REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
University Mohamed Boudiaf - M'SILA

Faculty: Mathematics & Computer Science

THESIS

Submitted in partial fulfillment of the requirements for the degree of DOCTOR

Field: Mathematics & Computer Science
Specialty: Computer Science

By:
WALID BEN ELMIR

Subject:

Optimizing Blood Donor Management System in Algeria

Acknowledgements

First and foremost, I express my deepest gratitude to Allah, the Almighty, for granting me the strength, patience, and perseverance to complete this research journey. Without His guidance and blessings, none of this would have been possible.

I also extend my sincere thanks to my supervisor, **Dr. Hemmak Allaoua**, for their guidance, insightful feedback, and constant support. Their expertise was invaluable to the progress and success of this thesis.

I am equally grateful to my co-supervisor, **Dr. Benaoumeur Senouci**, for their valuable advice and encouragement, which greatly enriched this work.

I also appreciate the collaboration of all staff involved in the practical deployment of the system and those who supported the experimentation phase.

Special thanks are due to the national health authorities for their assistance in providing data and logistical support.

Finally, I wholeheartedly thank my family and friends for their unwavering encouragement and patience throughout this academic journey.

Abstract

Blood supply chain management plays a vital role in maintaining the availability, quality, and safety of blood products, particularly in developing healthcare systems. This thesis proposes and evaluates a smart, AI-powered platform for blood donor management in Algeria, integrating predictive analytics, donor classification, and scheduling optimization. The platform was deployed and tested in a major blood donation center in Algiers, using data from 2017 to 2021. Through the application of machine learning algorithms such as Random Forest, ARIMA, and SVM, the system demonstrated improvements in forecasting blood demand (with a MAPE of 6.7%), classifying donor return behavior, and optimizing appointment scheduling (reducing no-shows by 23%). The platform also includes real-time dashboards and automated reporting tools. The findings show that AI-driven platforms can significantly enhance operational efficiency, reduce blood wastage, and ensure better preparedness in resource-constrained environments

Keywords: Blood Supply Chain, Artificial Intelligence, Machine Learning, Forecasting, Donor Classification, Scheduling Optimization.

الملخص

تُعد إدارة سلسلة إمداد الدم عنصرًا محوريًا لضمان توفر وسلامة منتجات الدم، لا سيما في الأنظمة الصحية في البلدان النامية. تهدف هذه الأطروحة إلى تصميم وتقييم منصة ذكية لإدارة المتبرعين بالدم في الجزائر، تعتمد على تقنيات الذكاء الاصطناعي والتعلم الآلي، وتشمل نماذج لتنبؤ الطلب، وتصنيف المتبرعين، وتنظيم مواعيد التبرع. تم اختبار المنصة ميدانيًا في مركز التبرع بالدم في الجزائر العاصمة باستخدام بيانات حقيقية للفترة الممتدة من 2017 إلى 2021. أظهرت النتائج تحسنًا كبيرًا في دقة التنبؤ بالطلب (بنسبة خطأ مطلق متوسط بلغت 6.7%)، وزيادة في فعالية تصنيف المتبرعين المتوقع عودتهم، وتقليص معدلات الغياب عن المواعيد بنسبة 23%. كما وفرت المنصة لوحات قيادة تفاعلية وتقارير مؤتمتة. تؤكد هذه الدراسة أن الحلول الذكية المعتمدة على الذكاء الاصطناعي قادرة على تحسين كفاءة إدارة الدم وتقليل الهدر وتحقيق استجابة أفضل في بيئات محدودة.

الكلمات المفتاحية : سلسلة إمداد الدم، الذكاء الاصطناعي، التعلم الآلي، التنبؤ بالطلب، تصنيف المتبرعين، تحسين الجدولة.

Résumé

La gestion de la chaîne d'approvisionnement en sang est cruciale pour assurer la disponibilité et la sécurité des produits sanguins, en particulier dans les systèmes de santé en développement. Cette thèse propose une plateforme intelligente de gestion des donneurs de sang en Algérie, intégrant l'intelligence artificielle pour la prévision de la demande, la classification des donneurs et l'optimisation des rendez-vous. Testée dans un centre majeur de don à Alger avec des données réelles de 2017 à 2021, la plateforme utilise des algorithmes tels que Random Forest, ARIMA et SVM. Les résultats ont montré une amélioration de la précision des prévisions (MAPE de 6,7 %), une réduction des absences aux rendez-vous (23 %), et une meilleure gestion des donneurs. La plateforme comprend aussi des tableaux de bord interactifs et des outils de rapport automatisés. Les résultats soulignent l'efficacité des solutions basées sur l'IA pour transformer la gestion du sang dans les environnements à ressources limitées.

Mots clés: Chaîne d'approvisionnement en sang, Intelligence Artificielle, Apprentissage Automatique, Prévision, Classification des donneurs, Optimisation des rendez-vous.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TERMS AND ABBREVIATIONS

CHAPTER 1

INTRODUCTION

1.1 Introduction

Blood transfusion services are critical components of healthcare systems worldwide, often determining patient outcomes in surgeries, trauma care, and chronic disease management. Ensuring a stable and safe blood supply remains a significant challenge, particularly in developing countries where resource constraints and logistical inefficiencies prevail.

Recent studies have demonstrated the effectiveness of artificial intelligence (AI) in predicting blood demand and optimizing supply chains. For instance, Kwon et al. (2024) developed a blood-demand prediction algorithm based on national public big data, providing medical facilities with appropriate blood volumes ahead of time. Similarly, Li et al. (2021) proposed an integrated strategy for short-term demand forecasting and ordering for red blood cell components, resulting in reduced inventory levels and ordering frequency.

In Algeria, the blood donation system faces recurring issues such as unpredictable donor turnout, inefficient scheduling, and inadequate demand forecasting. These challenges frequently lead to either critical shortages or wastage of blood units, which have a limited shelf life. Moreover, blood donation efforts are predominantly campaign-based and lack the support of intelligent planning systems that leverage data for proactive decision-making.

To address these challenges, this research proposes the development of a smart blood donor management platform. This platform integrates AI and machine learning technologies to forecast blood demand, classify donor behavior, and optimize donation scheduling. By building upon methodologies proven effective in other contexts, the proposed platform aims to enhance the efficiency of blood donation centers and contribute to the resilience of the national healthcare system.

1.2 Research Context

1.2.1 Global Perspective

Blood transfusion services are critical in modern healthcare systems, playing a vital role in surgeries, trauma care, cancer treatment, and chronic illness management. Ensuring a stable, safe, and timely supply of blood is a global challenge, often complicated by issues such as blood perishability, demand variability, and dependency on voluntary donations.

