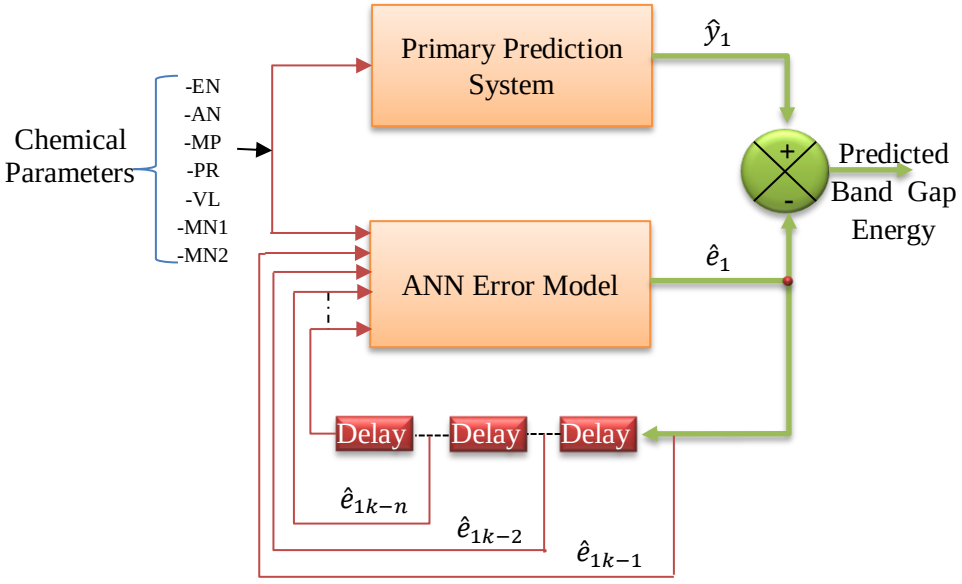

Figure.IV.3 Final prediction system.

IV.2.5. Simulation Results

In this section, we will present and discuss the simulation results of our applied method for predicting chalcopyrite compounds' band gap. A variety of features are used to evaluate our proposed system in terms of prediction errors, and we will investigate the relevance of features used in previous study reports in [3,11-12]. As a starting point, the original data consists of seven elementary descriptors for each atom present in the ternary compounds under consideration (ABC_2). Thus, a total of twenty-eight features (which were detailed in Section IV.2.3) are used to predict the chalcopyrite band gap. Next, we will evaluate whether each of the optimal features reported in [6] is relevant, meaningfully contributes to our prediction system, and offers a trustworthy comparative investigation. Finally, we evaluated our performance method by comparing it to the various benchmarking methods described in [3,11-12]. We used the prediction errors to assess the accuracy of our system.

As a first step, we split our labeled data, which included 46 compounds, into two phases. For the first section, we used 60% of theme as training and cross-validation data (D_{TR}) and 40% for test data (D_{TS}), and these data are shown in Tables IV.2 and IV.3, respectively. For each compound in this data, the experimental band gap energy is known. Consequently, we employ it to determine its correctness and explore the significance of the features and the proposed

approach. The unlabeled data, which comprises 266 compounds with unknown band gap energy, makes up the largest amount of the data in the second section. Since their prediction error cannot be computed, we provide the anticipated band gap energy prediction values.

Table.IV.2 Compounds used in training phase D_{TR} and their experimental band gap energy (Eg) (eV) [3].

N°	Compounds	Eg(eV)		Compounds	Eg(eV)
1	$CuAlS_2$	3.49	15	$CuGaS_2$	2.43
2	$AgAlS_2$	3.13	16	$AgGaS_2$	2.64
3	$CuAlSe_2$	2.67	17	$CuGaSe_2$	1.68
4	$AgAlSe_2$	2.55	18	$AgGaSe_2$	1.8
5	$CuAlTe_2$	2.06	19	ZnSiP2	2.07
6	$AgAlTe_2$	2.27	20	$ZnSiAs_2$	1.74
7	$CuGaTe_2$	1.12	21	$ZnGeP_2$	2.05
8	$AgGaTe_2$	1.32	22	$ZnGeAs_2$	1.15
9	$CuInS_2$	1.53	23	$CdSiP_2$	2.33
10	$AgInS_2$	1.87	24	$CdSiAs_2$	1.55
11	$CuInSe_2$	1.04	25	$CdGeP_2$	1.72
12	$AgInSe_2$	1.24	26	$CdGeAs_2$	0.57
13	$CuInTe_2$	1.06	27	$CdSnP_2$	1.17
14	$AgInTe_2$	0.95	28	$CdSnAs_2$	0.26

Table.IV.3 The Compounds used in the test phase D_{TS} and their experimental band gap energy Eg (eV) [3].

N	Compounds	Eg (eV)	N	Compounds	Eg (eV)
1	$GaInP_2$	2	10	$AuBSe_2$	1.53
2	$AgGaO_2$	4.1	11	$AuBTe_2$	1.33
3	$CuAlO_2$	3.5	12	$HgGeAs_2$	0.2
4	$CuBO_2$	2.2	13	$AgAlO_2$	3.6
5	$CuBS_2$	3.61	14	$CuInO_2$	3.9
6	$CuBSe_2$	3.13	15	$MgGeAs_2$	1.6
7	$CuGaO_2$	3.37	16	$MgSiAs_2$	2
8	$AgBO_2$	2.21	17	$MgSiP_2$	2.3
9	$AuBS_2$	2.55	18	$ZnGeN_2$	2.67

IV.2.6.Cross validation method

IV.2.6.1. Overview

One of the most popular techniques to evaluate a model's prediction performance is cross-validation, or CV. It is well known that using the same data for both algorithm training and statistical performance evaluation leads to an overly optimistic result. The initial idea behind CV's development was to address this problem by experimenting with new data to evaluate the algorithm's output and produce a more accurate assessment of its performance[13].Its basic idea is that we suppose have a dataset $D = \{(X_i, y_i), i = 1, \dots N\}$.We want to evaluate a regression model M in order to get a set of predicted values $\hat{y}_i$,i=1,..N. The dataset D may be divided into two sets:$D = D_1 \cup D_2$; D1 contains k data, whereas D_2 has N-k data. Using the dataset D_2 as the training set, we fit the model M; this process is often referred to as training the model. Next, we test the model using D1 as the test set and the trained model M to get predictions for observations $\hat{y}_{D1}$ given X_{D_1}. The data can be divided into $\binom{N}{k}$ possible divisions, and this operation can be carried out more than once. Applying our fitted function to an independent test set gives an estimate of the average generalization error, which is known as the CV estimate of error [14]. The choice among the many cross-validation techniques found in the literature is based on the regression model and the data. Given our small label data, we use the leave one out cross-validation technique (LOOCV).

IV.2.6.2.Leave-One-Out Cross Validation

The procedure known as leave-one-out cross validation (LOOCV), which we discuss in detail due to several significant linkages, is performed when k = 1 is utilized in the preceding formulation. In this instance, the model is trained and tested on every compound i(i=1….N) possible partition, where N is the total number of training data D_{TR}=28 compounds, and our test set always has cardinality 1. Finally, we calculate the overall errors MSCVE to be the overage of the N test. Figure.IV.4 illustrates the mechanism of LOOCV. Let $\hat{y}^{-i}$be the expected value of the missing observation for each i. The overage error estimate obtained via leave-one-out cross-validation is:

$$MSCVE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i^{-i}\right)^2 \quad \text{(IV.5)}$$

This is a special case of K-fold cross-validation in which the number of folds is the same as the number of observations. Although the first "official" declaration by Masteller and Tukey said that

this procedure pulls all information from provided data without additional reason, it appears that the overall principle of leave-one-out cross validation has been known for some time [15]. The LOOCV estimate is referred as the PRESS (prediction sum of squares) statistic in linear regression [14].

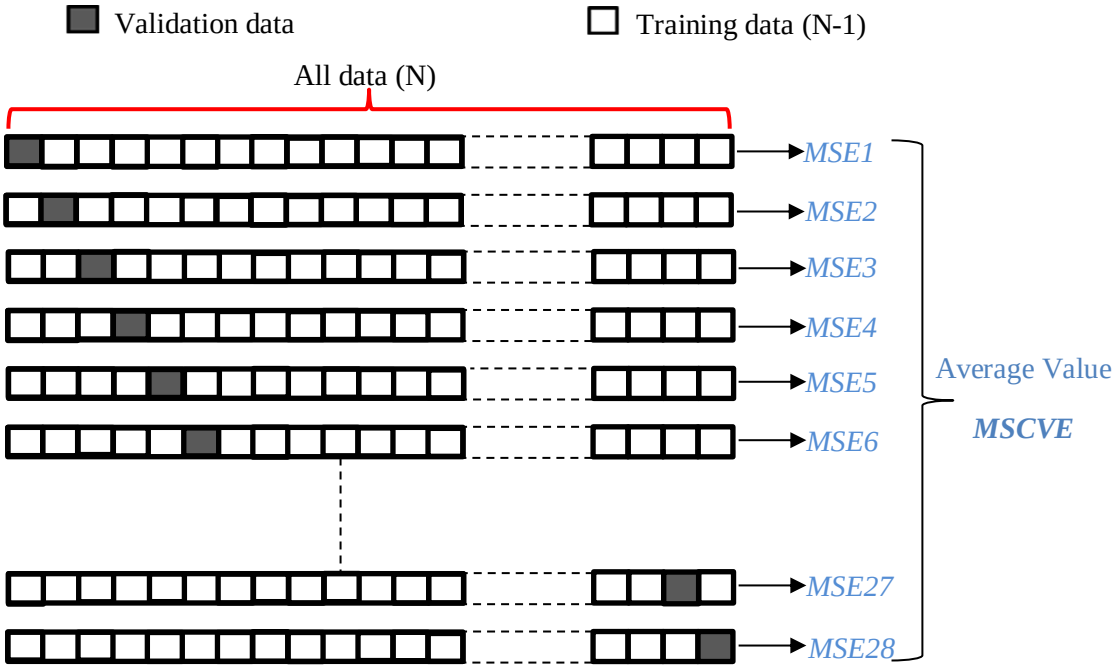

Figure.IV.4 Leave one out cross validation diagram.

IV.2.7.The effectiveness of the proposed approach's using the original data

In this section, we used the labeled data from tables IV.2 and IV.3 as a starting point for creating and evaluating the model. The original data set consisted of 28 features. We discussed the system phases in previous sections. Using the ANN-PSO as a basis for band gap energy prediction, Figures IV.5, IV.6, and IV.7 present the simulation result and output for each phase, respectively.

Figures.IV.5: presents a superposition between the experiment band gap and the output of the Primary-ANN prediction system.

Figure.IV.6: represents a superposition of the modeling error, the modeling error model and the final modeling error.

Figure.IV.7: presents a superposition between the experiment band gap and the output of the final model prediction

As shown in Figures IV.5, IV.6, IV.7, and IV.8, the final model is much better than the primary model, where we see that the final modeling error has been reduced, effectively improving the final model.

We can observe that there is a lack of good superposition between the output of the primary ANN-prediction system and the experimental band gap results by inspecting Fig. IV.5-a and its magnified version, Figure IV.5-b. Nevertheless, we see an improvement in the congruence of the two curves, as shown in Figure IV.7, when we include the recently created ANN-error model (Figure IV.6). Thus, it is clear from a visual inspection of Figures IV.5 and IV.7 that the output of the final prediction system is better than that of the main prediction system. This is confirmed by Figure IV.8, which shows that the primary error is smaller than the final error.

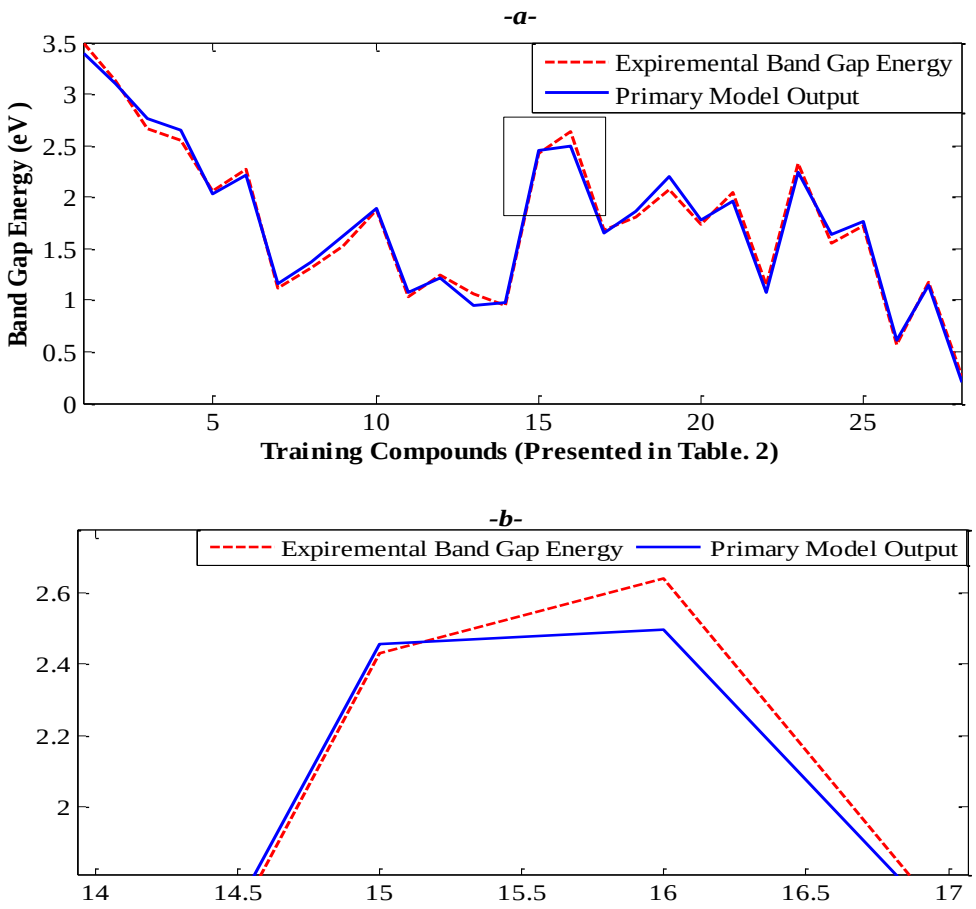

Figure.IV.5 PSO- ANN based primary prediction system: (a) Primary ANN-prediction system output , (b) Zoomed segment of primary ANN- prediction system output.

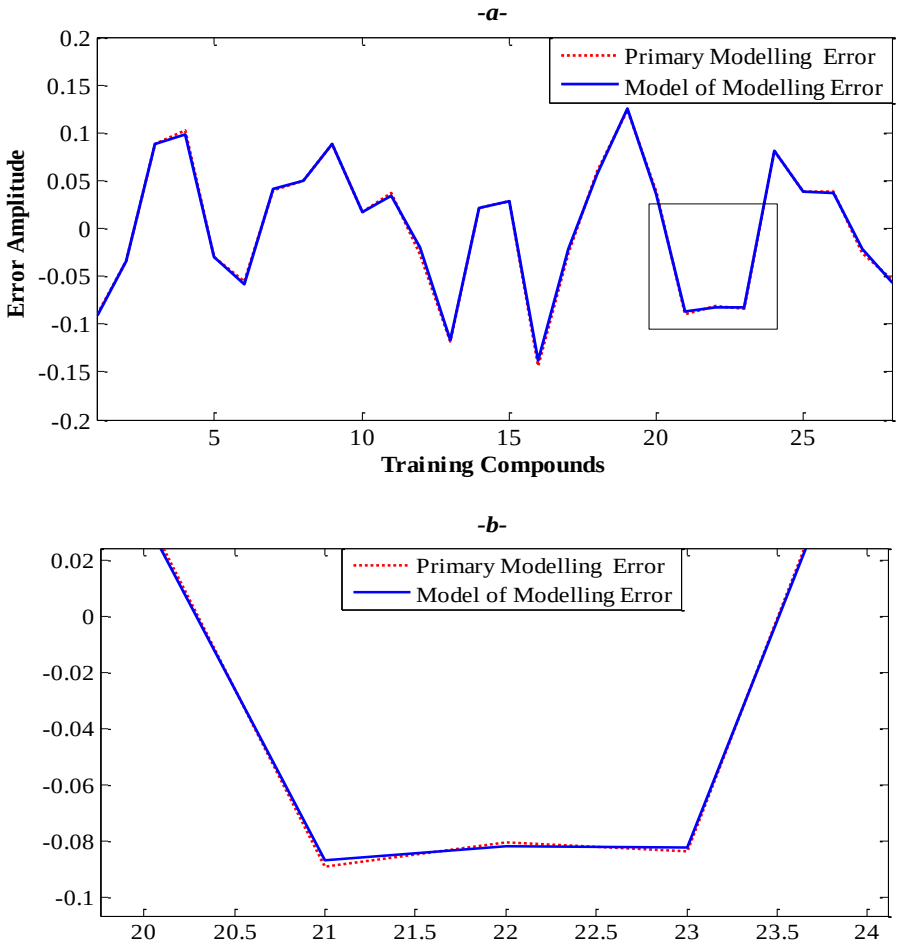

Figure.IV.6 PSO-based error process model: (a) Model of modelling error, (b) Zoomed segment of the model of modelling error.