1.2.2 Local Perspective: The Algerian Context

In Algeria, the blood donation system is heavily dependent on sporadic, campaign-based donations, leading to irregular supply levels. The lack of predictive analytics, donor retention strategies, and intelligent scheduling tools often results in both shortages and wastage. These challenges highlight the urgent need for a smart, data-driven solution adapted to Algeria's healthcare system.

1.3 Background and Motivation

1.3.1 Healthcare Supply Chain Management

Modern healthcare delivery depends heavily on robust and responsive supply chains. These systems manage a vast array of critical resources—from surgical instruments and medications to lifesaving products such as blood. Ensuring the timely availability and safe distribution of these resources is essential to sustaining clinical operations and improving patient outcomes.

1.3.2 Blood Supply Chain Management

Among all components of the healthcare supply chain, the Blood Supply Chain (BSC) stands out due to the unique characteristics of blood products: they are highly perishable, demand is variable, and safety standards are uncompromising. A typical BSC involves complex processes such as donor recruitment, blood collection, testing for transmissible diseases, component separation, inventory management, and delivery to hospitals under tight timelines.

In high-resource settings, these processes are often supported by advanced information systems and predictive analytics. However, in countries like Algeria, blood supply chain operations still rely on fragmented, paper-based workflows and static planning methods. This results in inefficiencies, unpredictable supply patterns, and elevated risk of shortages or wastage.

Addressing these challenges requires more than process refinement—it demands the integration of smart technologies capable of providing real-time insights, automating decisions, and adapting to fluctuating demand. This research builds upon that premise by exploring how Artificial Intelligence (AI)/Machine Learning (ML)-powered platforms can modernize blood management practices in a resource-constrained national context.

1.3.3 Blood Supply Chain Characteristics

The blood supply chain (BSC) is a unique segment of the healthcare supply chain characterized by perishability, tight safety regulations, and unpredictable demand. It includes processes such as donor recruitment, blood collection, testing, storage, and delivery—all under time-sensitive constraints.

1.3.4 Problem Statement

Despite the vital role of blood donation in saving lives, managing the blood supply chain remains a complex challenge in many developing countries. In Algeria, the current blood donation system is highly dependent on irregular and campaign-based donations, which lack consistency and long-term planning.

Several operational challenges exacerbate the problem:

- **Unpredictable demand:** There is no reliable system to forecast future blood needs, leading to either overstocking and expiration or severe shortages.
- **Donor retention issues:** Many donors do not return after their first donation, making it difficult to build a stable and reliable donor base.
- **Inefficient scheduling:** The absence of intelligent scheduling systems often causes appointment overlaps or low attendance rates.
- **Lack of data-driven decision-making:** Current management strategies do not rely on predictive models or donor behavior analytics.

These issues create a reactive system, where interventions occur only after problems surface.

To address these operational limitations, this thesis proposes a data-driven, intelligent solution that automates key processes in the BSC, including demand prediction, donor classification, and appointment scheduling. The smart platform, powered by AI and ML, aims to improve operational efficiency, reduce waste, and enhance the reliability of blood supply in the Algerian healthcare context.

Research Question: How can a machine learning-based smart platform enhance forecasting accuracy, improve donor retention strategies, and optimize appointment scheduling to strengthen the alignment of supply and demand in Algeria's blood supply chain?

1.4 Research Aims and Objectives

General Aim

To design, develop, and validate an AI-driven smart platform that enhances operational efficiency in Algeria's blood supply chain.

Specific Objectives:

1. To investigate and analyze the limitations of the current blood donation system in Algeria.
2. To design and implement forecasting models for predicting short- and medium-term blood demand using time-series and machine learning techniques.
3. To apply classification algorithms to segment donors based on historical behavior, frequency, and donation type.
4. To develop an intelligent scheduling module that reduces donor absenteeism and optimizes appointment flow.
5. To integrate all AI/ML components into a unified, web-based smart platform with real-time dashboards and decision support.
6. To validate the platform's performance and usability through deployment in a real-world setting (e.g., the Algiers Blood Donation Center).

1.5 Significance of the Study

The significance of this research lies in its multidisciplinary approach to solving a critical healthcare logistics problem in Algeria. The blood supply chain is a high-risk domain characterized by perishability, uncertainty in donor turnout, and fluctuating demand patterns. This study contributes to bridging the gap between healthcare operations and intelligent systems by applying AI and ML technologies to enhance predictive capabilities and decision-making processes.

From an academic perspective, the study contributes to the growing body of research on data-driven healthcare systems, particularly in low-resource environments. It integrates techniques from machine learning, time-series forecasting, and operational research within the context of public health.

From a practical perspective, the proposed smart platform provides:

- A real-time, adaptive system to manage donor data and appointment schedules.
- Improved accuracy in demand forecasting, thereby reducing both shortage and wastage of blood units.
- A scalable and transferable solution that can be adopted by other regional blood centers or public health systems.

The study also holds strategic importance for Algerian health authorities by offering a prototype that aligns with national digital transformation goals and can inform future public health planning.

1.6 Thesis Contributions

This thesis makes several original contributions to the field of intelligent healthcare systems, with a specific focus on blood supply chain management in resource-constrained environments. The key contributions are as follows:

1. **Contextual Adaptation:** Development of a smart blood donor management platform tailored to the Algerian healthcare context, addressing operational challenges specific to the local donation system.
2. **AI/ML Integration:** Application of machine learning models (e.g., Random Forest, ARIMA, and Support Vector Machines) for demand forecasting, donor classification, and behavioral prediction within a unified decision-support system.
3. **Scheduling Optimization:** Design and deployment of an intelligent appointment scheduling module aimed at reducing donor no-shows and improving service utilization.
4. **Real-world Implementation:** Deployment and validation of the proposed platform at a public blood donation center in Algiers, demonstrating tangible improvements in donor engagement and operational efficiency.
5. **Scalable Architecture:** Proposal of a modular, web-based platform architecture that supports future integration with national health information systems or expansion to other health facilities.

1.7 Structure of the Thesis

This thesis is organized into six chapters, each addressing a specific component of the research:

- **Chapter 1: Introduction**
  Introduces the research context, highlights the key challenges in blood supply chain management in Algeria, and presents the problem statement, objectives, significance, and contributions of the study.
- **Chapter 2: Literature Review**
  Surveys existing research on blood donation systems, supply chain optimization, AI/ML applications in healthcare, and identifies key gaps that this work aims to address.
- **Chapter 3: Background on Artificial Intelligence Models**
  Provides technical background on the machine learning and forecasting models used in the study, including ARIMA, SVM, Random Forest, and Neural Networks.
- **Chapter 4: Research Methodology**
  Describes the overall methodological framework (CRISP-DM), data collection and preprocessing steps, and the design of the smart platform.
- **Chapter 5: Implementation and Evaluation**
  Details the implementation of the smart platform, experimental setup, model performance evaluation, and real-world validation in a public blood center.
- **Chapter 6: Conclusion and Future Work**
  Summarizes the main findings, outlines the study's limitations, and proposes directions for future research and potential system expansion.

CHAPTER 2

LITERATURE REVIEW

The increasing complexity of blood supply chain (BSC) operations, combined with evolving demands on healthcare systems, necessitates a clear understanding of the existing research landscape. Before proposing an intelligent platform tailored to the Algerian context, it is essential to examine how similar challenges have been addressed in the literature, what methodological approaches have been employed, and where significant gaps remain.

This chapter critically reviews previous studies on BSC management, emphasizing both conventional and AI/ML-enhanced approaches. The insights derived from this review inform the methodological decisions and model selection for the smart platform developed in this thesis.

2.1 Introduction

The efficient management of the blood supply chain (BSC) has become increasingly important in both research and practice due to the life-critical nature of blood products and the operational complexity of handling them. This chapter presents a comprehensive review of recent literature related to the analysis and optimization of BSC processes, with a special focus on studies that utilize Artificial Intelligence (AI) and Machine Learning (ML) methods.

Over the past decade, academic interest in BSC management has grown significantly. As shown in Figure 2.1, there has been a notable increase in the number of published articles addressing various dimensions of blood inventory and supply chain systems, particularly between 2014 and 2020. This upward trend reflects a growing recognition of the strategic importance of blood logistics within healthcare systems.

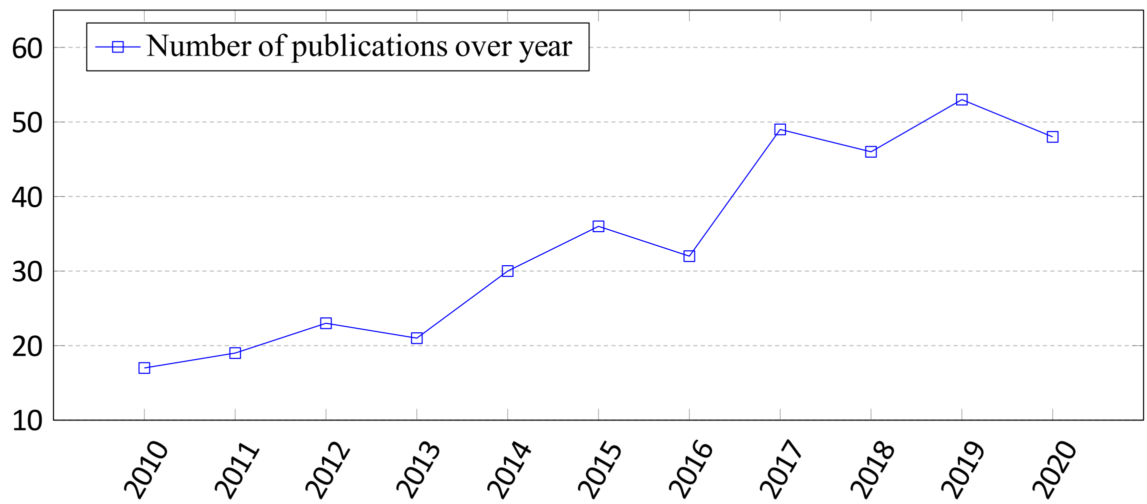


Fig. 2.1 The trend of paper publications in the BSC and blood inventory management between 2010 and 2020

A closer look at publication sources, as illustrated in Figure 2.2, reveals that the most active journals in this domain are *Transfusion*, *International Journal of Production Economics*, and *Vox Sanguinis*. While *Transfusion* and *Vox Sanguinis* focus predominantly on medical and clinical aspects of blood management, the *International Journal of Production Economics* highlights operations research and optimization techniques relevant to the BSC. The sharp drop-off in publication numbers beyond these top journals reflects a concentrated but growing body of research in this interdisciplinary field.

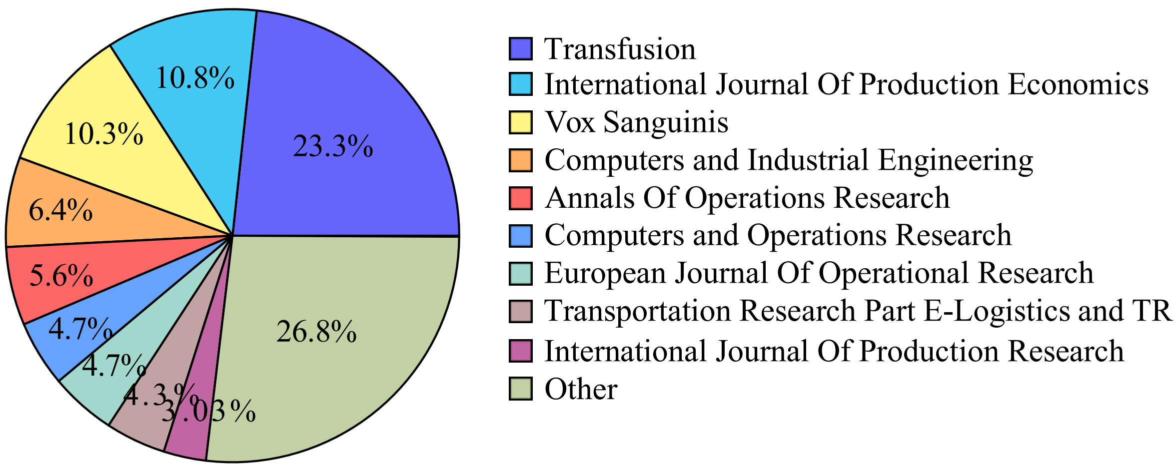


Fig. 2.2 Top nine active journals dealing with the idea of BSC in terms of number of articles published.

The review explores both traditional and data-driven approaches across key stages of the BSC, including blood collection, transportation, processing, inventory management, and transfusion. In doing so, it aims to:

- Identify the range of problem types addressed in BSC research;
- Examine the methodological frameworks used for decision-making and forecasting;
- Highlight existing gaps, particularly in the context of resource-constrained environments such as Algeria.

This literature review provides a foundation for the platform proposed in this thesis and supports the selection of forecasting, classification, and scheduling models in subsequent chapters.

2.2 Traditional and Data-Driven Approaches in BSC

Various modeling techniques have been employed to tackle the complexities of BSC management. These range from mathematical programming and simulation to modern AI/ML-based predictive systems.

2.2.1 Mathematical and Simulation Models

Traditional optimization techniques—such as linear programming, queuing theory, and discrete-event simulation—have been widely applied to problems related to inventory management, donor scheduling, and blood distribution. These models are valued for their interpretability and their capacity to capture logistics constraints.

2.2.2 AI and Machine Learning Applications in BSC

Recent applications of AI in BSC have leveraged time-series forecasting models such as ARIMA, Prophet, and LSTM to accurately predict blood demand across seasonal and emergency contexts.