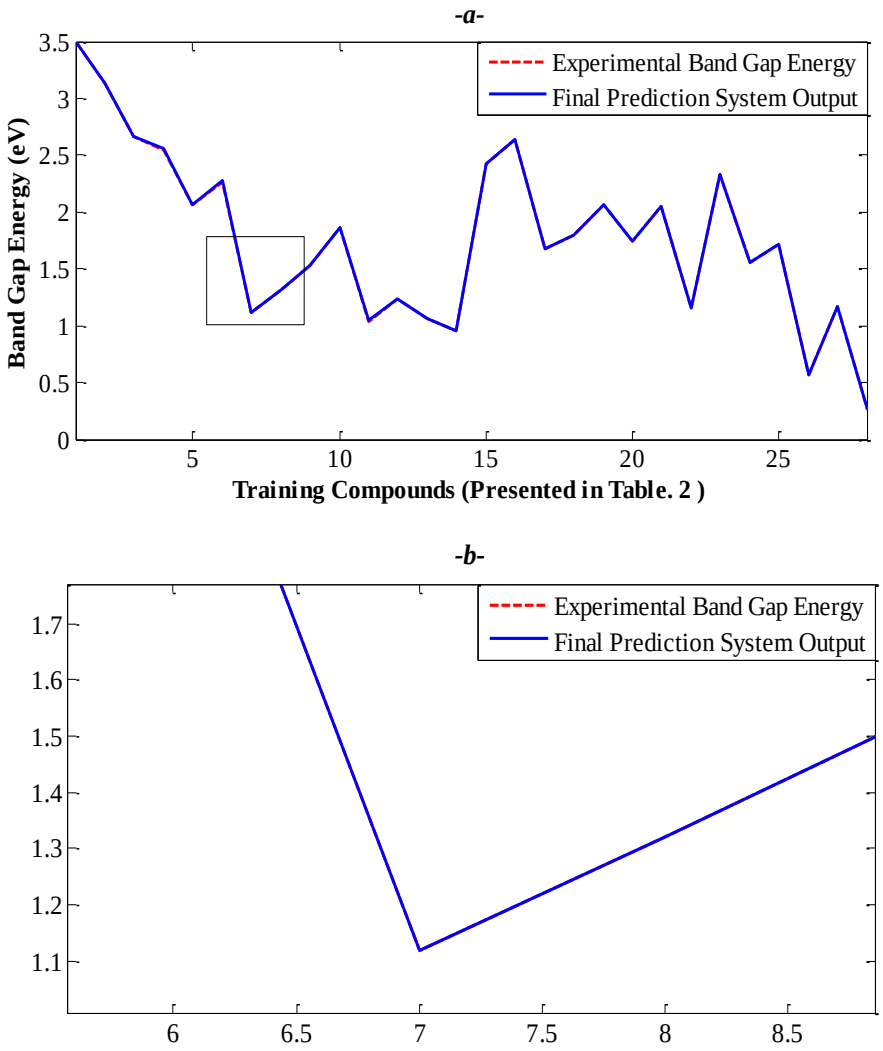

Figure.IV.7 PSO-based Final ANN-prediction system: (a) Final ANN-prediction system output, (b) Zoomed segment of the final ANN-prediction system output.

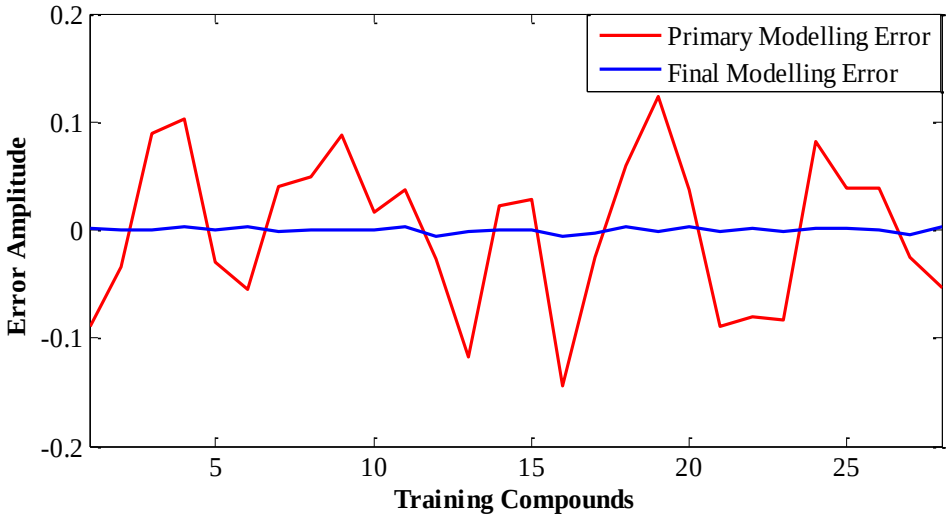

Figure.IV.8 Primary modelling error **Vs** Final modelling error.

Furthermore, to illustrate the significance of the added ANN-error model, we offer the regression curves without it (Figure IV.9-a) and with it (Figure IV.9-b), which display the prediction results for all training data. To assess the trained model's performance in terms of accuracy score, that is, the percentage of accurate predictions the model makes, we have employed the coefficient of determination (R^2), a metric that indicates how well a model predicts an outcome. The accuracy increases as it approaches 1. The coefficient of determination can be interpreted as follows:

- ✓ $R^2=0$ indicates that the model is not able to predict the result (low accuracy score);
- ✓ $0 < R^2 < 1$ denotes a partially predicted outcome by the model (Accuracy with Medium score);
- ✓ $R^2 \cong 1$, the model accurately predicts the result.(high accuracy score).

Thus, the accuracy score of the suggested prediction model was high, as shown in Figure IV.9-b, where the coefficient of determination for the training data was $R^2 = 0.9999$. The accuracy of the proposed technique suggests that the model's training data was sufficient to describe the band gap energy.

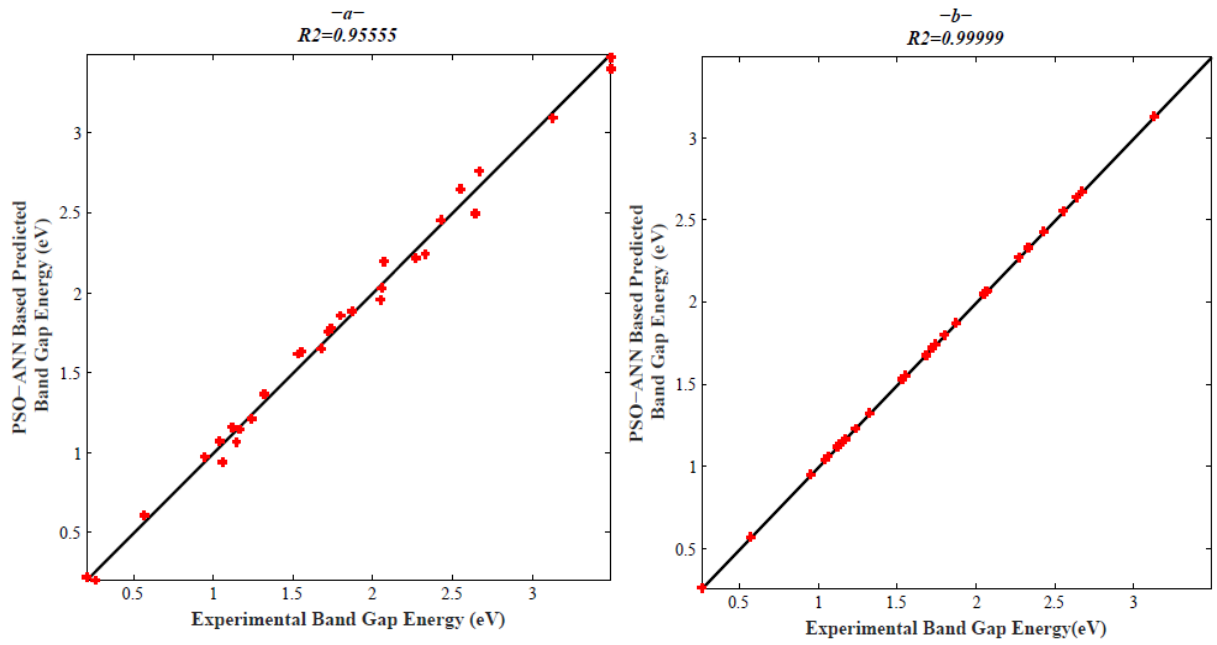

Figure.IV.9 The PSO-ANN base predicted Band Gaps Energy Vs Experimental band gap energy: (a)Without the ANN-error Model, (b) With the ANN-error Model.

IV.2.8. Comparative analysis and model evaluation

IV.2.8.1. The performance of our method on the basis of MSCVE

In the following analysis, we assess our system's efficiency quantitatively using the MSCVE criteria (Figure IV.4). To evaluate the significance of a subset of the features under consideration and the effectiveness of our model, as we said in the previous section, we use the training data D_{TR} (28 compounds) set. We employ the leave-one-out cross-validation procedure (LOOCV) due to our limited label data. Table IV.4 presents the average MSCVE for our approach compared to other benchmark approaches.

According to Table IV.4, it's evident that our suggested approach using the 28 features outperforms the other three benchmark techniques in terms of performance. We achieved a mean cross-validation error of 0.0454 eV^2. The exposure of "*a*" represents the significance of the features used in various studies, cited in Table IV.1 for each atoms A, B, and C. Thus, there are a total of 15 features, where b is the number of optimally selected features in [3].

Table.IV.4 Comparative analysis of the suggested approach and alternative benchmark methods based on the MSCVE(eV^2) criteria.

Models	**MSCVE (eV^2)**		
	OLS	PLS	LASSO
Suh C and rajan [11]	-	0.1849 [a]	-
Dey *et al* [12]	0.501 [a]	0.0619 [a]	0.0527 [a]
Khemisseia *et al* [3]	0.0731[a]	0.0750[a]	0.0658[a]
	0.0727[b]	0.0646[b]	0.0593[b]
		ANN-PSO	
Proposed method		0.0454	

[a]: MSCVE for the 15 features using in [3,11-12].

b: MSCVE for the 7 optimal features using by khmaissia et al [3].

IV.2.8.2.The performance of our method on the basis of MSE

In a second experiment, the regression model and the generalization of the chosen feature subset are learned using all the data D_{TR} including the 28 compounds. The data test, including 18 compounds D_{TS} that are not in the training set, will be used for validation.

For each substance tested in the data test, Table IV.5 presents the absolute errors between the anticipated band gap energy using the suggested approach and the experimental band gap energy. The range prediction error $\|predicted - experimental\|$ for solar cells must be less than 0.5 eV [3].

Table IV.6 shows that for 55% of the validation data, the expected error is less than 0.5 eV. This suggests that because of the training data's small size, the training algorithm does not fully suit the training set. However, compared to earlier research, its performance is superior; the final raw value displays the average squared error (MSE) of the prediction for each of the 18 test compounds. It was lesser than other works reported in [3, 12].

Table.IV.5 Absolute errors of predicted band gap energy values (eV) and MSE (eV^2) for each of the 18 D_{Ts} compounds, bolded values display errors within approximately a 0.5 eV interval for each of them.

Compounds	Experimental band gaps value (eV)	$\|\mathbf{predicted - experimental}\|$ (eV)
$GaInP_2$	2	**0.24**
$AgGaO_2$	4.1	1.25
$CuAlO_2$	3.5	**0.31**
$CuBO_2$	2.2	0.98
$CuBS_2$	3.61	**0.30**
$CuBSe_2$	3.13	**0.14**
$CuGaO_2$	3.37	**0.20**
$AgBO_2$	2.21	1.02
$AuBS_2$	2.55	**0.27**
$AuBSe_2$	1.53	1.04
$AuBTe_2$	1.33	**0.04**
$HgGeAs_2$	0.2	0.61
$AgAlO_2$	3.6	**0.31**
$CuInO_2$	3.9	0.88
$MgGeAs_2$	1.6	0.63
$MgSiAs_2$	2	**0.55**
$MgSiP_2$	2.3	**0.25**
$ZnGeN_2$	2.67	0.93
MSE(eV^2)	-	**0.50**

IV.2.8.3.The performance of our method using the features determined by the various filter selection techniques in [3] study.

Following that, we make use of the selected features determined by the various filter selection techniques provided in the earlier study [3], which are indicated in Table IV.6, to compare the performance of the current work with the performance of the previous model in terms of the MSE criteria. The suggested approach outperforms the regression techniques used by

Khmaissia et al. [6]. All six-filter selection procedures demonstrate a reduction in the mean squared projected errors of the validation data for each selected feature. When a subset of relevant features is selected using the correlation criterion approach, the lowest squared error average is reached. However, the prediction accuracy did not significantly improve, possibly due to the omission of the properties of the second atom in the third element of ABC_2 compounds. This is another theory that demonstrates the differences between the ABC and ABC_2 structures and strengthens the correlation between inputs and outputs. All of the features of each atom in the compounds under consideration must be defined. Note that the B and C elements are the most important features [12, 3].

Table.IV.6 Performance comparison between the developed system and Khmaissia et al. system [3] using the same data in their work.

Filter selection technique	Selected features	MSE (eV^2)			
		Regression methods			
		OLS[3]	PLS[3]	LASSO[3]	Proposed method
15 Features	AN(B), PR(B), AN(C), EN(C), PR(C), MP(B), MP(C), MP(A), PR(A), AN(A),VL(B),VL(C), EN(B), EN(A), VL(A).	19.63	1.05	0.79	0.73
Correlation	AN(B), PR(B), AN(C), EN(C), PR(C), MP(C), MP(A)	0.66	0.66	0.63	0.61
Lasso regularization	AN(B), PR(B), AN(C), EN(C), PR(A), MP(B), MP(C), MP(A), PR(C), AN(A),VL(B),VL(C),	1.32	2.06	1.32	1.21
SFS-OLS	AN(B), PR(B), AN(C), EN(C), PR(C), MP(C), VL(B) ,VL(C), VL(A)	1.23	-	-	0.75
SFS-PLS	AN(B), MP(B), EN(C), PR(C), MP(C), EN(A), AN(C),	-	0.99	-	0.71
SFS-Lasso	AN(B), PR(B), AN(C), EN(C), PR(C), MP(C), VL(B) ,VL(C), VL(A)	-	-	1.32	1.11

IV.2.9.The predicted band gap for unlabeled data

Our approach trained 60% of the labeled data in all previous experiments, using the remaining 40% for validation and evaluation. In order to get further insight into the training model and test the trained prediction model with 266 unlabeled compounds, we will use all of the labeled data for 46 compounds in this field. These investigations' inability to confirm the accuracy of the estimated band gap energy is a limitation. It simply enumerates the values that the

suggested system predicts. We believe that a reliable measure of the anticipated values' validity would be the consistency of the expected values as determined by the feature set and regression model. Other researchers will have access to these anticipated band gaps for further analysis and validation. According to the results referred to in Tables IV.7 and 8, given that the band gap energy cannot be negative, any negative values in the table should be deemed invalid, indicating poor prediction by the system. Consequently, the regression models learned are unable to accurately predict significant band gaps for all compounds. This suggests that our model requires more labeled training data to enhance its robustness and learning capability. Compared to the results found in [3,11–12], as shown in Table IV.9, the new ANN-PSO model only shows 3% of the 266 compounds it was made for having negative band gaps. Thus, it's reasonable to conclude that the proposed approach has surpassed previous models, reducing prediction errors, albeit still yielding negative values in some instances.

Table.IV. 7 The predicted band gap energy values (eV) for the 266 compounds.

Compounds	Eg (eV)	Compounds	Eg (eV)	Compounds	Eg (eV)	Compounds	Eg (eV)	Compounds	Eg (eV)

$AgBS_2$	3.40	$BeCP_2$	0.82	$MgPbBi_2$	2.51	$ZnPbSb_2$	0.95	$SrSnN_2$	4.52
$AgBSe_2$	2.65	$BeCSb_2$	-0.69	$MgPbN_2$	1.93	$ZnSiBi_2$	1.34	$SrSnP_2$	2.98
$AgBTe_2$	1.19	$BeGeAs_2$	0.35	MgPbP2	2.14	$ZnSiN_2$	2.19	$SrSnAs_2$	1.61
$AgTlS_2$	1.31	$BeGeBi_2$	1.00	$MgPbSb_2$	2.57	$ZnSnBi_2$	0.99	$SrSnSb_2$	0.95
$AgTlSe_2$	1.30	$BeGeN_2$	3.53	$BePbP_2$	1.54	$ZnSnN_2$	3.47	$SrSnBi_2$	0.77
$AgTlTe_2$	1.50	$BeGeSb_2$	1.02	$BePbSb_2$	2.83	$ZnSnSb_2$	0.67	$SrPbN_2$	3.53
$AuAlS_2$	2.95	$BePbAs_2$	1.62	$BeSiBi_2$	1.19	$ZnSnAs_2$	0.15	$SrPbP_2$	2.67
$AuAlSe_2$	2.88	$BePbBi_2$	2.68	$BeSiN_2$	3.50	$ZnSnP_2$	1.43	$SrPbAs_2$	2.43
$AuAlTe_2$	1.35	$BePbN_2$	1.36	$BeSiSb_2$	0.71	$AgBPo_2$	1.08	$SrPbSb_2$	2.11
$AuBO_2$	2.55	$CdSnN_2$	3.63	$BeSnAs_2$	1.84	$CaCN_2$	1.38	$SrPbBi_2$	1.36
$CuTlS_2$	1.49	$CdSnSb_2$	0.23	$BeSnBi_2$	1.35	$CaCP_2$	0.97	$BaCN_2$	0.89
$CuTlSe_2$	1.28	$CuAlPo_2$	1.69	$BeSnN_2$	3.24	$CaCAs_2$	1.78	$BaCP_2$	1.28
$CuTlTe_2$	1.06	$CuBPo_2$	1.31	$BeSnP_2$	1.39	$CaCSb_2$	1.37	$BaCAs_2$	1.60
$AuInS_2$	1.59	$CuInPo_2$	0.93	$BeSnSb_2$	1.65	$CaCBi_2$	1.16	$BaCSb_2$	0.90
$AuInSe_2$	1.18	$CuTlPo_2$	1.02	$CdCAs_2$	0.72	$CaSiN_2$	3.39	$BaCBi_2$	0.91
$AuInTe_2$	1.34	$HgCAs_2$	0.72	$CdCBi_2$	0.71	$CaSiP_2$	2.84	$BaSiN_2$	3.32
$CuBTe_2$	1.44	$HgCBi_2$	-0.26	$CdCN_2$	1.62	$CaSiAs_2$	1.95	$BaSiP_2$	2.27
$CuTlO_2$	3.28	$HgCN_2$	1.09	$CdCP_2$	1.09	$CaSiSb_2$	1.53	$BaSiAs_2$	1.97
$AuTlTe_2$	1.00	$HgCP_2$	0.66	$CdCSb_2$	0.25	$CaSiBi_2$	1.27	$BaSiSb_2$	1.03
$AgGaPo_2$	0.99	$HgCSb_2$	-0.29	$CdGeBi_2$	0.32	$CaGeN_2$	3.65	$BaSiBi_2$	0.99
$AgInO_2$	4.40	$HgGeBi_2$	0.41	$CdGeN_2$	2.86	$CaGeP_2$	3.22	$BaGeN_2$	3.98
$AgInPo_2$	1.36	$HgGeN_2$	3.06	$CdGeSb_2$	0.69	$CaGeAs_2$	1.55	$BaGeP_2$	2.49
$AgTlO_2$	3.35	$HgGeP_2$	1.23	$CdPbAs_2$	0.36	$CaGeSb_2$	0.75	$BaGeAs_2$	1.17
$AgTlPo_2$	1.08	$HgGeSb_2$	-0.35	$CdPbBi_2$	1.49	$CaGeBi_2$	0.57	$BaGeSb_2$	0.52
$AgAlPo_2$	1.11	$HgPbAs_2$	0.66	$CdPbN_2$	2.56	$CaSnN_2$	3.94	$BaGeBi_2$	0.48
$AuAlO_2$	3.52	$HgPbBi_2$	1.36	$CdPbP_2$	0.63	$CaSnP_2$	2.72	$BaSnN_2$	3.97
$AuAlPo_2$	0.60	$HgPbN_2$	2.04	$CdPbSb_2$	0.77	$CaSnAs_2$	2.22	$BaSnP_2$	2.91
$AuBPo_2$	1.29	$HgPbP_2$	1.14	$CdSiBi_2$	1.47	$CaSnSb_2$	1.46	$BaSnAs_2$	1.60
$AuGaO_2$	3.43	$HgPbSb_2$	1.07	$CdSiN_2$	2.18	$CaSnBi_2$	0.50	$BaSnSb_2$	0.97
$AuGaPo_2$	0.78	$HgSiAs_2$	1.15	$CdSiSb_2$	1.49	$CaPbN_2$	3.65	$BaSnBi_2$	1.13
$AuGaS_2$	2.26	$HgSiBi_2$	0.69	$CdSnBi_2$	1.22	$CaPbP_2$	2.62	$BaPbN_2$	2.92
$AuGaSe_2$	1.72	$HgSiN_2$	2.39	$MgSiBi_2$	0.67	$CaPbAs_2$	2.32	$BaPbP_2$	2.78
$AuInO_2$	4.34	$HgSiSb_2$	0.53	$MgSiSb_2$	1.24	$CaPbBi_2$	1.54	$BaPbSb_2$	2.37
$AuInPo_2$	0.91	$HgSnAs_2$	0.20	$MgSnAs_2$	2.29	$SrCN_2$	0.92	$BaPbBi_2$	0.81
$AuTlPo_2$	0.81	$HgSnBi_2$	0.46	$MgSnBi_2$	1.20	$SrCP_2$	0.94	$RaCN_2$	0.83
$Be_{0.5}C_{0.5}Sb$	-0.74	$HgSnN_2$	3.37	$MgSnN_2$	3.39	$SrCAs_2$	1.76	$RaCP_2$	0.98
$AuGaTe_2$	1.85	$HgSiP_2$	2.66	$MgSiN_2$	3.97	$CaPbSb_2$	3.08	$BaPbAs_2$	3.00

Table.IV.8 Continued.