Donor behavior prediction is another rapidly growing area, where classifiers like Random Forest, Support Vector Machines (SVM), and Gradient Boosting have been employed to identify likely repeat donors, enabling targeted retention strategies.

In terms of operations, intelligent scheduling using machine learning has shown promise in minimizing no-shows and balancing workload across donation centers.

Recent applications of Artificial Intelligence (AI) and Machine Learning (ML) in blood supply chain (BSC) management have demonstrated significant potential to overcome the limitations of traditional models. These methods offer greater adaptability, real-time responsiveness, and improved predictive capabilities—qualities especially valuable in resource-constrained healthcare environments.

Demand Forecasting: Time-series forecasting methods such as ARIMA, Prophet, and Long Short-Term Memory (LSTM) networks have been widely employed to capture the nonlinear, seasonal patterns in blood demand. These models can help blood centers anticipate shortages and reduce unnecessary overstocking, thus optimizing inventory levels and ensuring timely availability of blood products.

Donor Behavior Classification: Predictive modeling techniques—including Random Forests, Support Vector Machines (SVM), and Gradient Boosting—have been used to classify donors based on return likelihood, donation frequency, and eligibility status. Accurate classification models support targeted donor engagement strategies and improve donor retention rates.

Scheduling Optimization: Machine learning has also been applied to appointment scheduling, helping centers minimize no-show rates and reduce donor congestion. Reinforcement Learning and decision-tree-based algorithms are among the emerging approaches enabling dynamic, real-time adjustment of donation schedules.

Despite these advancements, few studies integrate these predictive capabilities into a unified operational platform—especially one tailored for environments with limited digital infrastructure, such as Algeria.

2.2.3 Inbound Problems

Inbound problems concern all processes prior to the distribution of blood, including donor recruitment, blood collection, initial testing, and inventory management. These issues are frequently exacerbated by structural limitations such as decentralized collection systems, inconsistent donor availability, and the lack of integrated data systems.

In Algeria, for example, blood collection centers often operate independently, with limited coordination across regions. This can result in oversupply in one location and critical shortages in another. Poor forecasting methods and manual scheduling further contribute to high rates of expired blood products.

Key inbound challenges include:

- Predicting donor turnout and ensuring a stable collection pipeline.
- Managing perishable inventory with a short shelf life.
- Ensuring timely and accurate testing under limited resources.

2.2.4 Outbound Problems

Outbound problems arise in the delivery and distribution of blood components to hospitals and transfusion units. These include transportation logistics, scheduling of deliveries, and responsiveness to emergency needs.

In geographically dispersed systems like Algeria's, transport inefficiencies and delayed dispatches frequently impact blood availability in rural and peri-urban hospitals. Additionally, the absence of real-time communication between collection centers and transfusion units hinders the alignment of supply with clinical demand.

Key outbound challenges include:

- Route optimization under time-critical conditions.
- Matching supply quantities with dynamic hospital needs.
- Avoiding delivery duplication and minimizing delays.

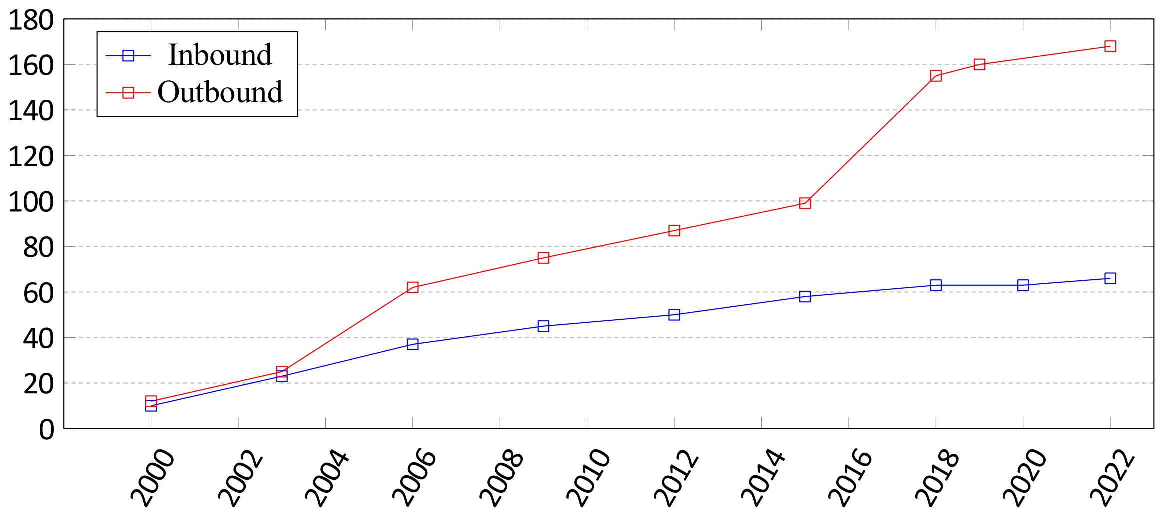


Fig. 2.3 Trends in Inbound versus Outbound Problems

Figure 2.3 illustrates a growing interest in outbound challenges over the past decade, reflecting the increasing complexity of last-mile delivery and emergency response in modern healthcare environments.

2.3 Research Streams and Methodological Trends in BSC

Research on the blood supply chain (BSC) has expanded considerably over the past decade, covering a wide range of operational and strategic problems. This section provides an overview of the main research streams and methodological approaches used in the literature. Rather than analyzing every individual study, we focus on categorizing prior works based on two key dimensions:

- The primary problem domain addressed (e.g., inventory control, transportation, donor behavior, forecasting).
- The type of methodology employed (e.g., optimization, simulation, machine learning, decision support systems).

Table 2.1 summarizes representative studies from each category, highlighting how different techniques have been applied to various BSC challenges.

Table 2.1 Classification of BSC research based on problem domain and methodology

Problem Domain	**Methodologies**	**Representative Studies**
Inventory Management	Optimization models, Simulation	Meneses, Santos and Barbosa-Póvoa (2023), Ahmadimanesh et al. (2023) Rahiminia et al. (2025)
Donor Behavior	Statistical analysis, ML classification	Epifani et al. (2023), Wu et al. (2022), Kauten et al. (2022)
Forecasting Demand	Time series models, ML forecasting	Twumasi and Twumasi (2022), Sarvestani et al. (2022)
Distribution and Transportation	Routing algorithms, Heuristics	Mansur et al. (2025), Ganesh and Narendran (2007), Agac et al. (2023)
Decision Support	DSS architectures, Fuzzy systems	Kayvanfar et al. (2024), Casal-Guisande et al. (2023)

This classification illustrates both the diversity and fragmentation of research in the field. While optimization and simulation dominate inventory-related work, machine learning is increasingly being applied in donor classification and demand forecasting. However, few studies address the integration of these techniques into a unified platform, especially in low-resource settings.

This gap underlines the contribution of this thesis, which seeks to combine forecasting, classification, and scheduling into an intelligent, deployable solution tailored to the Algerian context.

2.3.1 Traditional BSC Management Approaches

Before the emergence of Artificial Intelligence (AI) and Machine Learning (ML) techniques, the management of the blood supply chain (BSC) was largely addressed using conventional models—primarily based on mathematical programming, simulation, and heuristics. These traditional approaches focused on optimizing individual operations without integrated data-driven intelligence.

A systematic review conducted by (Pirabán et al. 2019) identified six key operational processes commonly analyzed in the literature: blood collection, transportation, component processing, inventory management, testing, and transfusion. These represent the core activities within the BSC and are often studied in isolation or in limited combinations.

Figure 2.4 shows the relative frequency with which each of these processes has been addressed in prior studies. Notably, inventory management dominates the focus of traditional research, reflecting its critical role in reducing waste and managing perishability—particularly for components such as platelets. In contrast, processes such as component processing and testing have received relatively less standalone attention, suggesting a potential research gap in the operational granularity of BSC modeling.

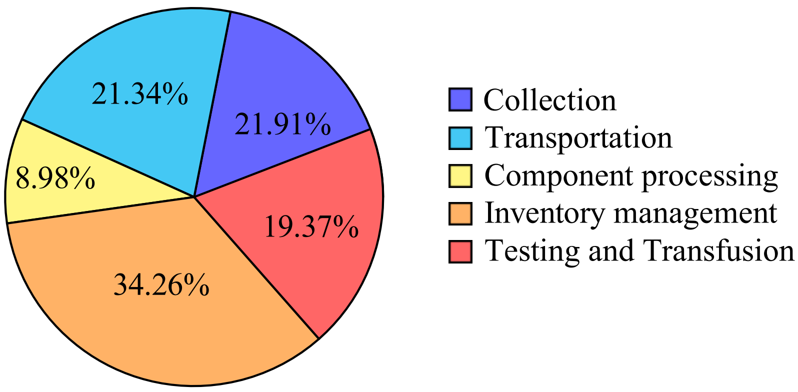


Fig. 2.4 Frequency of the papers referenced based on the processes in the BSC.

In the following subsections, we provide a detailed review of each operational process as addressed in conventional, non-AI literature, highlighting representative models and their practical contributions.

2.3.1.1 Collection

The collection of blood is a foundational process in the blood supply chain (BSC), involving the recruitment of donors, planning of collection events, and coordination of mobile and fixed-site blood drives. While vital to the success of the entire chain, relatively few studies have focused specifically on optimizing this phase in isolation.

One notable example is the work of (Ayer et al. 2018), who developed a collection model aimed at improving the efficiency of cryoprecipitate unit acquisition. The authors proposed a hybrid approach combining dynamic programming and a greedy heuristic to optimize the timing and location of mobile blood collections. The model was implemented as a decision support tool (Decision Support Tool (DST)) in the American Red Cross Southern Region, yielding a 40% reduction in per-unit collection cost and significantly increasing the number of units processed within an 8-hour workday.

Additionally, (Nagurney and Dutta 2019) examined blood collection in tandem with transportation planning, emphasizing the need for integrated decision-making to balance collection site operations with downstream logistics.

These studies underscore the operational complexity and cost sensitivity of the collection phase and demonstrate that even modest optimization efforts can lead to substantial gains in efficiency and service quality.

2.3.1.2 Transportation

Transportation within the blood supply chain (BSC) involves the timely and secure movement of blood products from collection centers to testing laboratories, storage facilities, and ultimately to hospitals and transfusion points. This phase is highly sensitive to time, geographic constraints, and demand variability—particularly during emergency or disaster scenarios.

(Rezaei Kallaj et al. 2021) proposed a routing model to optimize blood transport to a crisis-affected city using buses and helicopters. The model accounted for the compatibility of blood groups, vehicle capacity, and travel time under disrupted infrastructure conditions—offering a robust solution for emergency logistics during disasters.

Beyond direct blood transport, several studies in humanitarian logistics provide valuable frameworks that can be adapted to the BSC context. For example, (Paul and Wang 2019) designed a robust system for allocating relief equipment under earthquake conditions, incorporating uncertainty in facility damage and casualty rates. Similarly, (Adarang et al. 2020) proposed a time- and cost-efficient delivery plan using ambulances and helicopters to reduce emergency medical response time. (Davoodi and Goli 2019) presented a comprehensive model for location, allocation, and vehicle routing tailored to rural disaster settings, employing decomposition methods for tractability and validating the model with a real-world case study.

While these studies vary in scope and focus, they collectively highlight the importance of robust, time-sensitive transportation planning in high-risk environments—lessons that are increasingly relevant for blood logistics in both routine and emergency contexts.

2.3.1.3 Component Processing

After collection, donated blood must undergo a series of processing steps to separate it into its primary components (red blood cells, platelets, plasma) and to ensure product safety through rigorous screening. This stage is crucial for maximizing the usability of each donation and minimizing the risk of transfusion-transmitted infections.

(El-Amine et al. 2018) proposed a robust optimization model to design effective blood screening strategies for resource-constrained donation centers. The model integrates both regret- and expectation-based objectives to manage uncertainty in test outcomes and infection risk. Compared to several FDA-approved testing schemes, the proposed approach significantly reduced the probability of distributing infected blood while maintaining cost-effectiveness. Moreover, the model demonstrated superior performance under conditions of data uncertainty and limited test reliability.

This work highlights the critical trade-off between safety, cost, and logistical feasibility in blood component processing—a dimension that remains underexplored in much of the conventional literature.

2.3.1.4 Inventory Management

Inventory management is one of the most extensively studied areas in the blood supply chain (BSC), given the high perishability of blood products and the critical consequences of both shortages and overstocking. Recent studies have focused on integrating multiple operational goals such as cost minimization, service level improvement, and donor satisfaction.

(Aghsami et al. 2023) proposed a comprehensive mathematical model that addresses multiple aspects of the BSC, including donor allocation, facility placement, inventory control, and queuing systems at collection centers. Their objective was to reduce total costs while maximizing donor satisfaction. The model was validated using a real-world case study and solved using a meta-heuristic algorithm, supported by sensitivity analyses and managerial insights.

In a related effort, (Meneses, Marques and Barbosa-Póvoa 2023) developed a two-stage stochastic programming model to optimize hospital ordering policies for blood components. The model accounted for uncertain demand, ABO-substitutions, and trade-offs between shortages, wastage, and holding costs. Results from a Portuguese hospital demonstrated that the model could effectively support tactical and operational decision-making by reducing inventory risk.