Compounds	Eg (eV)	Compounds	Eg (eV)	Compounds	Eg (eV)	Compounds	Eg (eV)	Compounds	Eg (eV)
$Be_{0.5}Ge_{0.5}As$	0.21	$HgSnP_2$	0.59	$MgSnP_2$	1.73	$SrCSb_2$	0.89	$RaCAs_2$	0.84
$Be_{0.5}Ge_{0.5}P$	1.65	$HgSnSb_2$	0.88	$MgSnSb_2$	1.99	$SrCBi_2$	1.18	$RaCSb_2$	0.74
$Be_{0.5}Ge_{0.5}Sb$	0.68	$MgCAs_2$	0.28	$ZnCAs_2$	0.90	$SrSiN_2$	3.34	$RaCBi_2$	0.74
$Be_{0.5}Si_{0.5}As$	1.08	$MgCBi_2$	0.57	$ZnCBi_2$	0.44	$SrSiP_2$	2.52	$RaSiN_2$	3.03
$Be_{0.5}Si_{0.5}P$	2.05	$MgCN_2$	0.89	$ZnCN_2$	1.91	$SrSiAs_2$	1.88	$RaSiP_2$	2.32
$Be_{0.5}Si_{0.5}Sb$	1.11	$MgCP_2$	0.58	$ZnCP_2$	1.58	$SrSiSb_2$	1.54	$RaSiAs_2$	1.96
$Be_{0.5}Sn_{0.5}As$	1.44	$MgCSb_2$	0.22	$ZnGeBi_2$	0.37	$SrSiBi_2$	1.12	$RaSiSb_2$	0.72
$Be_{0.5}Sn_{0.5}P$	1.53	$MgGeBi_2$	0.76	ZnGeSb2	0.59	$SrGeN_2$	3.83	$RaSiBi_2$	0.79
$Be_{0.5}Sn_{0.5}Sb$	1.72	$MgGeN_2$	3.35	$ZnPbAs_2$	0.16	$SrGeP_2$	2.50	$RaGeN_2$	3.51
$BeCAs_2$	-0.19	$MgGeP_2$	1.84	$ZnPbBi_2$	1.44	$SrGeAs_2$	1.14	$RaGeP_2$	1.85
$BeCBi_2$	-0.46	$MgGeSb_2$	1.23	$ZnPbN_2$	2.72	$SrGeSb_2$	0.58	$RaGeAs_2$	1.15
$BeCN_2$	2.04	$MgPbAs_2$	2.78	$ZnPbP_2$	1.00	$SrGeBi_2$	0.56	$RaGeSb_2$	0.65
$RaGeBi_2$	0.92	$RaSnSb_2$	1.45	$RaPbAs_2$	2.07	$BeSiAs_2$	1.02	$AuTlSe_2$	1.07
$RaSnN_2$	4.21	$RaSnBi_2$	1.36	$RaPbSb_2$	2.32	$BeGeP_2$	1.45	$CuGaPo_2$	1.17
$RaSnP_2$	2.53	$RaPbN_2$	2.66	$RaPbBi_2$	1.05	$AuTlO_2$	3.04	$ZnCSb_2$	0.31
$RaSnAs_2$	1.09	$RaPbP_2$	2.91	$BeSiP_2$	2.28	$AuTlS_2$	2.03	$ZnSiSb_2$	1.30
$CdGeBi_2$	0.61								

Table.IV.9. The Performance comparison of the developed models and the existing models using the negative number Band gap percentage .

New compounds	The negatives numbers Band gap percentage						
	LASSO		OLS		PLS		Our approach
	a	B	a	B	a	b	28 features
205 compounds [11]	-	-	-	-	37%	-	
202 compounds[12]	18%	-	22	-	21	-	-
156 compounds[3]	31%	18%	34%	21%	34%	18%	-
266 compounds	-		-		-		3%

a:Refers to the 15 features used in previous works [3,11-12].

b:Refers to the optimal features relate with band gap used in khmaissia models [3].

IV.3. Perovskite Lattice Constant Prediction: simple and double perovskites

IV.3.1. Perspective

This study suggests the right models to make lattice parameter predictions more accurate and useful for cubic and pseudo-cubic ABC_3 perovskite and double perovskite $A_2BB'O_6$ structures. Firstly, we predict the lattice parameters of cubic and pseudo-cubic ABC_3 perovskite structures using a combination of metaheuristic algorithms with artificial neural networks (ANN) and fuzzy logic methods, which have proven to be effective in a wide range of fields [16–18]. Hence, they mimic natural processes like genetic mutation, natural selection, or animal behavior to search the search space and identify the best answer. Additionally, the population-based methodology of these algorithms allows for the examination of a wide variety of behaviors, increasing the likelihood of finding the most suitable solution. In our studies, we proposed particle swarm optimization [19], invasive weed optimization [20], and the imperialist competitive algorithm [21] as optimization algorithms for the artificial neural network (ANN) and fuzzy logic computational intelligence methods. Secondly, we used the extreme learning machine method to predict the lattice constant for the $A_2BB'O_6$ double perovskite.

IV.3.2.The simple perovskites ABC_3 lattice constant prediction

IV.3.2.1.Data description

The first section of our second application uses the ABC_3 cubic and pseudo-cubic simple perovskites as the consideration materials for lattice constant prediction; as descripted in Section I.4.1, A and B stand for alkaline earth (Ca, Ba, Sr, etc.) and transition metal (Fe, Ti, Ni, etc.), respectively. The consideration elements are cited in the periodic table from Chapter I (Figure I.8**),** which created a total of 127 compounds. The 6-fold effective ionic radii (rA, rB, and rC) of each atom A, B, and C, the tolerance factor (t) (Eq.I.6), and their corresponding experimental lattice parameters (LC in Å) comprise the dataset utilized in this work for modeling and simulation. The data presented here has been collected from earlier research, as reported in [22–24].Based on these descriptors, a correlation between the lattices constant (LC) and the studied data may be established, as the function below illustrates:

$$LC = f\left(r_{A,B,3C}, t\right) \tag{IV.8}$$

The three atoms are represented by A, B, and C. As we mentioned in our first application, to differentiate between ABC_3 and ABC, we used the properties specific to each atom. Hence, this is the total of six descriptors for each compound.

IV.3.2.2.Contribution

The lattice constant is an important parameter in the cubic perovskite structure that is used to identify the material's crystal structure as well as for the creation of materials for some technological applications. That's why this research offers effective methods for predicting the lattice constants of cubic and pseudo-cubic perovskites using fuzzy logic and artificial neural networks (ANNs).

This work primarily contributes to the development of a new general framework for lattice parameter prediction by combining ANN/fuzzy-based computational intelligence (CI) approaches with metaheuristic approaches. This allows for the comparison of the two approaches based on the accuracy of the lattice constant prediction for the materials under consideration. This work consists of two phases. The first phase involves comparing the effectiveness of many metaheuristic methods for adjusting the ANN model's parameters in order to identify the optimal optimization strategy. The ANN and fuzzy logic models are compared in the second phase, where we combine the optimal optimization method found in the first phase with the fuzzy logic technique to predict the lattice parameters.

IV.3.2.3.Proposed approaches

The aim of this section is to assess the efficiency of employing artificial neural networks and fuzzy logic, along with optimization algorithms, to accelerate the identification of novel materials with greater accuracy. Specifically, the focus is on predicting the lattice constants of cubic and pseudo-cubic perovskites.

Among the various available architectures of ANN, in this thesis we use the feedforward neural networks (FNNs) which are explained in section II.5.2.1. Generally, optimal FNNs may be viewed as: optimal weights, optimal hidden layers, optimal hidden neurons, and optimal learning algorithm, and so on. The important purpose of optimizing FNN is to enhance its generalized performance. This study aims to cover a wide range of FNN optimization approaches by using different optimization algorithms.

On the other hand, A fuzzy model can represent any unknown system or function, $y = f(x)$ using input/output data $x_k = [x_{1k}, x_{2k}, \dots x_{nk}]^T$ and y_k. The index k denotes the number of samples and n the dimension of the vector (the number of regressors). The fuzzy model used in this thesis is the type-1 fuzzy system discussed in Section II.6, the idea is to find a relationship between the input and output of the fuzzy system to progressively adjust its parameters by using the optimal adjustment mechanism (optimization algorithm) found in the optimization of FNN. To achieve the desired objective, our study consists of three stages:

- ✓ **First stage:** Data preprocessing.
- ✓ **Second stage:** ANN prediction model-based metaheuristics optimizations algorithms.
- ✓ **Third stage:** Fuzzy prediction model-based the optimal optimization algorithm.

In order to determine the ideal ANN parameter, the study's first part evaluated the performance of many metaheuristic algorithms in adjusting the parameters of the ANN prediction structure in such a way that the objective function (discussed in Section IV.2) reaches its minimum. In the second part, we combined the optimal optimization algorithm with the fuzzy logic methodology and compared the efficacy of the two methods—ANN and fuzzy logic—using various metrics.

IV.3.2.3.1.Data preprocessing

Pre-processing and normalizing the data is essential before including it into developed models. This requires converting features or data to a common scale, which is a crucial step that might improve the model's training stability and performance. As in the first application, equation (IV.3) has been utilized to pre-process and normalize all of the data used for modeling and simulation to the range [0, 1].

For additional analysis, the pre-processed and normalized data is usually divided into three groups (Figure. IV.10). In particular, 80% of the data is kept for model training, 10% is used for testing, and the remaining 10% is kept for validation. In machine learning, it is standard procedure to divide the data into several sets so that the models may be tested and trained on various sets.

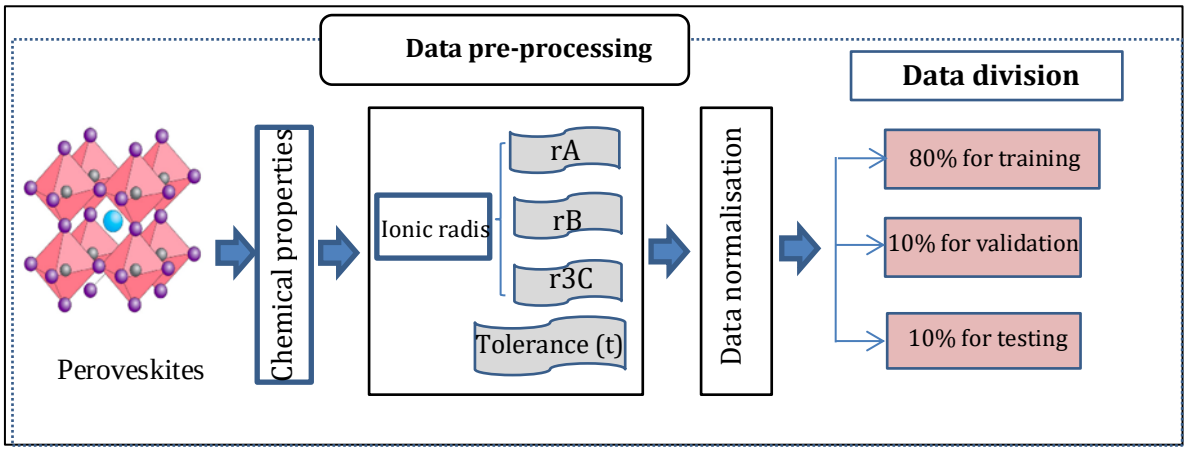

Figure.IV.10 Data description and preprocessing.

IV.3.2.3.2.ANN prediction model-based metaheuristics optimizations algorithms

The ANN-prediction model is determined by the input-output data set; its strategy is essentially an online adaptation of this model. The PSO, IWO, or ICA algorithms, which are the optimization algorithms shown in Figure IV.11, are used to adjust the model's parameters in order to minimize the error between the experimental lattice constant and the estimated lattice constant, which becomes a cost function for the optimization algorithms that are applied.

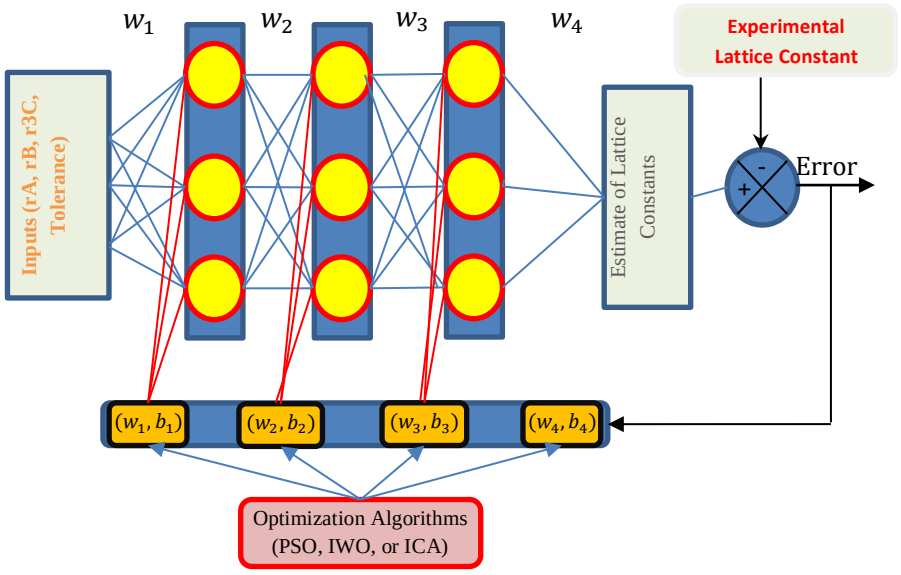

Figure.IV.11 Proposed ANN prediction model.

IV.3.2.3.3.Fuzzy prediction model-based metaheuristics optimizations algorithms

To construct a fuzzy system prediction model, one must identify and ascertain the values of the unknown parameters in both the antecedent and consequent sections of the fuzzy if-then rules (Eq. (II.11)). The antecedent portion divides the input space into a number of fuzzy regions, and the subsequent part automatically designs the system behavior inside those regions. Uncertainties in Gaussian type-1 fuzzy sets can be linked to the standard deviation (STD) and the mean (center). Gaussian membership functions were used for this work because they are useful for universal approximation and can uniformly estimate continuous functions [25]. We can express the Gaussian MF mathematically as follows:

$$\mu(x) = exp\left(\frac{1}{2}\frac{(x-c)^2}{\sigma^2}\right); \quad \text{(IV.9)}$$

Where x is the input vector and c and σ are the membership function's center and width. In this work, type-1 fuzzy sets with a fixed STD σ are created by just taking into account the uncertainty on the mean c.

The fuzzy-prediction model is determined by the input-output data set; this model's strategy is mostly an online adaption of it as in ANN model. The model's parameters are adjusted using the optimal optimization algorithm which found in the first part (see section IV.2.3.3.2.) As shown in Figure IV.12, so that the cost functions reaches a minimal value.

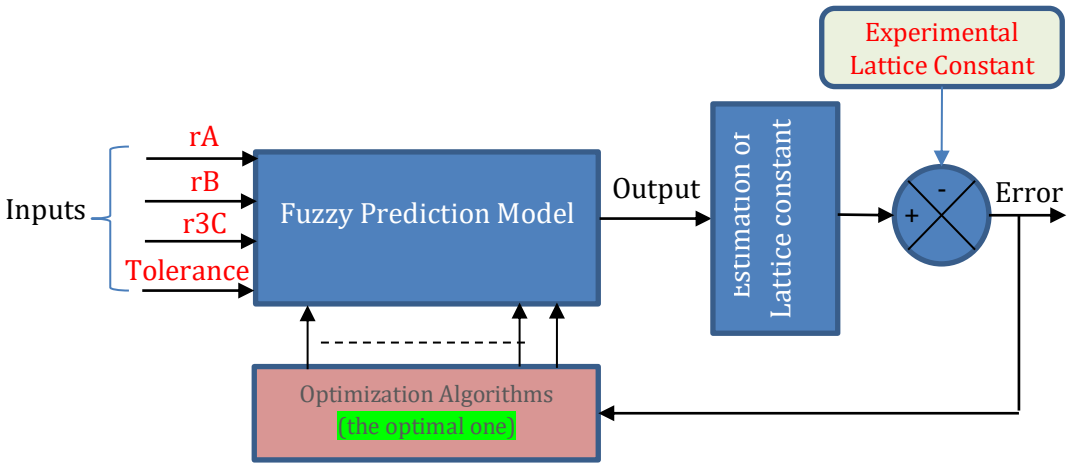

Figure.IV.12 Proposed Fuzzy prediction model.