Furthermore, (Hosseini et al. 2023) introduced a multi-objective mixed-integer linear programming (MIP) model aimed at minimizing cost, transportation time, and unmet demand, while maximizing system reliability. Due to the model's complexity, the NSGA-II algorithm was employed, and its effectiveness was confirmed through extensive computational experiments and a real-world case study.

Collectively, these studies reflect the growing sophistication of inventory models in BSC research, moving from cost-centric approaches toward multi-criteria optimization frameworks that balance efficiency, responsiveness, and reliability.

2.3.1.5 Testing and Transfusion

Testing and transfusion represent the final and most patient-facing stages of the blood supply chain (BSC). Testing includes all safety screening procedures prior to product release, while transfusion involves the clinical administration of blood components to recipients. Both processes are subject to high regulatory scrutiny and are critical to ensuring patient safety.

(Cagliano et al. 2021) proposed a structured framework to improve blood bag traceability in hospital wards. Their approach integrates logistics risk identification with proactive mitigation strategies and the use of key performance indicators (KPIs) to monitor risk response. The method has practical implications for improving reliability and accountability in clinical settings.

Building on this work, (Cagliano et al. 2022) introduced an enhanced healthcare supply chain risk management framework tailored to blood transfusion. The model incorporates Fault Tree Analysis and causal mapping to better understand the origin and impact of adverse events across multiple BSC echelons. Tested in real-world settings, the approach was shown to improve the early detection and prevention of high-risk activities.

From a clinical audit perspective, (Wong et al. 2015) investigated risk factors associated with allogenic blood transfusion in total hip replacement (THR) surgeries. The audit identified older age, female sex, and preoperative anemia as primary predictors of transfusion need. It also found that spinal anesthesia reduced transfusion risk, while conventional blood-saving techniques were largely ineffective.

These studies highlight the multifaceted nature of risk in testing and transfusion stages from logistical vulnerabilities to clinical decision-making. They underscore the need for integrated, cross-disciplinary approaches to enhance both operational efficiency and patient safety within the BSC.

Concluding Remarks

The reviewed literature on traditional blood supply chain (BSC) management demonstrates substantial progress in modeling individual processes such as collection, transportation, inventory control, and transfusion. These studies have contributed valuable insights through optimization techniques, simulation models, and risk management frameworks. However, they often treat BSC operations in isolation, with limited integration across stages, and rely heavily on deterministic assumptions or static planning methods.

Moreover, many of the proposed solutions assume ideal data availability, fixed system parameters, and stable demand conditions—assumptions that are frequently violated in real-world contexts, especially in developing countries such as Algeria. In such environments, supply and demand are often volatile, infrastructure is fragmented, and operational data may be sparse or unreliable.

These limitations highlight the need for integrated, data-driven, and adaptive systems that can respond to real-time constraints—especially in fragile healthcare infrastructures such as Algeria's.

2.4 Classification of Research by Decision Environment

Studies in the blood supply chain (BSC) domain can be broadly classified based on the decision environment assumed in the modeling approach. The three primary categories are:

- **Deterministic Models:** These models assume all input parameters (e.g., demand, lead time, donation rate) are known with certainty. While computationally tractable and easier to implement, they often fall short in representing real-world complexity.
- **Stochastic Models:** These incorporate randomness and probabilistic variation, making them suitable for modeling uncertainty in demand, donor turnout, or supply disruptions. Stochastic models are widely used in inventory planning and emergency blood distribution.
- **Fuzzy Models:** These address ambiguity or vagueness in parameter values and decision criteria. They are particularly useful in situations where expert judgment or linguistic variables (e.g., "high risk", "sufficient stock") are involved, such as donor eligibility assessment.

Figure 2.5 illustrates the historical trend in the usage of deterministic versus uncertain decision environments (including stochastic and fuzzy models) in BSC research. While deterministic models dominated early years with steady but low publication rates, uncertain models began to see increased attention starting around 2013.

A sharp rise in uncertain model usage occurred between 2016 and 2021, coinciding with the growing availability of computational tools and a broader shift toward modeling real-world complexity. Although deterministic approaches also gained slightly during this period, the preference for uncertainty-aware frameworks became increasingly evident.

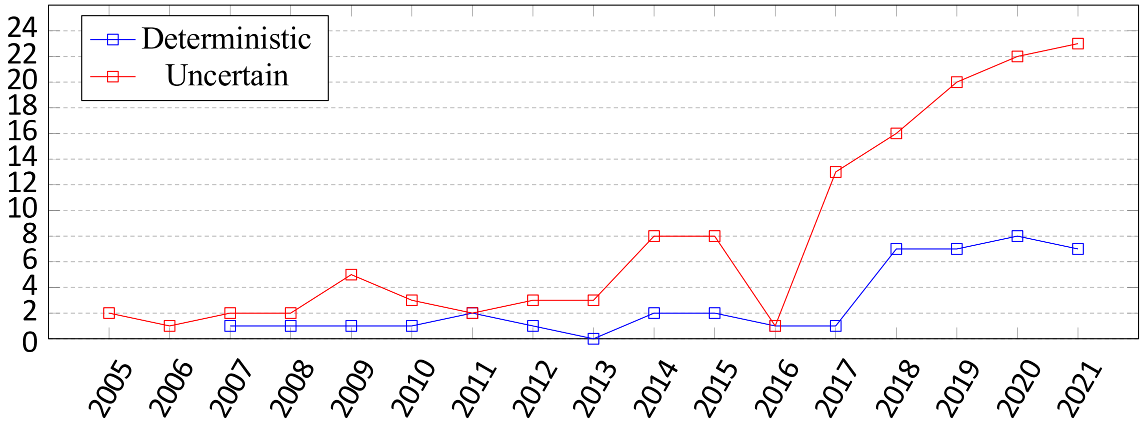


Fig. 2.5 The frequency and trends of the reference papers based on decision-making and forecasting environments

Understanding these decision environments is essential for choosing the right modeling framework. Given the inherent uncertainty in donor behavior and fluctuating blood demand in Algeria, this thesis emphasizes the use of data-driven and stochastic techniques to enhance planning robustness and responsiveness.

2.5 AI and ML-based Approaches in BSC

Recent years have seen a significant increase in the use of Artificial Intelligence (AI) and Machine Learning (ML) techniques in solving complex problems across healthcare and logistics domains. Within the blood supply chain (BSC), these techniques offer promising solutions to key challenges such as demand forecasting, donor classification, and real-time decision support.

To better understand how AI and ML have been utilized in this domain, we classify the literature according to their primary application within the BSC workflow. This includes forecasting methods for anticipating demand, classification models for managing donor behavior, and decision support systems for assisting operational planning. Each category reflects a distinct set of technical challenges and modeling approaches.