IV.3.2.4.The evaluation criteria

The performance of the prediction models was assessed in this work using three widely used regression indicators. Among these measures are the mean absolute error (MAE), the root mean square error (RMSE), and the coefficient of determination (R^2). The average size of the mistakes in the collection of predictions is represented by the MAE, as indicated by Eq. (IV.10). However, the RMSE, as demonstrated by Eq. (IV.11), is helpful in penalizing large errors, which makes it especially helpful in situations when large errors are undesirable. Lastly, the R^2 statistic, which is based on percentage and can be found in Eq. (IV.12), indicates how well the data fits a regression line. Lower values of the MAE and RMSE scores, which are negatively orientated, signify better model predictions in the variable's units (in this case, Angstrom or Å). It is significant that these measures are frequently employed in statistical modeling and machine learning to evaluate prediction models.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad \text{(IV.10)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{(IV.11)}$$

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}\right) \quad \text{(IV.12)}$$

Where $\bar{y}_i$ is the mean value with summation over *n* number of samples, i.e. $i = 1, 2, \ldots, n$.

IV.3.2.5.Results and Discussion

In this section, the simulation results of our proposed approaches to estimating the lattice constants of cubic and pseudo-cubic perovskite will be shown and discussed. Both prediction and optimization performance are assessed in order to evaluate how well these strategies perform. First, the 127 data set are split into three sets, using an 8:1:1 ratio for the training, testing, and validation sets, respectively. Random division of data is done to avoid selection bias by creating different training, testing, and validation sets. Random selection ensures that the data used for training, testing, and validations are representative of the whole dataset, which helps to increase the models' ability to generalize to new data. This is accomplished through the use of the cross-validation technique, a popular method for evaluating the predictive performance of a regression mode. Test-set cross-validation is the cross-validation technique employed in this study, which

allows one to create a regression function using training data and assess the model's predictive capability and future generalizability using testing data. This validation method is simpler and has less computational complexity than the laborious leave-one-out cross-validation. Therefore, 103 compounds were used to train the model, 12 compounds were used to adjust the hyper parameters, and the remaining compounds were used to validate the model. Table IV.15and 16. Lists all the data used in different phases.

IV.3.2.5.1. Determination of metaheuristics algorithms parameters

Metaheuristic algorithms employ empirical testing and domain knowledge to determine acceptable values for their parameters. The technique involves iteratively adjusting the parameters to improve performance on a given issue. Trial and error was used to choose these parameters for the present section. One of the most important parameters influencing the optimal solution is population size. Although a larger population results in faster calculations, it also allows for better exploration of the solution space. Every algorithm was evaluated with a population size of 50 to guarantee an unbiased assessment. The settings of the algorithmic control were adjusted in accordance with Table IV.10.

Table.IV.10 Parameters setting of the used optimization algorithms.

Algorithms	**PSO**	**IWO**	**ICA**
Parameters Values	-Maximum number of populations $nPop = 50$; -The personal and social cognizance coefficients $C_1 = C_2 = 2$ - Inertia factor $w = 0.75$	-Maximum number of populations $nPop = 50$. -Maximum Number of Seeds $S_{max} = 5$; -Initial and final Value of Standard Deviation: [0.5;-0.001]. - Variance Reduction Exponent=1.5.	-Maximum number of populations $nPop = 50$. -Number of Empires / Imperialists $nEMP = 1$. -Selection Pressure $alpha = 0.5$. -Assimilation Coefficient $beta = 1$. -Colonies Mean Cost Coefficient $zeta = 0.25$. -Revolution Rate= 0.2 -Probability Revolution = 0.05

IV.3.2.5.2. ANN prediction model evaluation

The model created utilizing an artificial neural network combined with PSO, IWO, and ICA is shown in this section. This analysis's goal is to assess each optimization algorithm's capacity to converge to ideal solutions. We used the ANN model to estimate the lattice constant in the subsequent example. We considered three possible evaluations, and in each, we used the consideration methods of optimization (PSO, IWO, or ICA). Figures IV.13 through IV.21 display the simulation findings in relation to PSO, IWO, and ICA. First, the artificial neural network model is constructed using 80% of the data as a training set listed in Tables IV.15and 16. The training procedure was performed using 500 epochs for every case. For all three cases of optimization techniques, the MSE is shown as a function of epoch number in Figure IV.13. Figure IV.13 shows that each MSE for the PSO-ANN, ICA-ANN, and IWO-ANN models saw a sharp drop at the beginning of the training process (rapid training), and the training concluded after approximately 100, 200, and 350 epochs, respectively.

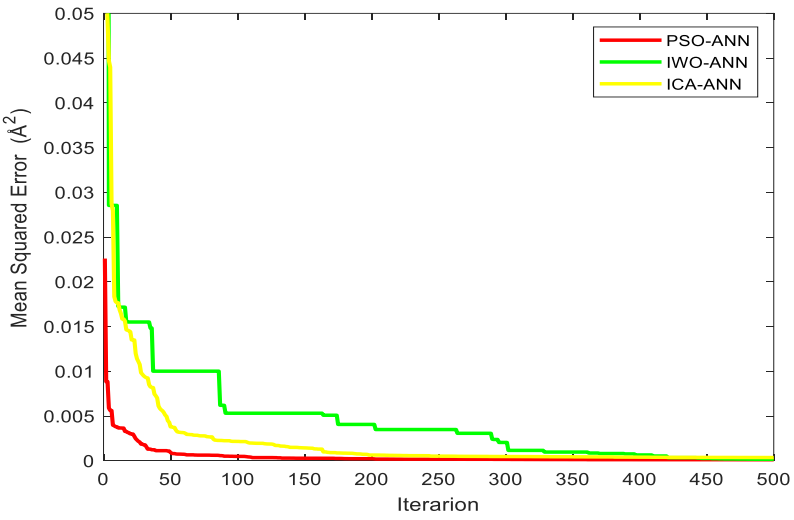

Figure.IV.13 The MSE vs Iteration for PSO-ANN, IWO-ANN and ICA-ANN.

IV.3.2.5.3.The evaluation of the optimization algorithms based on accuracy metrics

We used the evaluation criteria outlined in Section IV.3.2.4 to quantitatively assess the effectiveness of the optimization technique. The numerical results for the three optimization techniques that were utilized at each stage (training, testing, and validation) are shown in Table IV.11. Because the dataset was extremely simple, all of the methods performed similarly, as

Table IV.11 illustrate. As a result, it is impossible to conclude which optimization technique is better than the others based on these results.

Table.IV.11 Evaluations of the predicted effectiveness of the created ANN-based metaheuristics optimization methods.

Prediction Model	Data sets	MSE	$RMSE$	MAE	R^2
ANN-PSO	Training	$1.1712e^{-04}$	0.0108	0.0085	0.9997
	Validation	$1.3459e^{-04}$	0.0116	0.0095	0.9981
	Tests	$1.3672e^{-04}$	0.0117	0.0104	0.9971
ANN-IWO	Training	$1.5878e^{-04}$	0.0126	0.0098	0.9997
	Validation	$1.6776e^{-04}$	0.0130	0.0112	0.9981
	Tests	$2.0331e^{-04}$	0.0143	0.0115	0.9951
ANN-ICA	Training	$3.5980e^{-04}$	0.0190	0.0113	0.9993
	Validation	$5.5330e^{-04}$	0.0235	0.0198	0.9944
	Tests	$8.2431e^{-04}$	0.0287	0.0214	0.9934

In order to make it easier to perform, additional analysis was employed. First, we've provided Figures IV. 14, 15, and 16, which, for each of the three possible situations of the optimization techniques utilized, display a correlation-cross plot between the experimental lattice parameter and the value predicted by the created ANN model.

Plotting one set of data against the other to illustrate the link between the two is known as a cross-correlation plot. Pattern recognition and connection analysis are widely applied fields in signal processing and data analysis. The data points are shown on a two-dimensional graph in a cross-correlation plot, with one variable on the x-axis and the second variable on the y-axis. If there is a positive or negative correlation between the two variables, it may be seen in the resultant pattern. The effectiveness of the ANN model and the applied optimization procedures is confirmed by the positive and strong correlations we can see between the experimental lattice constant and the projected value based on Figures IV. 14, 15, and 16. However, these graphs do not allow us to conclude which optimization technique is better than the others.

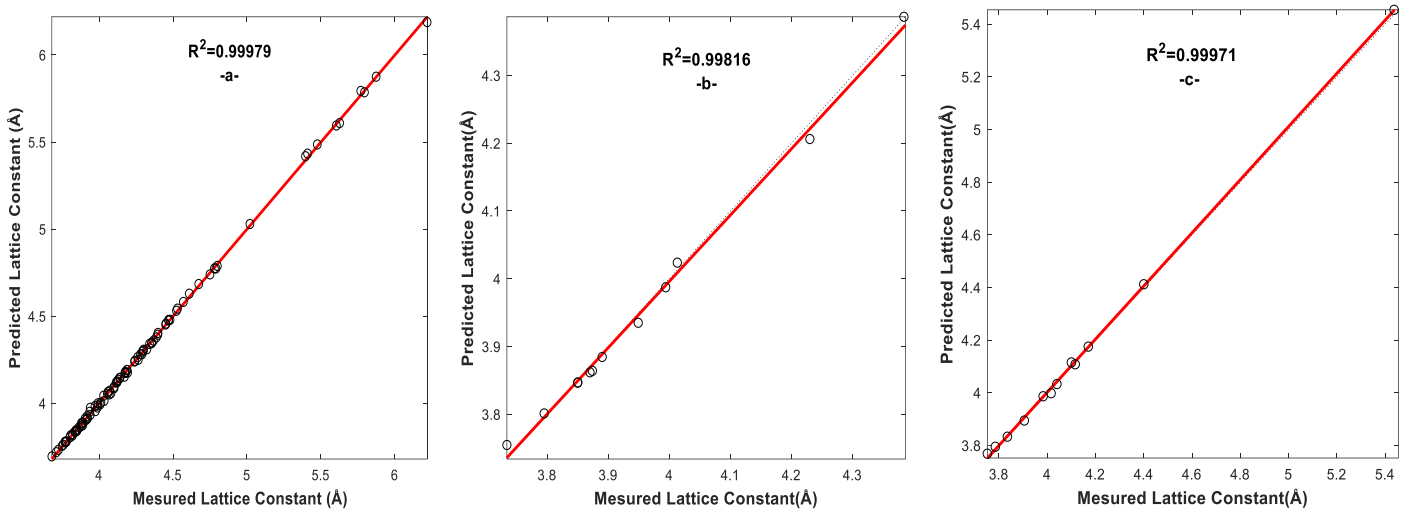

Figure.IV.14 The correlation-cross plot between the experimental lattice parameter and the predicted value for ANN-based PSO model (-a-) training, (-b-) validation, (-c-) testing.

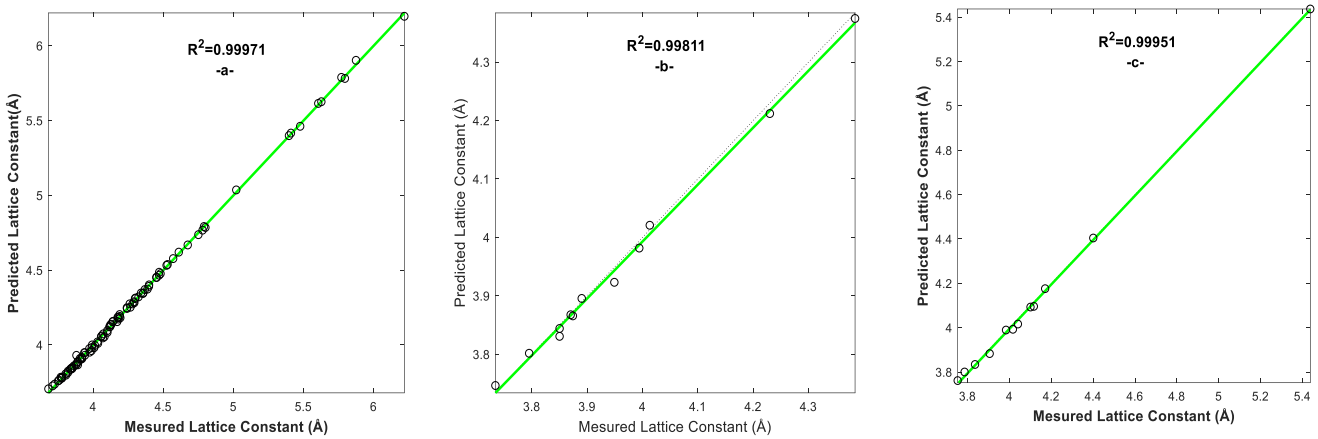

Figure.IV.15 The correlation-cross plot between the experimental lattice parameter and the predicted value for ANN-based IWO model (a) training, (b) validation, (c) testing.

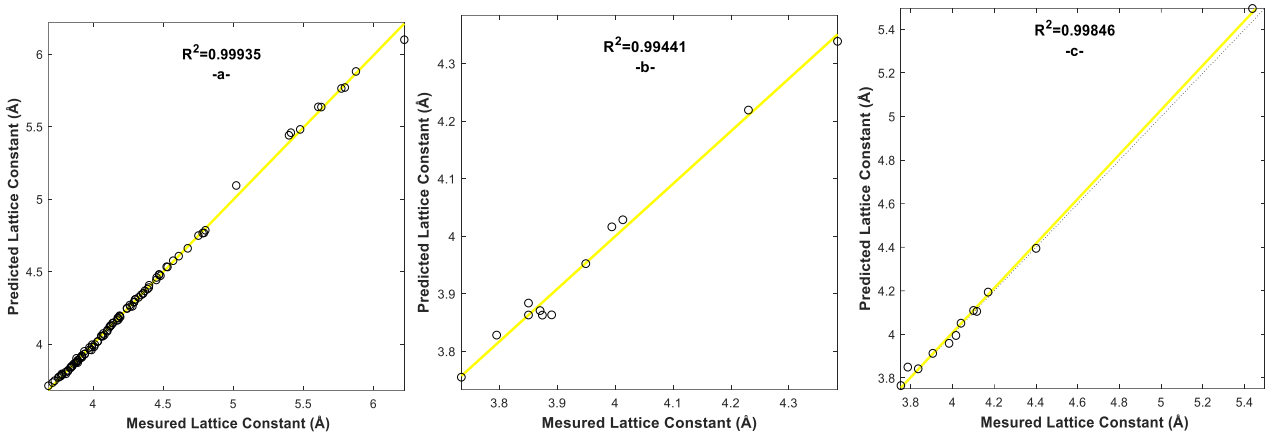

Figure.IV.16 The correlation-cross plot between the experimental lattice parameter and the predicted value for ANN-based ICA model (a) training, (b) validation, (c) testing.

To assess the efficiency of the suggested optimization methods, we generate a graphical comparison between the outcomes of our proposed approach and the results of the experiment. Therefore, Figure IV.17 displays the results of the PSO, IWO, and ICA based algorithms' validation phase.

A superposition of the output between the experimental validation set and the predicted validation set is shown in Figure IV.17-a. Visual inspection reveals that the experimental Lattice Constants findings and the output of the prediction techniques are not properly superposed and are somewhat close to one another. Consequently, a closer look at its zoomed version (FigureIV.17-b) reveals that the experimental output more closely approaches the predicted lattice constant of the ANN based on PSO method than that of the ICA and IWO based algorithms.

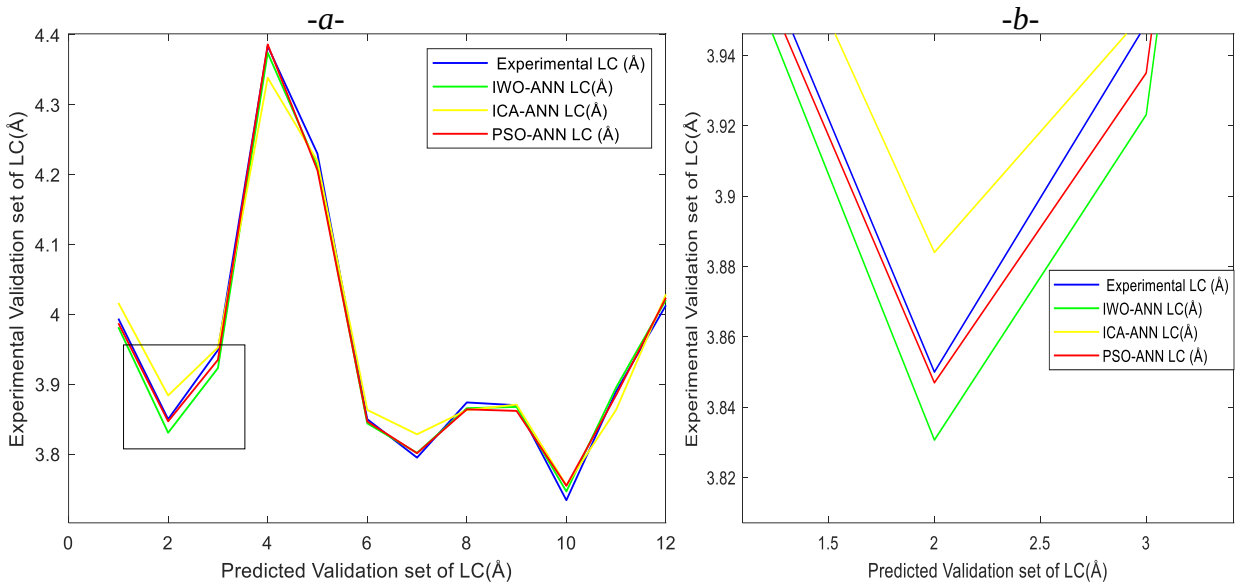

Figure.IV. 17 (a) The experimental validation data set compared to the prediction validation data set for the ANN-based models (PSO, IWO, and ICA); (b) the magnified section.

In the following, we conducted a quantitative comparison of the ANN prediction model's performances utilizing the PSO, IWO, and ICA optimization techniques. Twenty independent trials of the algorithm were conducted. Statistical performance metrics, including the fitness function's worst, mean, best, and standard deviation (SD), were calculated and displayed in Table IV.12. As we observe that this last suggests that during all of the independent experiments, the PSO algorithm was unquestionably better and operated at the greatest level.

Table.IV.12 statistical performance accumulated during 20 separate runs.

Algorithms	*Mean*	*Best*	*Worst*	*SD*
PSO	3.032 e $^{-3}$	0.481 e $^{-3}$	0.124 e $^{-3}$	2.5218e^{-04}
IWO	3.303 e $^{-3}$	0.539 e $^{-3}$	0.121 e $^{-3}$	2.9539e^{-04}
ICA	4.406 e $^{-3}$	0.6729 e $^{-3}$	0.2083 e $^{-3}$	3.2849e^{-04}

Finally, we took into consideration the error bars for the parameter optimization to assess the correctness of the parameter optimization for further statistical analysis and thorough interpretation. This method gives a broad notion of the accuracy of the parameter values by graphically representing the variability of the estimated parameters on graphs to show the estimation uncertainty. Figures IV.18, 19, 20, and 21 show error bars for the optimized parameters of the ANN prediction model employing the three optimization techniques (PSO, IWO, and ICA). It is evident that the ICA method's error bar widths are the narrowest when compared to those of the PSO and IWO techniques, indicating that the ICA algorithm is more stable than the PSO and IWO algorithms.

According to the comparison analysis, the three optimization algorithms (PSO, IWO, and ICA) that were employed to resolve the neural model's parameter optimization problem for the lattice constant prediction are almost equally effective. Therefore, we can say that the algorithms' respective performances in this specific task do not significantly differ.

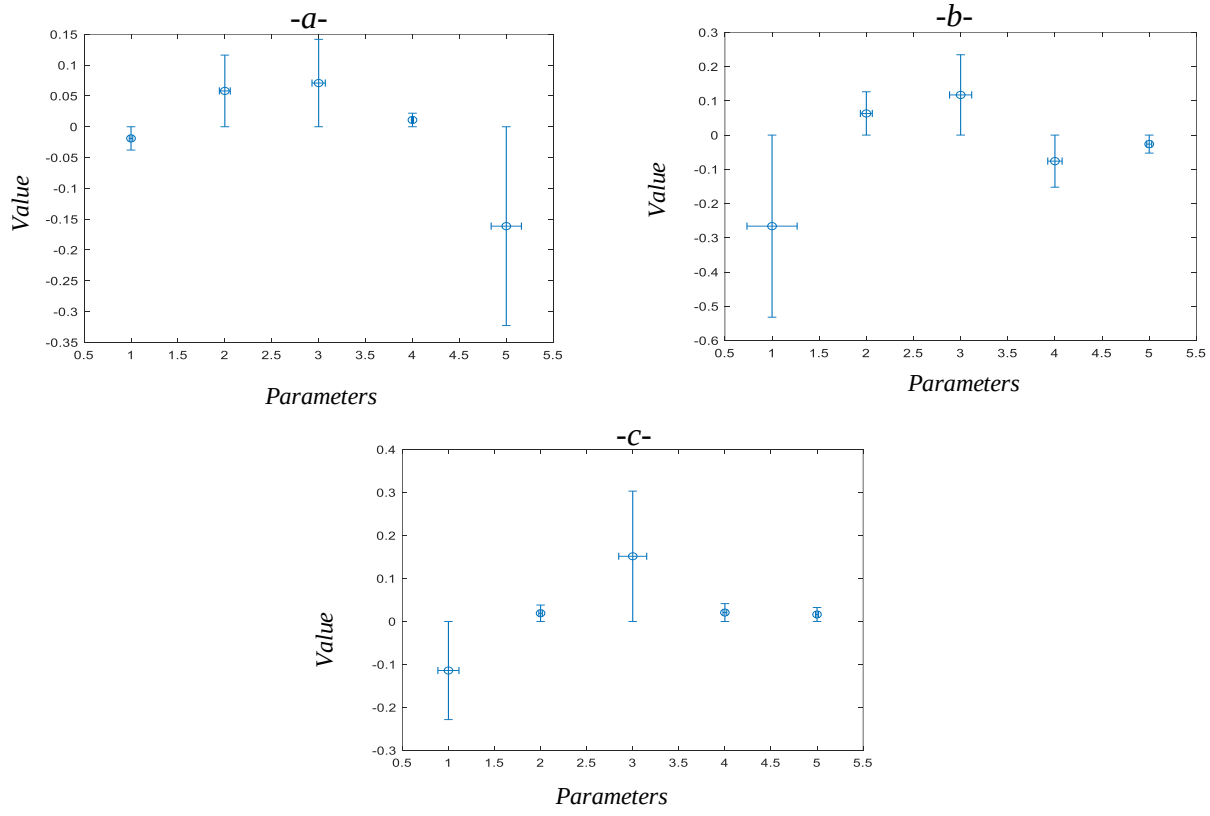

Figure.IV.18 The input layer parameters' error bars (a) PSO, (b) IWO, (c) ICA.

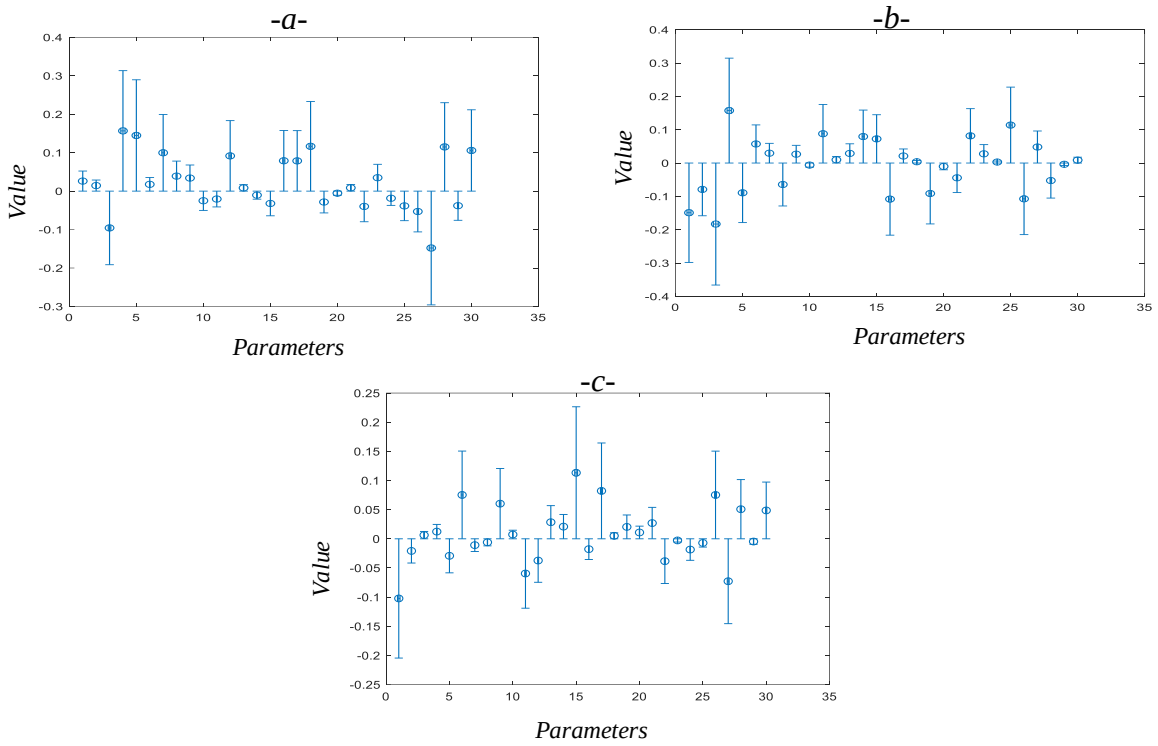

Figure.IV.19 The first hidden layer parameters' error bars (a) PSO, (b) IWO, (c) ICA.

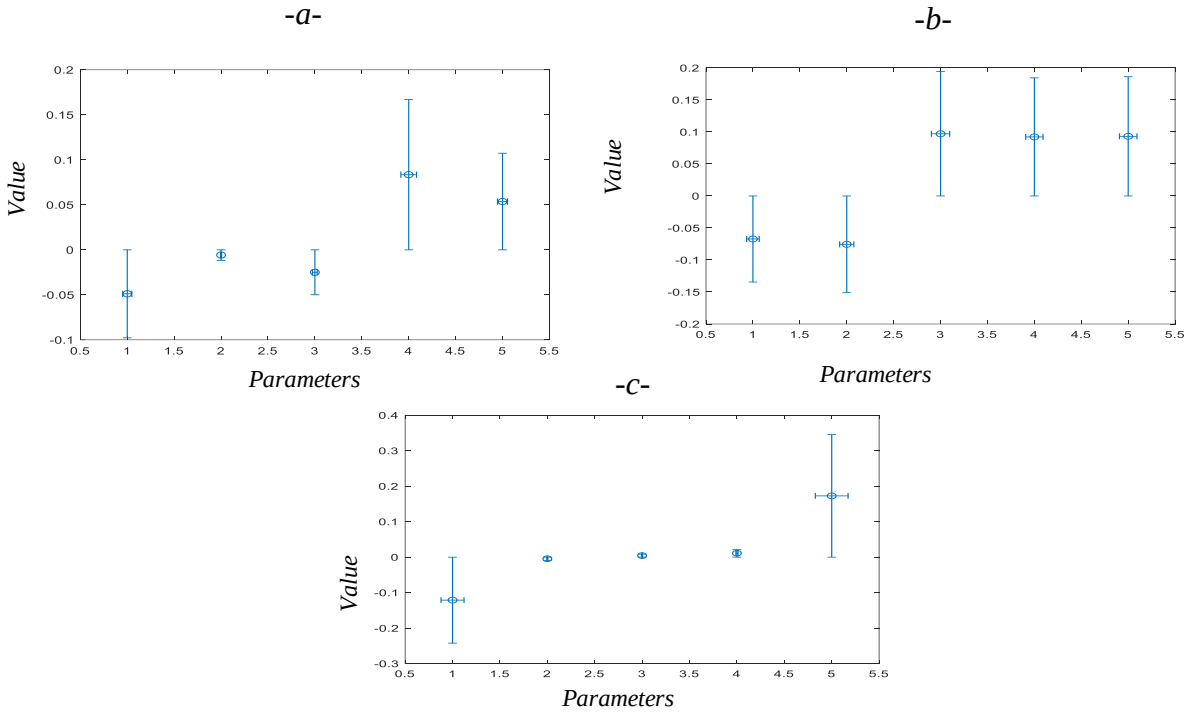

Figure.IV.20 The second hidden layer parameters' error bars (a) PSO, (b) IWO, (c) ICA.

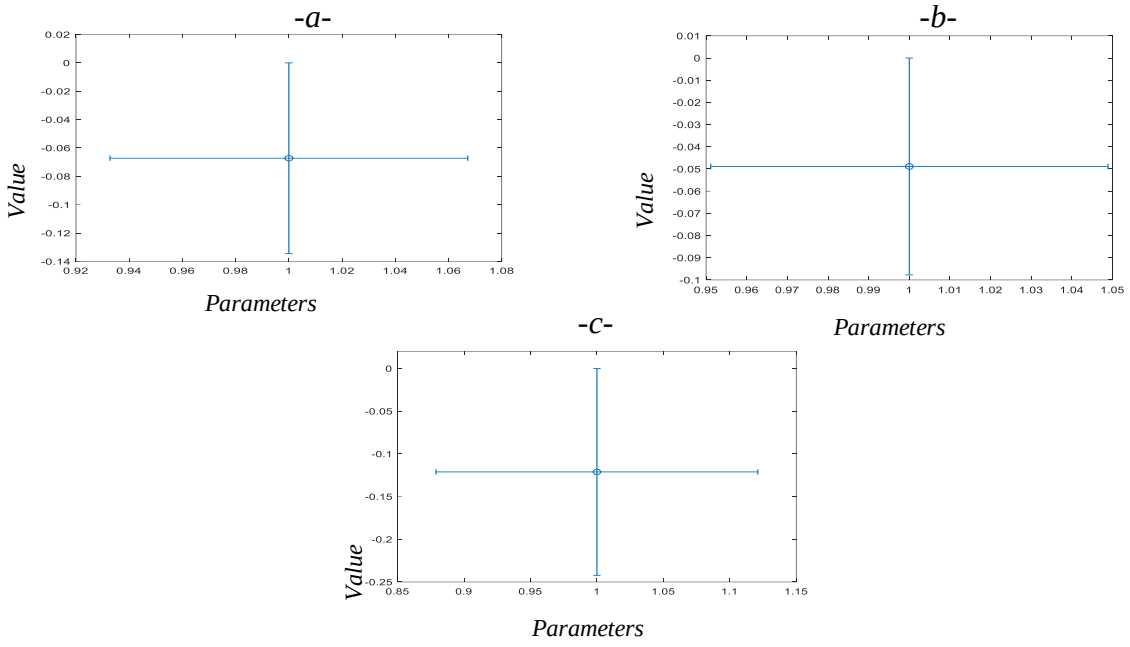

Figure.IV.21 The output layer parameters' error bars (a) PSO, (b) IWO, (c) ICA.

IV.3.2.5.4. Comparative study

IV.3.2.5.4.1.ANN model against Fuzzy model

In order achieve the second goal of this section, we evaluated how well two prediction models “the fuzzy prediction model and the artificial neural network model «predicted lattice constants. To be more precise, the lattice constant was predicted using the particle swarm optimization technique and the fuzzy scheme discussed in IV.3.2.3.4. The results were then compared to those obtained using the ANN-based PSO prediction model. The results presented in Figure IV.22 demonstrate a steady decrease in mean squared error as the number of iterations increased for the Fuzzy-based PSO model contrast to the ANN-based PSO model. This observation validates the faster convergence of the Fuzzy-based PSO model in comparison to the ANN-based PSO model.

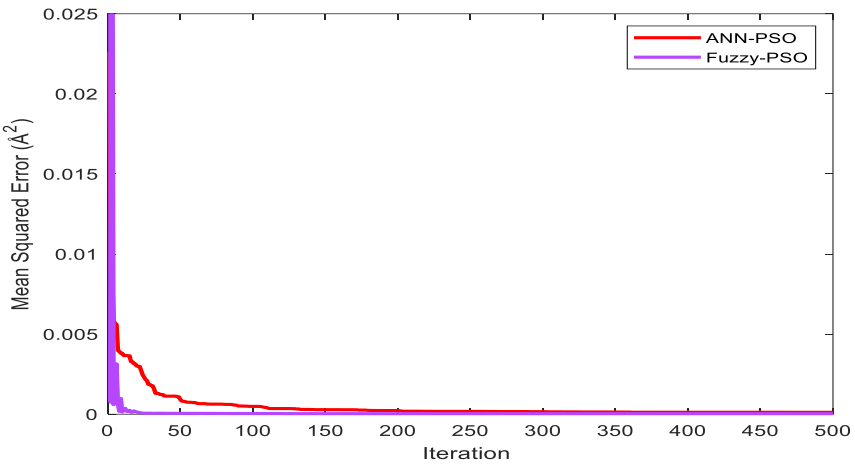

Figure.IV.22 The MSE **Vs** iteration for the Fuzzy and ANN based PSO model.

The performance of the ANN-based PSO model and the fuzzy-based PSO model was quantitatively evaluated using the evaluation criteria .The numerical results for the two models at each stage are shown in Table IV.13. It is verified that the PSO-Fuzzy model outperforms the PSO-ANN model in terms of efficiency based on the data reported in [26-28].

Table.IV.13 Measurement of the PSO-ANN and PSO-Fuzzy models' performances based on several criteria.

Prediction Model	Data sets	*MSE*	*RMSE*	*MAE*	R^2
PSO-ANN	Training	$1.1712e^{-04}$	0.0108	0.0085	0.9997
	Validation	$1.3459e^{-04}$	0.0116	0.0095	0.9981
	Tests	$1.3672e^{-04}$	0.0117	0.0104	0.9971
PSO-Fuzzy	Training	$1.1386e^{-05}$	0.0034	0.0026	0.9979
	Validation	$5.6568e^{-06}$	0.0024	0.0019	0.9989
	Tests	$5.7657e^{-06}$	0.0029	0.0019	0.9909

IV.3.2.5.4.2. Performance comparison of the developed models and the existing models.

The study results have been compared with previously published studies in order to assess the performance of the Fuzzy-based PSO model in comparison to existing benchmark approaches of the lattice constant prediction. Figures IV.23-a-and 23-b show a comparison of our model's MAE and RMSE with those of the Ubic's model [22], Sidey's model [23], and Owolabi's model

[24]. Table IV.14 displays all of the performance measurement parameters for both our models and the benchmark models, along with the current models' percentage improvement. Regarding the mean absolute errors shown in Figure IV.23-a-. Upon visual inspection, it is evident that the Fuzzy-PSO model outperforms the PSO-ANN model, with an improvement percentage of 71.26%, however it performs better than the Ubic, Sidey, and Owolabi models, with performance enhancements of 90.31%, 89.58%, and 85.02%, respectively. Furthermore, the root mean squared error criteria is used to compare our models with the benchmark models described in Figure IV.23-b- With a 90.36% improvement in performance over the Ubic model, the new PSO-Fuzzy outperforms the more modern Sidey and Owolabi models, which have performance percentages of 84.46% and 89.74%, respectively. It is obvious from a comparison of our suggested models that the PSO-Fuzzy outperforms the PSO-ANN, with a 70.90% improvement in performance. These results allow us to validate the PSO-Fuzzy models' superiority over the PSO-ANN and other benchmark models. Lastly, all of the used cubic and pseudo cubic ABC_3 perovskites compounds' predicted lattice constant parameter values are listed in Tables IV.15 and IV.16.

Table.IV.14 Performance comparison of the developed models and the existing models.

Model	MSE	RMSE	MAE
PSO-ANN	$1.2054e^{-04}$	0.0110 b)70.90%	0.0087 b)71.26%
PSO-Fuzzy	$1.0336e^{-05}$	0.0032	0.0025
Sidey model [22]	0.0011	0.0332 a)66.86% b)90.36%	0.0258 a)66.27% b)90.31%
Ubic model [23]	0.0010	0.0312 a)64.74% b)89.74%	0.0240 a)63.75% b)89.58%
T. O. Owolabi [24]	$4.424\ e^{-4}$	0.0206 a) 46.60% b)84.46%	0.0167 a)47.90% b)85.02%

a) Percentage improvement of ANN-PSO-based model over the existing model.
b) Percentage improvement of Fuzzy-PSO model over the existing model.