2.5.1 Forecasting Applications

Time series forecasting models enhanced by ML, such as Support Vector Regression (SVR), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) networks, have shown improved accuracy over traditional models in predicting blood demand Shokouhifar and Ranjbarimesan (2022), Li et al. (2021). These methods are capable of capturing nonlinear patterns and seasonal trends, which are common in donation and consumption cycles. For instance, Ghaderzadeh et al. (2023) proposed a hybrid deep learning model using LSTM combined with CNN to forecast blood demand during pandemic disruptions. Similarly, Rezaei et al. (2023) conducted a comparative study between classical ARIMA and deep neural networks, showing superior performance of neural models in handling demand volatility.

2.5.2 Donor Classification

Classification algorithms like Decision Trees, Random Forests, and Support Vector Machines have been applied to identify eligible donors, predict return behavior, and reduce deferrals. For instance, Kauten et al. (2021) developed predictive models for donor return using historical donation data, improving targeting for recruitment campaigns. Similarly, Wu et al. (2022) applied machine learning models during the COVID-19 pandemic to estimate willingness to donate, achieving high predictive accuracy. In addition, Epifani et al. (2023) proposed a Bayesian model to classify donors based on donation frequency and demographic variables, supporting personalized engagement strategies in blood collection centers.

2.5.3 Decision Support Systems

AI-driven Decision Support Systems (DSS) integrate multiple predictive and optimization modules to assist blood bank managers in scheduling, stock monitoring, and contingency planning. Fuzzy logic, rule-based engines, and hybrid AI systems are used to enhance interpretability and adaptability, particularly in dynamic or uncertain environments Alabdulkarim et al. (2019).

Despite these advances, the adoption of AI/ML in BSC remains limited in developing countries. Challenges include lack of data infrastructure, model generalizability, and stakeholder trust. As outlined by Ghaderzadeh and Asadi (2024), key barriers include poor digitization of health records, limited computing infrastructure, and insufficient training of medical personnel in AI-assisted decision making. Moreover, Salem et al. (2024) emphasize that ethical concerns, data sovereignty, and technology access gaps further hinder AI adoption in the Global South.

This thesis responds to these challenges by proposing the design and implementation of an integrated AI/ML-based platform tailored to the Algerian healthcare context. The system will include dedicated modules for demand forecasting, donor classification, and appointment scheduling—each powered by machine learning algorithms to support more accurate and responsive decision-making within blood banks.

By addressing context-specific limitations such as data scarcity, organizational fragmentation, and limited automation, this research seeks to demonstrate the feasibility and added value of AI-driven tools in enhancing the efficiency, equity, and resilience of blood supply chains in developing countries.

Concluding Remarks

The reviewed literature demonstrates a growing interest in applying Artificial Intelligence (AI) and Machine Learning (ML) techniques to blood supply chain (BSC) management, particularly in forecasting, donor classification, and decision support. These approaches offer clear advantages over traditional models in terms of adaptability, accuracy, and responsiveness to uncertainty.

However, despite promising results, the application of AI/ML in BSC remains fragmented and largely concentrated in high-resource settings. Challenges such as data availability, limited technical capacity, and contextual misalignment often hinder implementation in low- and middle-income countries. Moreover, most studies tend to address isolated problems, without integrating AI/ML across multiple decision layers.

These gaps underscore the need for holistic, deployable solutions that unify predictive modeling with operational planning in real-world settings. The platform proposed in this thesis aims to fill that gap by embedding AI/ML techniques into an integrated system that supports forecasting, donor profiling, and scheduling—tailored specifically to the Algerian healthcare infrastructure and data realities.

2.6 Review of Existing Blood Bank Management Systems

As a complement to the literature on conceptual and methodological developments in blood supply chain (BSC) management, this section presents a set of real-world blood bank information systems currently deployed in various countries. These systems demonstrate how digital tools—ranging from cloud computing to integrated hospital registries—have been operationalized to manage key aspects of the BSC, including donor registration, inventory tracking, testing, and reporting.

Each example reflects a different national approach, shaped by infrastructure capacity, policy environment, and healthcare system integration. Their functionalities and limitations provide useful insights for the design of the AI-powered platform proposed in this thesis.

2.6.1 BBIS: Blood Bank Information System Based on Cloud Computing (Indonesia)

(Ramadhan et al. 2019) proposed a linking system that could include Indonesian society to be personal donors for the sake of ensuring blood supply availability. The system is based on a web and mobile application utilizing cloud computing technology, allowing potential donors to register and manage their information. It includes features such as blood request applications, donor registration, and medical history management, aimed at supporting non-emergency donation activities.

2.6.2 OBDRMS: Online Blood Donation Reservation and Management System (Saudi Arabia)

(Hashim et al. 2014) proposed a blood-donation management system targeting donors, recipients, and hospitals in Jeddah. The platform allows users to access real-time blood type availability across hospitals, improving coordination and reducing human error in blood allocation. The system enhances transparency and efficiency in hospital-specific distribution operations.

2.6.3 BBMS: Blood Bank Management System (Malaysia)

(Sulaiman et al. 2015) developed a web-based platform for the Sultanah Nur Zahirah Hospital to manage blood bags and test results. The system follows the Inception–Elaboration–Construction–Transition model and supports inventory reporting, donor details management, and automated blood test evaluations for transfusion eligibility. It contributes to improving traceability and stock control within the hospital's blood bank.

2.6.4 Blood System in Australia: National Blood Authority (Australia)

To ensure the national availability and safety of blood products, the Australian National Blood Authority (NBA) operates several integrated ICT systems, including the Australian Bleeding Disorders Registry (ABDR), BloodPortal, BloodNet, MyABDR, and BloodSTAR (NBA 2023). These systems span the entire BSC—from donor to distribution—providing real-time data on inventory, product use, authorization, ordering, and contract performance. The systems support 24/7 accessibility for healthcare providers and stakeholders.

2.6.4.1 Comparative Insights

While the reviewed systems vary in scope and technological depth, they all aim to improve operational visibility, reduce errors, and streamline decision-making in blood services. However, most of them remain limited in terms of predictive intelligence and integration of AI/ML capabilities. Furthermore, none explicitly address the challenges of fragmented data infrastructure, low donor retention, or decentralized healthcare logistics—issues that are particularly pressing in the Algerian context. The solution proposed in this thesis builds on these lessons, aiming to combine digital integration with machine intelligence to deliver a more adaptive and context-sensitive platform.

2.7 Literature Review Findings

The review of existing literature in blood supply chain (BSC) management reveals several critical gaps that remain underexplored and insufficiently addressed. Foremost among these is the limited application of data-driven technologies and intelligent systems capable of agile, cost-effective, and scalable decision-making. While optimization models abound, few systems integrate real-time data collection, analysis, and adaptive control mechanisms—particularly in low-resource environments.

One major area requiring further investigation is workforce planning at blood donation centers. Most current studies neglect the dynamic nature of staff scheduling and its impact on donor satisfaction, service time, and operational productivity. Aligning staff availability with fluctuating donor behavior throughout the day is a practical yet overlooked dimension of BSC efficiency.