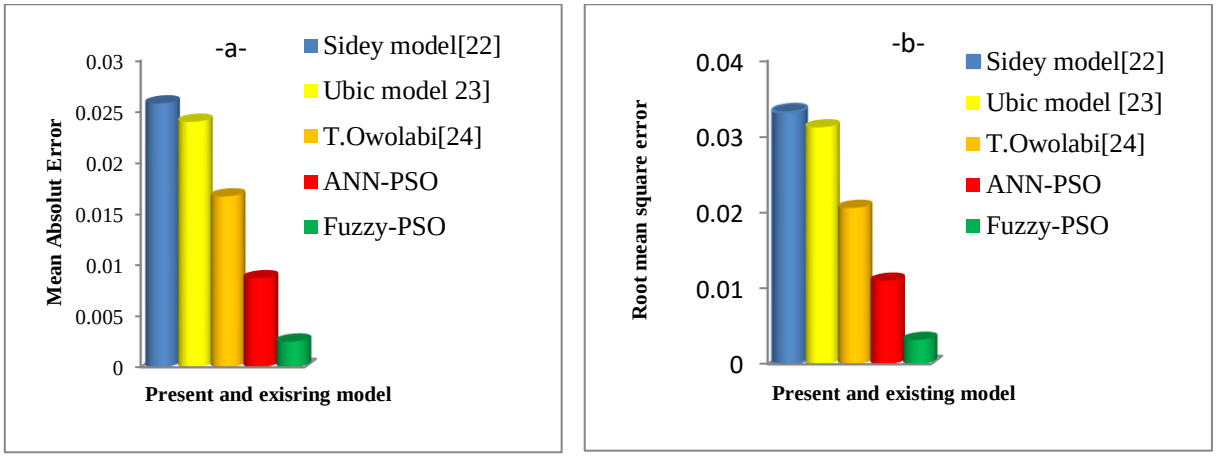

Figure.IV.23 Performance comparison between the existing model and the PSO-Fuzzy model: -a-in terms of mean absolute error,-b- in terms of mean absolute error.

Table.IV.15 Predicted values for lattice parameters (a) of 127 cubic-pseudocubic ABC_3 perovskites using PSO-ANN and PSO-Fuzzy. (a) and (b) exponents Indicates validation and testing data sets.

Compounds	PSO-ANN	PSO-Fuzzy	Compounds	PSO-ANN	PSO-Fuzzy
	a(Å)	a(Å)		a(Å)	a(Å)
$CsIO_3$	4.6847	4.612	$BaMoO_3$ [b]	4.033	4.0547
$RbUO_3$	4.31	4.2805	$BaPrO_3$	4.3453	4.3356
KUO_3	4.2819	4.221	$BaCeO_3$	4.3937	4.3631
$RbPaO_3$	4.3598	4.3097	$BaAmO_3$	4.3489	4.3356
$KPaO_3$	4.341	4.2494	$BaNpO_3$ [a]	4.3865	4.3631
$KTaO_3$	3.9759	4.0544	$BaUO_3$	4.3784	4.3911
$KNiF_3$ [a]	4.0238	4.0089	$BaPaO_3$	4.4524	4.405
$BaNbO_3$	4.0575	4.0954	$BaSnO_3$ [b]	4.1081	4.1094
$BaHfO_3$	4.1527	4.1368	$BaThO_3$	4.4795	4.4583
$BaZrO_3$	4.1763	4.1511	$SrTbO_3$	4.1752	4.1274
$BaIrO_3$	4.0935	4.0286	$SrTiO_3$ [b]	3.8944	3.9262
$EuTiO_3$	3.9126	3.8873	$SrPuO_3$	4.2815	4.2666
$NaWO_3$	3.8422	3.8602	$SrCoO_3$ [a]	3.8467	3.8326
$SnTaO_3$	3.8868	3.9517	$BaTiO_3$	4.0039	4.0031
$SrMnO_3$	3.8147	3.8326	$CaTiO_3$	3.8357	3.8375
$SrNbO_3$ [b]	3.9981	4.0176	$CeAlO_3$	3.7755	3.7605
$SrVO_3$	3.8747	3.9262	$EuAlO_3$	3.7331	3.7291
$SrMoO_3$	3.9552	3.9777	$EuCrO_3$	3.8043	3.8234
$SrSnO_3$	4.014	4.031	$EuFeO_3$	3.838	3.8549
$SrAmO_3$ [a]	4.2061	4.2526	$GdAlO_3$	3.7193	3.724
$SrHfO_3$	4.0509	4.0582	$GdFeO_3$	3.8236	3.8498

Table.IV.16 Continued

Compounds	PSO-ANN	PSO-Fuzzy	Compounds	PSO-ANN	PSO-Fuzzy
$BaPbO_3$	4.2501	4.2361	$GdCrO_3$ [a]	3.8013	3.8183
$KNbO_3$	3.9935	4.0544	$RbCdF_3$	4.4039	4.4384
$LaAlO_3$	3.7803	3.7704	$RbCaF_3$	4.4577	4.5064
$LaCrO_3$	3.8663	3.8651	$RbHgF_3$	4.4746	4.5323
$LaFeO_3$	3.9141	3.8971	$KCdF_3$	4.2963	4.3752
$LaRhO_3$	3.9338	3.9296	$KMgF_3$	3.9876	4.0494
$LaTiO_3$	3.9117	3.9296	$KZnF_3$	4.055	4.077
$LaVO_3$	3.9035	3.8907	$KCoF_3$	4.0717	4.077
$NaTaO_3$	3.8732	3.8854	KVF_3 [b]	4.1151	4.1476
$NdFeO_3$ [a]	3.8619	3.871	$KFeF_3$	4.1206	4.1335
$NdCoO_3$	3.7814	3.7561	$KMnF_3$	4.1911	4.2051
$NdAlO_3$ [b]	3.7685	3.7451	$AgMgF_3$	3.9256	3.944
$NdCrO_3$ [b]	3.8331	3.8394	$AgCoF_3$ [b]	3.9866	3.9709
$PrAlO_3$	3.7617	3.7501	$AgNiF_3$	3.9511	3.9042
$PrCrO_3$	3.849	3.8444	$AgZnF_3$	3.9844	3.9709
$PrFeO_3$	3..8852	3.8764	$AgMnF_3$	4.0434	4.0953
$PrGaO_3$	3.861	3.8444	$NaVF_3$	3.9745	3.973
$PrMnO_3$	3.8149	3.8764	$RbPdF_3$	4.302	4.3098
$PrVO_3$	3.889	3.87	$RbVF_3$ [b]	4.1752	4.2073
$LaGaO_3$ [a]	3.8641	3.8651	$NH4MgF_3$	4.065	4.09
$SmCoO_3$	3.7557	3.7456	$TlPdF_3$	4.3073	4.301
$SmVO_3$ [a]	3.8846	3.8541	$LiBaF_3$	4	3.9828
$SmAlO_3$ [a]	3.7547	3.7346	$RbYbF_3$	4.5436	4.5323
$SmFeO_3$	3.8456	3.8605	$CsEuF_3$	4.7748	4.7672
$SrZrO_3$	4.0866	4.0719	$CsPbF_3$	4.7889	4.7861
$YAlO_3$	3.6949	3.7028	$CsYbF_3$	4.6295	4.5967
$YFeO3$ [b]	3.7949	3.8283	$RbPbF_3$	4.7742	4.7224
$YCrO_3$	3.7786	3.7968	$CsCaCl_3$	5.4186	5.4075
$CsCdF_3$	4.4786	4.502	$CsPbCl_3$	5.5949	5.5852
$CsCaF_3$	4.5314	4.5708	$CsHgCl_3$	5.4333	5.4326
$CsHgF_3$	4.5819	4.5967	$TlMnCl_3$	5.0296	5.0812
$CsSrF_3$	4.7401	4.7771	$TlCdF3$ [b]	4.4122	4.4295
TlCoF3	4.1438	4.1276	$CsEuCl_3$	5.609	5.5736
$BaFeO_3$ [a]	3.9874	3.9781	$CsTmCl_3$	5.4861	5.4202
$SrFeO_3$ [a]	3.847	3.9016	$CsYbCl_3$ [b]	5.4545	5.4326
$TlFeF_3$	4.1908	4.185	$CsHgBr_3$	5.7935	5.7716
TlMnF3	4.2645	4.2577	$CsPbBr_3$	5.8751	5.8335
$NH4ZnF_3$	4.119	4.1181	$CsSnBr_3$	5.7856	5.8214
$NH4CoF_3$	4.1347	4.1181	$CsSnI_3$	6.1884	6.2579
$NH4FeF_3$	4.1796	4.1748	$SrTcO_3$ [a]	3.935	3.9777
$NH4MnF_3$	4.2437	4.2472	$RbCoF_3$	4.1468	4.1357

IV.3.4.The double perovskites oxide A_2BCO_6 lattice constant prediction

IV.3.4.1.Presepective

Double perovskite oxides have garnered considerable attention recently due to their unique and versatile material properties, as discussed in Section I.5. Within the cubic structure, the lattice constant stand out as one of the parameters that significantly influences the development of materials tailored for specific technological applications and distinctly characterizes the material's crystal structure [26]. This study employs the extreme learning machine (ELM) to establish correlations between the lattice constant of A_2BCO_6 cubic perovskite compounds and various parameters, with the goal of discovering an efficient technique for double perovskite lattice constant prediction.

IV.3.4.2.Data description

The double perovskites, which have a general formula, contain one type of A-cation element, specifically alkaline earth metals B and C. These elements include [Ag, Ba, I, U, Bi, Ta, Ca, Mo, Ca, Os, Re, Te, W, Cd, Ce, Pa, Co, Cr, Dy, Nb, Er, Ru, Fe, Gd, Ho, Sb, In, La, Li, Lu, Mg, Na, Ni, Pr, Rh, Sc, Sm, Tb, Tl, Tm, Yb, Y, Zn, Al, Ga, Mn]. The input variables are ionic radii (rA(Å), rB(Å), and rC(Å)), electrostatics (xB and xC), and oxidation states (zB). This experimental data set has been reported in [37].

IV.3.4.3.Proposed method

IV.3.4.3.1.Computational development of the ELM-based model

The Extreme Learning Machine (ELM) was suggested by Huang et al. [28]. This is a new version of the single-hidden layer feed-forward neural network. The approach overcomes the limitations of the single-hidden layer feed-forward neural network by randomly distributing the biases and weights between the input layer and the hidden layer [29].

The modeling and simulation in this study utilized data consisting of six descriptors for 147 cubic complex perovskite materials, along with their corresponding lattice constants. This dataset was randomly split into training and testing sets with an 8:2 ratio, respectively, ensuring an even distribution of data points for accurate calculations. The computational implementation of the Extreme Learning Machine (ELM) generated random input weights and hidden layer biases. Activation functions for the hidden layer neurons, including Sin, Sig, hardlim, tribas, and radbas,

were chosen from a selection. The hidden node range was set from 1 to 100. Subsequently, weights were computed to link the hidden and output layers. With these generated biases and weights, the crystal lattice parameters of the test dataset were predicted. These predicted parameters were then compared with the measured values using RMSE (Eq. IV.11).Figure IV.24 illustrates the schematic diagram depicting the architecture of the ELM-based models created in this study.

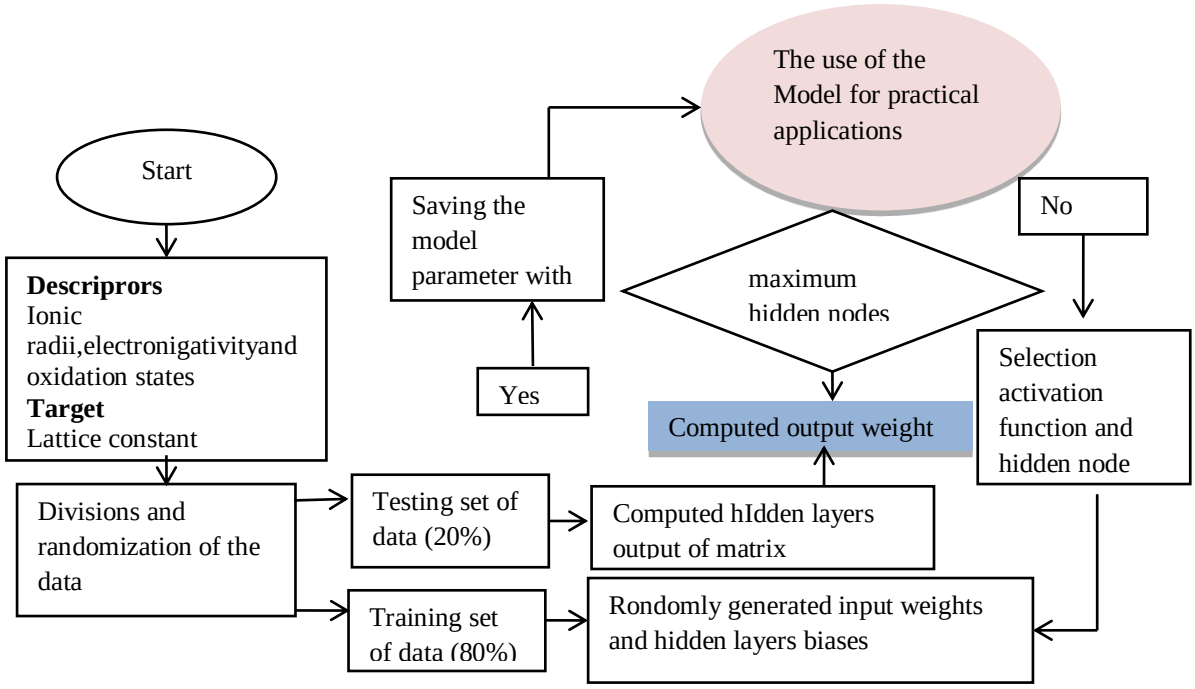

Figure.IV.24 Computational diagram of the developed extreme learning machine based model.

IV.3.4.4.Results and Discussion.

IV.3.4.4.1.Performance Comparison Between the Present Model and Existing Model

Figure IV.25 (a) illustrates the comparison of the developed ELM model with Sandra et al.'s [30] models, including ANN and MLR, based on mean absolute errors (MAE). The ELM model exhibits superior performance compared to MLR and ANN models, with percentage improvements of 10% and 14% respectively. When evaluating using RMSE, as depicted in Figure IV.25 (b), our model demonstrates a performance improvement of 24% and 22% respectively over the ANN and MLR methods. Figure IV.25 (c) presents a comparison of the correlation coefficients between the present work and the Sandra models. The ELM model exhibits a correlation coefficient of 98.74%, surpassing that of MLR and ANN models, which

stand at 98.43% and 98.53% respectively.Table .IV.17 Provides the actual values of all performance measuring indices for the compared models along with the percentage improvement of the present model.

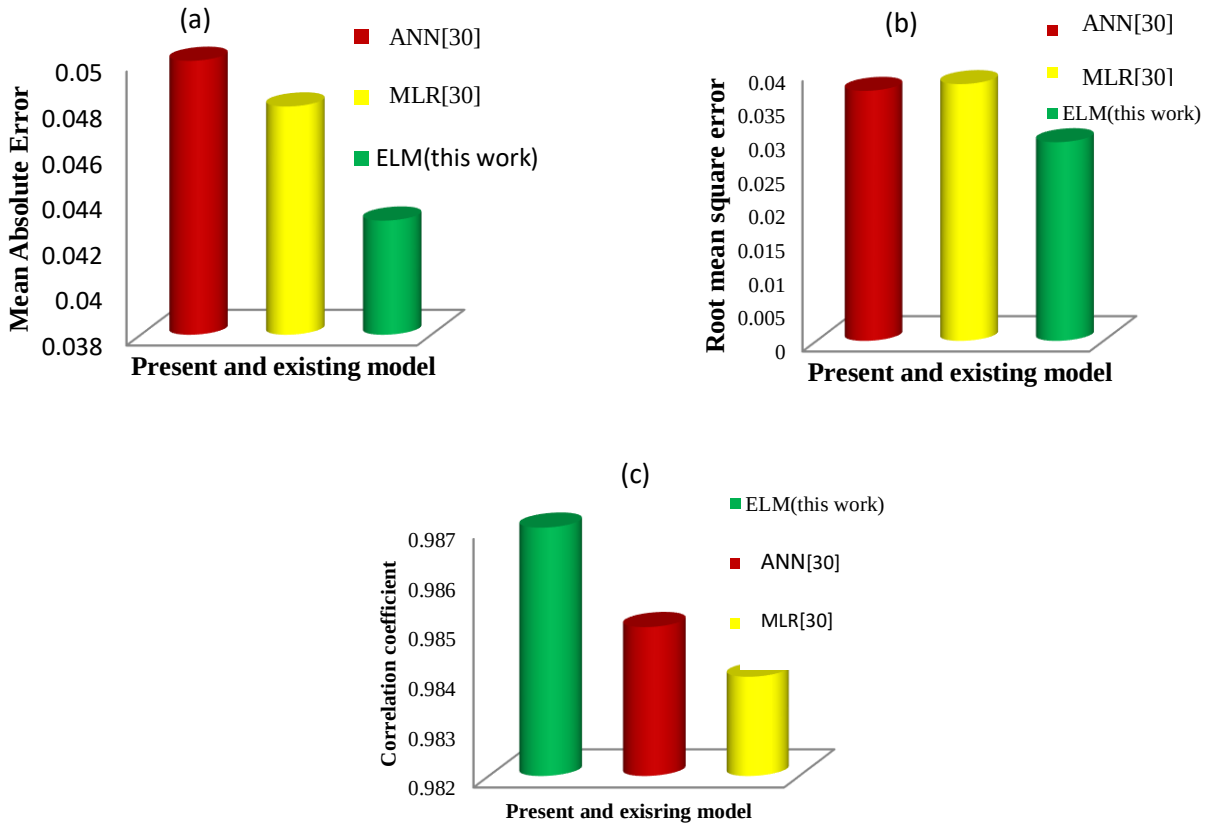

Figure.IV.25 Performance comparison between the existing model and the present model on the basis of :(a)Root mean square error,(b) Root mean square error and (c)correlation coefficient.

Table.IV.17 Values of performance measuring parameters for the proposed and existing model.

Models		MAE(Å)	RMSE(Å)	CC%
Sandra et al model [30]	MLR	0.048(10%)	0.038(24%)	98.43
	ANN	0.050(14%)	0.037(22%)	98.52
ELM(present work)		0.043	0.029	98.74

IV.3.4.4.2.Performance Comparison between the estimates and measured lattice Parameters.

The correlation crossplot between the measured crystal lattice parameters and the estimated values for the current model is presented in Fig.IV.26. The data point correspond to our developed model are well aligned.

Furthermore, Tables IV. 18 and19 list the actual and the prediction values of lattice constant parameters for all the investigated $A_2^{+2}BCO_6$ cubic perovskites compounds .The percentage error for each of the compounds is presented in Tables IV. 18 and 19. The maximum error and minimum error associated with our model is 2.90% and 0%, which corresponds respectively to Ba_2LiOsO_6 andBa_2BaUO_6, Ca_2MgWO_6 cubic $A_2^{+2}BCO_6$ compounds. On the other hand, the *MAPE* of the developed model is 0.356%, which is a compelling argument for its performance superiority to the existing models [30].

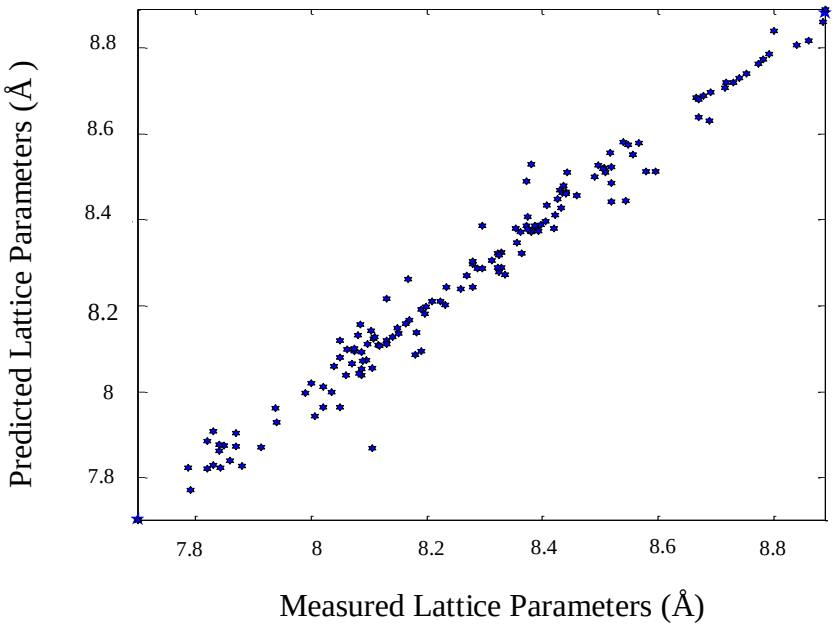

Figure.IV.26 The Correlation cross-plot between the mesured lattice constant and the predicted value for the present work.

Table.IV. 18 The experimental [27] and predicted value for the lattice parameter of A_2BCO_6 perovskites.

Compounds	y_{exp}(Å)	y_{pred}(Å)	error%	Compounds	y_{exp}(Å)	y_{pred}(Å)	error%
Ba_2AgIO_6	8.460	8,455	0,052	Ba_2ErPaO_6	8.716	8,707	0,099
Ba_2LiOsO_6	8.105	7,869	2,909	Ba_2ErRuO_6	8.323	8,322	0,001
Ba_2NaIO_6	8.330	8,323	0,074	Ba_2ErTaO_6	8.423	8,411	0,137
Ba_2NaOsO_6	8.287	8,286	0,011	Ba_2EuNbO_6	8.507	8,520	0,157
Ca_2LiOsO_6	7.830	7,828	0,017	Ba_2EuPaO_6	8.783	8,772	0,124
Ca_2LiReO_6	7.830	7,907	0,989	Ba_2FeMoO_6	8.075	8,094	0,237
Sr_2LiReO_6	7.870	7,872	0,030	Ba_2FeReO_6	8.050	8,078	0,357
Sr_2NaOsO_6	8.130	8,119	0,123	Ba_2GdPaO_6	8.774	8,762	0,129
Ba_2BiTaO_6	8.568	8,578	0,123	Ba_2GdReO_6	8.431	8,467	0,438
Ba_2CePaO_6	8.800	8,838	0,440	Ba_2HoNbO_6	8.434	8,463	0,348
Ba_2DyNbO_6	8.437	8,478	0,495	Ba_2HoPaO_6	8.730	8,718	0,133
Ba_2DyPaO_6	8.740	8,729	0,121	Ba_2InNbO_6	8.279	8,296	0,209
Ba_2ErNbO_6	8.427	8,448	0,249	Ba_2InOsO_6	8.224	8,210	0,161
Ba_2InReO_6	8.258	8,238	0,239	Ca_2CaWO_6	8.000	8,018	0,234
Ba_2InSbO_6	8.269	8,270	0,021	Pb_2FeWO_6	8.050	7,964	1,067
Ba_2InUO_6	8.520	8,486	0,395	Pb_2MgTeO_6	7.990	7,997	0,089
Ba_2LaPaO_6	8.885	8,859	0,287	Sr_2CaOsO_6	8.210	8,209	0,005
Ba_2LuNbO_6	8.364	8,321	0,510	Sr_2CoUO_6	8.190	8,093	1,174
Ba_2LuPaO_6	8.666	8,684	0,213	Sr_2FeOsO_6	7.850	7,873	0,305
Ba_2MnReO_6	8.180	8,085	1,155	Sr_2FeUO_6	8.110	8,126	0,205
Ba_2NdNbO_6	8.540	8,581	0,481	Sr_2MgUO_6	8.190	8,191	0,018
Ba_2NdReO_6	8.510	8,509	0,005	Sr_2MnUO_6	8.280	8,242	0,455
Ba_2NdTaO_6	8.556	8,552	0,039	Ba_2LiReO_6	8.118	8,106	0,147
Ba_2RhNbO_6	8.170	8,166	0,047	Ba_2NaReO_6	8.296	8,287	0,107
Ba_2ScNbO_6	8.234	8,243	0,111	Sr_2LiOsO_6	7.860	7,839	0,255
Ba_2ScPaO_6	8.549	8,574	0,298	Sr_2NaReO_6	8.130	8,215	1,051
$Ba2ScReO_6$	8.163	8,157	0,067	Ba_2CoReO_6	8.086	8,091	0,067
Ba_2ScTaO_6	8.231	8,201	0,355	Ba_2DyTaO_6	8.545	8,443	1,190
Ba_2ScUO_6	8.490	8,500	0,122	Ba_2ErReO_6	8.354	8,379	0,308
Ba_2SmPaO_6	8.792	8,784	0,081	Ba_2ErUO_6	8.670	8,679	0,107
Ba_2SmTaO_6	8.519	8,522	0,035	Ba_2EuTaO_6	8.506	8,518	0,146
Ba_2TlSbO_6	8.381	8,370	0,122	Ba_2GdNbO_6	8.496	8,527	0,369
Ba_2TlTaO_6	8.420	8,378	0,487	Ba_2GdSbO_6	8.440	8,465	0,304
Ba_2TmPaO_6	8.692	8,696	0,050	$Ba2HoTaO_6$	8.442	8,509	0,796
Ba_2TmTaO_6	8.406	8,395	0,125	Ba_2InPaO_6	8.596	8,513	0,962
Ba_2YPaO_6	8.718	8,718	0,009	Ba_2InTaO_6	8.280	8,303	0,278
Ba_2YReO_6	8.372	8,386	0,177	Ba_2LaReO_6	8.580	8,512	0,785
Ba_2YUO_6	8.690	8,630	0,687	Ba_2LuTaO_6	8.372	8,489	1,401
Ba_2YbNbO_6	8.374	8,406	0,388	Ba_2NdPaO_6	8.840	8,806	0,382
Ba_2YbTaO_6	8.390	8,382	0,083	Ba_2PrPaO_6	8.862	8,816	0,509
Pb_2ScTaO_6	8.140	8,126	0,160	Ba_2ScOsO_6	8.152	8,136	0,190
Sr_2AlTaO_6	7.791	7,770	0,260	Ba_2ScSbO_6	8.197	8,181	0,183
Sr_2CoSbO_6	7.880	7,827	0,667	Ba_2SmNbO_6	8.518	8,554	0,432
Sr_2CrOsO_6	7.840	7,861	0,280	Ba_2TbPaO_6	8.753	8,739	0,148
Sr_2CrWO_6	7.820	7,884	0,821	Ba_2TmNbO_6	8.408	8,432	0,293
Sr_2GaOsO_6	7.820	7,821	0,012	Ba_2YNbO_6	8.441	8,461	0,238
Sr_2GaReO_6	7.843	7,823	0,253	Ba_2YTaO_6	8.433	8,427	0,061
Sr_2InReO_6	8.071	8,064	0,081	Ba_2YbPaO_6	8.678	8,688	0,125
Sr_2InUO_6	8.330	8,288	0,497	Sr_2AlNbO_6	7.786	7,823	0,482

Table.IV. 19 Continued

Compounds	y_{exp}(Å)	y_{pred}(Å)	error%	Compounds	y_{exp}(Å)	y_{pred}(Å)	error%
Sr_2ScBiO_6	8.182	8,137	0,541	Sr_2CrMoO_6	7.840	7,875	0,458
Sr_2ScOsO_6	8.020	8,011	0,103	Sr_2FeBiO_6	8.063	8,099	0,447
Sr_2RhTaO_6	7.939	7,962	0,292	Sr_2InOsO_6	8.060	8,039	0,257
Sr_2CrNbO_6	7.870	7,903	0,429	Sr_2RhNbO_6	7.914	7,871	0,537
Ba_2CaMoO_6	8.380	8,374	0,070	Sr_2ScReO_6	8.020	7,963	0,701
Ba_2CaOsO_6	8.362	8,371	0,113	Ba_2BaUO_6	8.890	8,890	0
Ba_2CaTeO_6	8.393	8,373	0,237	Ba_2CaReO_6	8.356	8,347	0,099
Ba_2CaUO_6	8.670	8,638	0,367	Ba_2CaWO_6	8.388	8,385	0,030
Ba_2CdMoO_6	8.324	8,318	0,070	Ba_2CdReO_6	8.322	8,289	0,391
Ba_2CdOsO_6	8.325	8,277	0,565	Ba_2CoUO_6	8.374	8,378	0,048
Ba_2CoMoO_6	8.086	8,052	0,411	Ba_2FeReO_6	8.050	8,119	0,864
Ba_2CoReO_6	8.086	8,038	0,582	Ba_2MgOsO_6	8.080	8,131	0,631
Ba_2CoWO_6	8.108	8,122	0,178	Ba_2MgUO_6	8.381	8,529	1,776
Ba_2CrUO_6	8.297	8,385	1,061	Ba_2MnWO_6	8.199	8,198	0,011
Ba_2FeUO_6	8.312	8,306	0,068	Ba_2NiReO_6	8.040	8,059	0,244
Ba_2MgMoO_6	8.084	8,156	0,897	Ba_2ZnMoO_6	8.103	8,142	0,484
Ba_2MgReO_6	8.082	8,041	0,499	Ba_2ZnUO_6	8.397	8,388	0,099
Ba_2MgTeO_6	8.130	8,110	0,234	Ca_2MgWO_6	7.700	7,700	0
Ba_2MgWO_6	8.098	8,110	0,159	Pb_2MgWO_6	8.006	7,944	0,773
Ba_2MnMoO_6	8.168	8,261	1,147	Sr_2CrUO_6	8.090	8,071	0,225
Ba_2MnUO_6	8.520	8,442	0,907	Sr_2MgTeO_6	7.940	7,929	0,130
Ba_2NiMoO_6	8.035	7,999	0,436	Sr_2NiUO_6	8.150	8,148	0,021
Ba_2NiUO_6	8.336	8,272	0,761	MAPE			0,356
Ba_2NiWO_6	8.075	8,101	0,324	Maximum Error			2,909
Ba_2ZnReO_6	8.106	8,055	0,623	Minimum Error			0

IV.4. Calculation of the Lattice parameter and the Band gap energy for the double perovskite compound Ba_2BiTaS_6 using the DFT approach.

IV.4.1. Perspective on the material selected for the study

The chalcogenide perovskites have lately been suggested as possible absorbers for solar cells, because the Coulomb interaction in chalcogenides is predicted to be four times greater than that in halides for purely ionic systems. An ABX_3 family of perovskites (where A = Ca, Sr, Ba; B = Ti, Zr, Hf; and X = S, Se) has been theoretically explored by Sun et al [31] for potential use as solar cell absorbers. Among these, experimentally synthesized BaZrS3 was suggested as a viable contender, with a direct bandgap of 1.73−1.85 eV [32-33]. Another eighteen chalcogenide perovskites, designated ABX_3 (A = Ca, Sr, Ba; B = Ge, Sn, Te; X = S, Se), were studied by Ju et al [34]. They suggested that the bandgap might be tuned for optimal absorption of sunlight by properly combining $SrSnS_3$ and $SrSnSe_3$. In this study, one may create lone-pair s orbitals on a B site and produce a family of chalcogenide double perovskites $A_2B(III)C(V)X_6$ by chemically changing B-site cations based on $BaZrS_3$, i.e.,Zr (IV) → B(III) + C(V), where B(III) = Bi, C(V) = Ta and X=S.

In this section, we have studied and calculated the structural and electronic properties of the chalcogenide double perovskite Ba_2BiTaS_6 compound. Our interest in this material emerged because of its expected applications in various fields such as opto-electronics devices, particularly in solar cells.

IV.4.2. Calculation details

In this study, the calculations are performed using the Wien2k code [35], to investigate the structural and electronic properties of double perovskites Ba_2BiTaS_6 at thermodynamic equilibrium. This code is widely used in calculations of the physicochemical properties of materials because in most cases its results are very close to those of experiments. Based on the FP-LAPW method implemented in the Wien2k code [36], we have adopted as exchange and correlation functional the generalized gradient approximation (GGA: Generalized Gradient Approximation) in the framework of (PBE: Perdew-BurKe -Ernzerhof). To obtain more accurate values for the band gap, the Tran-Blaha modified Becke-Johnson approximation (TB-mBJ) [37-39] was employed. This approach is based on the modified Becke-Johnson potential and is optimized for total energy and the corresponding potential.

In the FP-LAPW method, the space is divided into non-overlapping Muffin-tin (MT) spheres and an interstitial region between the spheres. The Kohn-Sham wave functions are expanded in terms of spherical harmonics inside the MT spheres with a maximum value of *lmax*= 10 and in Fourier series in the interstitial region. The plane waves are limited to *RMT.Kmax* = 7 (where RMT is the smallest radius of the MT sphere and *Kmax* is the maximum wave vector K).For the number of k points used to sample the first Brillouin zone in reciprocal space, a grid of K points equal to 1000, corresponding to a mesh of (10×10×10), is sufficient to converge the total energy. The Muffin-Tin radii and the valence states adopted in our work are listed in Table IV.20.

Table.IV. 20. The atoms' valence states and RMT radii in the material under consideration.

Atoms	Electronic configuration	RMT(a.u)
Ba	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^6\ 4d^{10}\ 5s^2\ 5p^6\ 6s^2$	2
Bi	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^6\ 4d^{10}\ 5s^2\ 5p^6\ 4f^{14}\ 5d^{10}\ 6s^2\ 6p^3$	2
Ta	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^6\ 4d^{10}\ 5s^2\ 5p^6\ 4f^{14}\ 5d^3\ 6s^2$	2
S	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^4$	1.7

IV.4.3. Result and Discussion

IV.4.3.1. The lattice parameter

The unit cell of our compounds, which contains ten atoms, crystallizes in the cubic structure space group (N° = 225) without any deformation, described by $a = b = c$ and $\alpha = \beta = \gamma = 90°$. The positions of the Ba, Bi, Ta, and S atoms are: (0.25, 0.25, 0.25), (0, 0, 0), (0.5, 0, 0) and (0.25, 0, 0) respectively. The lattice parameter corresponding to our compound is 8.342 Å. Figure IV.27 shows the cubic structure of Ba_2BiTaS_6.

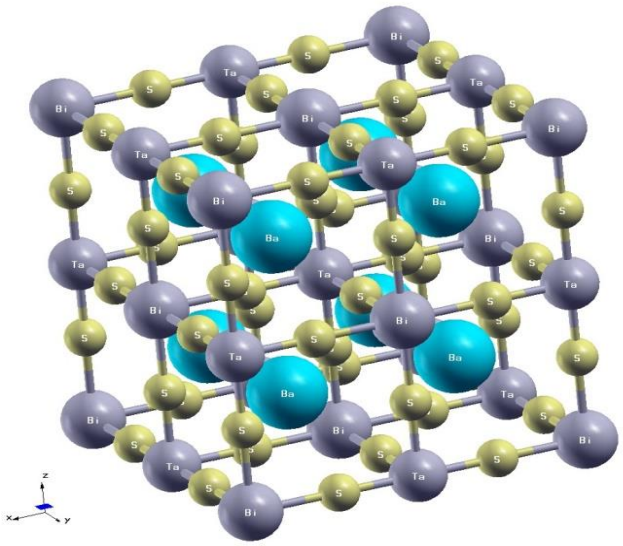

Figure.IV.27 The structure representation of the Ba_2BiTaS_6 material.

IV.4.3.2. The band gap energy

In this section, we are interested in calculating the energy band structures for our double perovskite compounds Ba_2BiTaS_6, using the optimized lattice parameters. The calculations were carried out along various high-symmetry points and directions in the first Brillouin zone to determine the values of the energy band gaps of this material. The calculations were performed using the GGA-PBE Sol and TB-mBJ approximations. Where, the TB-mBJ approximation is used to find the band gap in better agreement with experimental data. The obtained results for the electronic band structure of Ba_2BiTaS_6 are depicted in figures (IV.28-a-) and (IV.28-b-) respectively for both the GGA-PBE Sol and TB-mBJ approximations. By visual inspection, it is evident that the considerate compound Ba_2BiTaS_6 are indirect bandgap semiconductors. Specifically, along the Γ - X direction, the maximum of the valence band is located at the point Γ, while the minimum of the conduction band is at the point X. In Table IV.21, we compile the various gap values for our compound obtained using

both GGA and TB-mBJ approximations, along with those published by Sun et al. in 2019 [40].

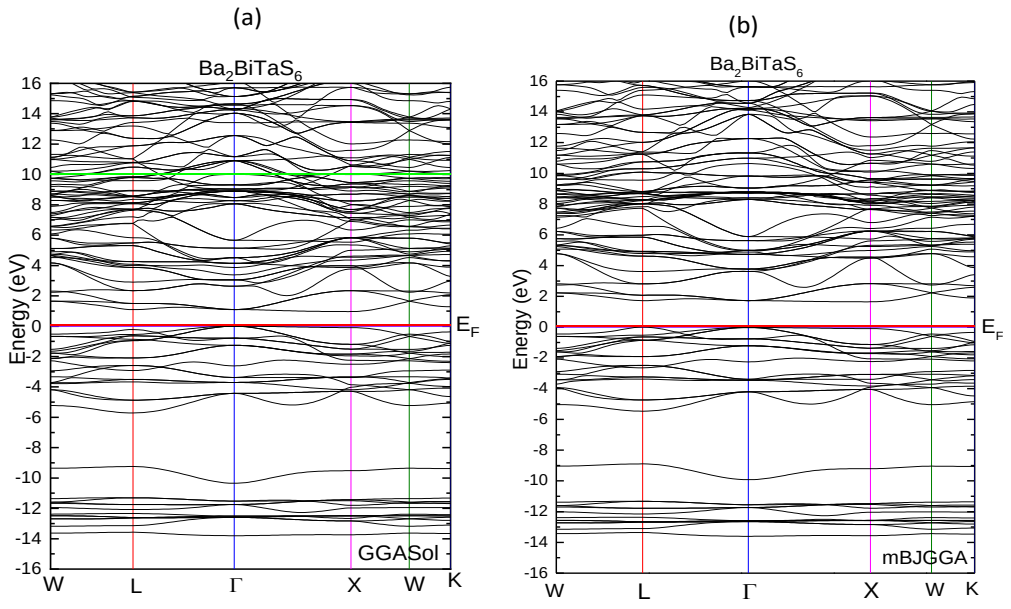

Figure.IV.28 The band structure of Ba_2BiTaS_6 :(a)-using the GGA SOL approximation and (b)using mBJ-GGA.

Table.IV.21 The calculated values of the gaps compared To the previous experimental result [40].

Compound	Type	Direction	Eg(ev)		
			GGA	mBJ-GGA	others [40]
Ba_2BiTaS_6	Indirect	Γ - X	0.97	1.67	1.56

IV.5. Conclusion

This chapter presents new results concerning the band gap energy and the lattice constant properties of chalcopyrite, single perovskites, and double perovskites using different approaches based on computational intelligence and machine learning methods combining with metaheuristics algorithms and DFT methods.

Firstly, we present different methods based on machine learning to enhance the prediction of materials' properties with higher accuracy and accelerate their discovery, which is less computationally costly. In the first part, we introduce a novel approach combining artificial neural networks and particle swarm optimization algorithms to predict the band gap energy

for 266 ABC_2 compounds with high precision. Our method addresses the issue of local minima in neural networks while maintaining fitting quality through a new module termed the ANN-error model. We also include features relating to the two atoms of the final element, which are critical for accurate band gap energy prediction. With the addition of feature sets and our suggested methodology, the prediction error is lowered by 0.50 eV in comparison to the findings that have been previously published, demonstrating the training dataset's consistency in producing meaningful predictions. Furthermore, we evaluate our method using previous research data, and we find that all five-feature selections have better prediction errors, although they are somewhat greater than those found in the original study using the original 28 features. This underscores the strong correlation between the features of the second atom for the third element (C) in the compounds under examination and the band gap energy. In the final section, we maximize the training set size by training the model on all labeled data (46 compounds) and testing it on a larger unlabeled dataset (266 compounds). While the real band gap energy of these unlabeled data cannot be verified, we observe a decrease in the number of invalid predictions (negative values) to 3% compared to previously reported results.

Subsequently, the second study employs computational intelligence models to predict and demonstrate the empirical relationship between a collection of descriptors and crystal lattice properties, utilizing two families of perovskites, ABC_3 and A_2BCO_6, as reference compounds. By using the simple perovskite ABC_3 as the consideration material, we propose a hybrid model based on ANN, fuzzy logic type 1, and metaheuristic algorithms. Fuzzy logic and artificial neural networks are often and effectively used in engineering domains for prediction tasks. This section aims to perform a thorough analysis of hybrid models in order to identify the optimal model with a high degree of accuracy for lattice constant prediction. This part's main goal is to examine the effectiveness of PSO, IWO, and ICA metaheuristic algorithms, which are important for adjusting the ANN prediction model's parameters. After conducting a thorough comparison analysis, we have shown that PSO algorithms are effective in modifying the ANN model's parameters, which are used for predicting the lattice constant, according to a number of criteria. Furthermore, the second goal is to assess how well prediction models based on artificial neural networks and fuzzy logic perform. According to the performance evaluation, the PSO-Fuzzy logic outperforms the PSO-ANN, with percentage improvements in terms of RMSE and MAE criteria of 70.90% and 71.26%, respectively. Moreover, when compared to current models, the PSO-ANN and PSO-Fuzzy

models that were built perform better. Concerning the third study based on $ABCO_6$ materials, the ELM method was used to predict their lattice constant. Compared to the existing model, the proposed model demonstrates superior performance across various criteria. Due to its efficacy and simplicity, the developed ELM model holds significant promise for diverse property prediction. Because of their excellent performance, both suggested models should prove to be very helpful for a wide range of technical applications, such as heteroepitaxial film fabrication and solar cells.

Finally, based on the DFT calculation, we present new results concerning the lattice parameter and the band gap energy of the double perovskite compound Ba_2BiTaS_6. These results are obtained using the FP-LAPW method with the TB-mBJ-GGA approximation. In addition, the present calculation shows that the Ba_2BiTaS_6 compound has an indirect gap with a value of 1.67 eV.

References

[1] J.Keller, K. V.Sopiha, O.Stolt ,L.Stolt & al. (2020). Wide-gap (Ag, Cu)(In, Ga) Se2 solar cells with different buffer materials—A path to a better heterojunction. Progress in Photovoltaics: Research and Applications, 28(4), 237-250.

[3] Y.Zeng, S. J.Chua & P. Wu, (2002). On the prediction of ternary semiconductor properties by artificial intelligence methods. Chemistry of materials, 14(7), 2989-2998,

[3] F.Khmaissia, H. Frigui, M. Sunkara, J. Jasinski & al. (2018). Accelerating band gap prediction for solar materials using feature selection and regression techniques. Computational Materials Science, 147, 304–315.

[4] P.Villars, K. Brandenburg, M. Berndt, S.LeClair & S. Iwata, (2000). Interplay of large materials databases, semi-empirical methods, neuro-computing and first principle calculations for ternary compound former/nonformer prediction. Engineering Applications of Artificial Intelligence, 13(5), 497-505.

[5] P.Villars, K.Brandenburg, M.Berndt, S.LeClair & S. Iwata, (2001). Binary, ternary and quaternary compound former/nonformer prediction via Mendeleev number. Journal of alloys and compounds, 317, 26-38.

[6] P.Villars, K.Cenzual, J.Daams, Y. Chen & S. Iwata, (2004). Data-driven atomic environment prediction for binaries using the Mendeleev number: Part 1. Composition AB. Journal of alloys and compounds, 367(1-2), 167-175.

[7] S. K. Lahiri & K. C. Ghanta, (2009). Artificial neural network model with the parameter tuning assisted by a differential evolution technique: The study of the hold up of the slurry flow in a pipeline. Chemical Industry and Chemical Engineering Quarterly/CICEQ, 15(2), 103-117.

[8] S.Azizi, M. M. Awad & E. Ahmadloo, (2016). Prediction of water holdup in vertical and inclined oil–water two-phase flow using artificial neural network. International Journal of Multiphase Flow, 80, 181-187.

[9] E.Jorjani, S. C.Chelgani & S. H. Mesroghli, (2008). Application of artificial neural networks to predict chemical desulfurization of Tabas coal. Fuel, 87(12), 2727-2734.

[10] S. A. Kalogirou, (2001). Artificial neural networks in renewable energy systems applications: a review. Renewable and sustainable energy reviews, 5(4), 373-401.

[11] C.Suh, A.Rajagopalan, X.Li & K. Rajan, (2003). Combinatorial Materials Design through Database Science. MRS Proceedings, 804.

[12] P.Dey, J.Bible, S.Datta, S.Broderick & al. (2014). Informatics-aided bandgap engineering for solar materials. Computational Materials Science, 83, 185–195.

[13] Feng.ZHANG, (2011).Cross-validation and regression analysis in high-dimensional sparse linear models. Stanford University.

[14] Ali R.SYED, (2011) A review of cross validation and adaptive model selection.

[15] C. R.Shi & R. Adnan, (2014). Modified cross-validation as a method for estimating parameter. AIP Conference Proceedings.

[17] K.Balasamy& S.Ramakrishnan, (2018). An intelligent reversible watermarking system for authenticating medical images using Wavelet and PSO. Cluster Computing, 22(S2), 4431–4442.

[18] A. H.Gandomi, X.-S.Yang, S.Talatahari & A. H. Alavi, (2013). Metaheuristic Algorithms in Modeling and Optimization. Metaheuristic Applications in Structures and Infrastructures, 1–24.

[19] A.Kaur& Y. Kumar, (2021). A new metaheuristic algorithm based on water wave optimization for data clustering. Evolutionary Intelligence, 15(1), 759–783.

[20] J.Kennedy & R.Eberhart, (n.d.). Particle swarm optimization. Proceedings of ICNN'95 - International Conference on Neural Networks.

[21] A. R.Mehrabian & C.Lucas, (2006). A novel numerical optimization algorithm inspired from weed colonization. Ecological Informatics, 1(4), 355–366.

[22] R. Ubic, (2007). Revised Method for the Prediction of Lattice Constants in Cubic and Pseudocubic Perovskites. Journal of the American Ceramic Society, 90(10), 3326–3330.

[23] V.Sidey, (2019). A simplified empirical model for predicting the lattice parameters of the cubic/pseudocubic perovskites. Journal of Solid State Chemistry, 279, 120951.

[24] T. O. Owolabi, (2020). Extreme learning machine and swarm-based support vector regression methods for predicting crystal lattice parameters of pseudo-cubic/cubic perovskites. Journal of Applied Physics, 127(24).

[25] D.Kukolj, (2002). Design of adaptive Takagi–Sugeno–Kang fuzzy models. Applied Soft Computing, 2(2), 89–103.

[26] A.Majid, A.Khan, & T.-S Choi. (2011). Predicting lattice constant of complex cubic perovskites using computational intelligence. Computational Materials Science, 50(6), 1879–1888.

[27] Y.Zhang & X.Xu, (2020). Machine learning lattice constants for cubic perovskite $A_2BB'O_6$ compounds. CrystEngComm, 22(38), 6385–6397.

[28] Y.Wang, F.Cao & Y. Yuan, (2011). A study on effectiveness of extreme learning machine. Neurocomputing, 74(16), 2483–2490.

[29] Z.-F.Liu, L.-L.Li, M.-L.Tseng & M. K. Lim, (2020). Prediction short-term photovoltaic power using improved chicken swarm optimizer - Extreme learning machine model. Journal of Cleaner Production, 248, 119272.

[30] S.Dimitrovska, S.Aleksovska & I. Kuzmanovski, (2005). Prediction of the unit cell edge length of cubic $A_2BB'O_6$ perovskites by multiple linear regression and artificial neural networks. Open Chemistry, 3(1), 198–215.

[31] Y.-Y.Sun, M. L.Agiorgousis, P. Zhang & S. Zhang, (2015). Chalcogenide Perovskites for Photovoltaics. Nano Letters, 15(1), 581–585.

[32] W.Meng, B.Saparov, F.Hong, J. Wang & al. (2016). Alloying and Defect Control within Chalcogenide Perovskites for Optimized Photovoltaic Application. Chemistry of Materials, 28(3), 821–829.

[33] S.Perera, H.Hui, C.Zhao, H. Xue & al. (2016). Chalcogenide perovskites – an emerging class of ionic semiconductors. Nano Energy, 22, 129–135.

[34] S.Niu, H.Huyan, Y.Liu, M.Yeung & al. (2016). Bandgap Control via Structural and Chemical Tuning of Transition Metal Perovskite Chalcogenides. Advanced Materials, 29(9). Portico.

[35] M.Ju, J.Dai, L. Ma & X. C. Zeng, (2017). Perovskite Chalcogenides with Optimal Bandgap and Desired Optical Absorption for Photovoltaic Devices. Advanced Energy Materials, 7(18).

[36] P.Blaha, K.Schwarz, F.Tran, R. Laskowski & al. (2020). WIEN2k: An APW+lo program for calculating the properties of solids. The Journal of Chemical Physics, 152(7).

[37] Omkar Jani, (2008). ‘’Development of wide band gap InGaN solar cells for high-efficiency photovoltaics‘’, Doctorat thesis, Georgia Institute of Technology,
[38] Omkar Jani et al, (2007)‘’Effect of phase separation on performance of III-V nitride solar cells‘’, 22nd European Photovoltaic Solar Energy Conference, Milan, Italy.

[39] C.Honsberg, W. A.Doolittle & I. Ferguson, (2010). Novel High Efficiency Photovoltaic Devices Based on the III-N Material System: December 7, 2005 - June 6, 2008. Office of Scientific and Technical Information (OSTI).

[40] Q.Sun, H. Chen & W. J. Yin, (2018). Do chalcogenide double perovskites work as solar cell absorbers: a first-principles study. Chemistry of Materials, 31(1), 244-250.

GENERAL CONCLUSION

Conclusion

The research conducted in this thesis was devoted to studying various properties of materials. This was performed using two computational axes: artificial intelligence -based algorithms, and the first principle method, based on the linearized augmented plane wave method with a total potential (FP-LAPW) within the framework of density functional theory (DFT), implemented in the Wien2k code.

To accelerate data-driven material property discovery, we presented a new approach. We developed computational intelligence methods guided by fundamental material science principles to predict and simulate important properties of various materials. In particular, we concentrated on the following two crucial aspects of machine learning: learning algorithms (i) and optimization algorithms (ii). Accordingly, we introduced a new model called the error model and combined different optimization techniques with ANN and Fuzzy logic type-1 methods, to build robust regression models. These models take into account the intrinsic properties of the training data as well as the impacts of the descriptors in all atoms that constitute the matter.

Firstly, we presented an original approach that combined particle swarm optimization methods with artificial neural networks. Our approach used a novel bloc, called the ANN-error model, to solve the problem of local minima in neural networks while preserving fitting quality. It was successfully applied and tested to enhance the prediction of chalcopyrite's band gap. By using feature sets and our proposed methodology, we significantly improved previous results in informatics-based band gap prediction. Compared to previously published findings, the prediction error value is lowered by 0.50 eV, indicating the training dataset's consistent ability to produce meaningful predictions. Moreover, based on prior research data, we assessed our strategy and determined that while all five-feature selections had better prediction errors, they were still relatively higher than those obtained in the original study with the original 28 descriptors. This highlights the close relationship between the band gap energy and the characteristics of the second atom for the third element (C) in the researched compounds. Accordingly, the tests demonstrated that our proposed method may increase band gap prediction accuracy with the same dataset.

Regarding the second part, we thoroughly analyzed hybrid models to identify the most accurate model with a significant level of accuracy for the lattice constant prediction of cubic

and pseudo-cubic perovskites. We demonstrated that PSO algorithms are efficient in altering the ANN model parameters. These parameters were utilized to forecast the lattice constant based on many criteria following the performance of an extensive comparative investigation. Evaluating prediction models based on Fuzzy logic and artificial neural networks was our second objective. The performance evaluation indicated that the PSO-Fuzzy logic performed better than the PSO-ANN. Additionally, the built-in PSO-ANN and PSO-Fuzzy models outperformed the existing models. Thus, hybrid systems that integrate artificial neural networks and particle swarm optimization, or Fuzzy logic, surpass the performance of other models.

In addition, the empirical link between the crystal lattice parameter and a set of descriptor properties of A_2BCO_6 oxide perovskites was predicted and explained using Extreme Learning Machine (ELM). The suggested ELM model performed better than the current model according to several indicates accuracy. The created ELM model offered remarkable potential for various technical applications because of its effectiveness and simplicity. Finally, we presented a detailed ab-initio study of the structural and electronic properties of the double perovskite compound Ba_2BiTaS_6. For this, we used the linearized augmented plane wave method (FPLAPW) in the framework of density functional theory (DFT), with a combination of approximations such as TB-mBJ by Tran and Blaha. This work's contribution lies in its examination of the theoretically untreated compounds. Our choice was based on the paucity of theoretical work on this material. The lattice parameters of this material, which were obtained through our calculations, are in good agreement with the experimental values. For the electronic properties, the used method provided an excellent description of the energy band structure for this material, where the energy gaps have an indirect nature, with a value of 1.67 which is in good agreement with the experimental measurements.

Comparatively, the AI computational tools demonstrate a significant acceleration in the discovery and prediction of various properties with a good accuracy when contrasted with other methods reliant on Density Functional Theory (DFT). The former typically requires substantially less computational time, rendering it more efficient, while DFT-based approaches may require between four days and a week for the analysis of a single material property. This efficiency shaped our motivation to utilize machine learning tools as a basis of this work.

Although our computational intelligence approaches have shown promising results, better results can still be achieved. In what follows, we present primary issues for future investigation to expand on the present work. The biggest obstacle we have encountered was the scarcity of data. Increases in the amount of labeled training data can inevitably lead to more accurate predictions. We have looked at the possibilities of adding more hosts, features, co-dopants, and data augmentation techniques to our data sets. This can support the development of stronger models and, in particular, the formulation of more perceptive conclusions.

It is difficult to generalize a specific regression model, particularly when the data's intrinsic structure is not enough. Using the semi supervised technique may be highly beneficial when there is a limitation in the available labeled data, as in our applications. Examining the distributions of the training and testing data sets provides additional information about the generalized model's validity. Before the data mining process, it facilitates the detection of outliers and the prevention of performance degradation caused by error propagation and model misspecification.

Techniques for group learning comprise another important subject for future research. Combining numerous regression models is necessary to create a robust regression model that considers both the inherent characteristics of the training data and the effects of various perturbations. Thus far, we have experimented with many regression methods and examined their consistency. However, depending on how it considers the intrinsic data structure, each approach has its advantages and drawbacks. By eliminating the shortcomings of the basis models, integrating these various approaches with the right fusion techniques might enhance the ensemble's performance and other prediction techniques will be also used. Finally, based on a DFT study for the proposed compound Ba_2BiTaS_6, we will predict other optical properties (such as dielectric function, refractive index, absorption, reflectivity, and optical conductivity) to confirm that this component has potential in technological applications (including optoelectronic sensors and the manufacturing of optical components for controlling the polarization state of light, and for solar energy conversion applications).