Another persistent challenge is balancing supply with demand. Oversupply, without predictive alignment to actual needs, often results in storage congestion, increased waste due to product expiry, and inefficiencies in downstream distribution. While inventory models have addressed this to some extent, real-time demand sensing and integrated planning remain limited.

Additionally, donor behavior modeling is largely absent from conventional BSC research. Factors such as frequency of donation, motivations, and preferred donation times or locations have rarely been studied in depth, yet they hold immense potential for improving donor retention and repeat participation.

Finally, despite substantial effort to incorporate uncertainty into BSC models—through stochastic and fuzzy approaches—the issue of unpredictability continues to plague practical implementations. Better uncertainty modeling, along with hybrid AI techniques, could offer more robust planning frameworks.

In conclusion, human blood remains a critical and perishable resource. Given the global reliance on voluntary donations, it is imperative that future research shifts toward intelligent, donor-centric, and operationally integrated solutions. Addressing these gaps not only enhances the resilience and efficiency of blood services but also supports more ethical and effective healthcare delivery.

2.8 Discussion and Chapter Outlook

This chapter has presented a structured and critical review of the literature on blood supply chain (BSC) management. It categorized existing research across key dimensions, including the type of operational problem addressed, the decision environment assumed, and the methodologies applied. Conventional models have proven useful in specific contexts, yet they often fall short in terms of integration, adaptability, and real-time responsiveness—particularly in resource-constrained environments.

In contrast, recent studies have begun to explore the potential of Artificial Intelligence (AI) and Machine Learning (ML) techniques in tackling long-standing challenges in the BSC. These approaches offer promising advances in forecasting, donor profiling, and decision support, but their adoption remains limited and fragmented, especially in low-resource settings like Algeria.

The chapter also surveyed existing blood bank management systems around the world, revealing a growing shift toward digitalization—but still lacking intelligent, predictive capabilities.

These findings collectively motivate the work presented in the next chapter, which focuses on the application of AI/ML in the blood donation domain. Chapter 3 outlines the core algorithms, models, and learning strategies used to enhance the planning, classification, and scheduling functions of a smart blood donor management platform.

2.9 Research Gaps and Opportunities

Despite the growing body of research on blood supply chain (BSC) management, several critical gaps remain—particularly in the context of low-resource environments such as Algeria.

1. Fragmentation of AI/ML Applications: Most existing studies focus on isolated aspects of the BSC, such as forecasting or inventory optimization. However, very few integrate multiple predictive components—such as demand forecasting, donor behavior classification, and scheduling—into a unified operational platform. This lack of integration limits the practical utility of AI/ML solutions in real-world healthcare settings.

2. Contextual Limitations: A significant portion of the literature is based on assumptions and datasets derived from high-resource healthcare systems. There is a clear scarcity of studies tailored to the constraints and operational realities of developing countries, where digital infrastructure, data availability, and resource allocation pose unique challenges.

3. Limited Real-World Implementation: While many models report high accuracy in experimental or simulated environments, very few are deployed or validated in real-world donation centers. The lack of field validation raises questions about the feasibility and robustness of proposed approaches under practical constraints.

4. Underutilization of Behavioral Data: Although donor behavior is acknowledged as a key factor in BSC sustainability, relatively few studies leverage longitudinal donor data to model return patterns, predict absenteeism, or optimize re-engagement strategies using intelligent systems.

Research Opportunity: These gaps present a clear opportunity to develop a comprehensive, AI-enabled platform that addresses multiple facets of BSC operations in an integrated manner. Specifically, there is a need for a solution that combines demand forecasting, donor classification, and scheduling optimization—tested and validated in a real-world, resource-constrained setting. This thesis aims to fill that void and contribute to both the academic and practical advancement of intelligent blood management systems.

2.10 Conclusion

The literature reviewed in this chapter highlights both the evolution and fragmentation in blood supply chain management strategies, especially regarding the use of AI and ML. These findings collectively motivate the work presented in the next chapter, which focuses on the application of AI/ML in the blood donation domain. Chapter 3 outlines the core algorithms, models, and learning strategies used to enhance the planning, classification, and scheduling functions of a smart blood donor management platform.

CHAPTER 3

AI/ML FOR BLOOD DONATION MANAGEMENT

Building upon the gaps and opportunities identified in the previous chapter, this chapter transitions from reviewing existing literature to defining the technical foundation of the proposed intelligent platform. It focuses on how Artificial Intelligence (AI) and Machine Learning (ML) techniques can be applied to enhance key components of blood donor management, particularly in forecasting demand and classifying donors.

The chapter presents and evaluates a range of modeling approaches that were selected based on their relevance to the Algerian context and the nature of the available data. These models serve as core building blocks of the AI-based decision support system introduced in later chapters.

3.1 Introduction

Artificial Intelligence (AI) and Machine Learning (ML) offer advanced capabilities for building predictive models that support complex decision-making across healthcare systems. Within blood supply chain (BSC) management, two key areas where such models can deliver significant impact are blood demand forecasting and blood donor classification.

Demand forecasting helps reduce inefficiencies such as shortages and overstocking by predicting blood usage patterns in advance. Inaccurate forecasts may amplify the Bullwhip Effect[1], leading to instability across the entire system.

[1]The Bullwhip Effect refers to demand variability amplification upstream in the supply chain due to distorted information or delayed signals.

ML models are particularly well-suited to handle non-linear patterns and variable dynamics in healthcare data. These models fall into two main categories:

- **Supervised learning**, which relies on labeled datasets with input-output mappings to train models for prediction tasks such as regression and classification.
- **Unsupervised learning**, which aims to discover hidden structures or patterns in data without predefined labels, often used in clustering and segmentation tasks.

This chapter applies both statistical and AI-based models to the challenges of blood demand forecasting and donor behavior modeling. Algorithms explored include Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), Support Vector Regression (SVR), Linear Regression (LR), Decision Tree (DT), Random Forest (RF), and others. These models are evaluated and selected based on their performance characteristics, ability to generalize, and alignment with the data conditions of the Algerian blood system.

The following sections provide a detailed examination of the forecasting and classification models used in this thesis, as core components of the proposed intelligent decision support platform.

3.2 Blood Demand Forecasting

One of the most critical challenges in managing blood supply chains (BSC) is the ability to predict future demand accurately. Blood demand is influenced by complex temporal patterns, seasonal effects, emergency events, and local dynamics—making simple forecasting models insufficient in many contexts.

This section explores two distinct families of forecasting approaches: traditional statistical time-series models and machine learning-based techniques. The objective is to compare their strengths, limitations, and applicability to the blood donation environment in Algeria. The insights gained from this comparative analysis guide the selection of appropriate models for integration into the intelligent platform proposed by this research.

3.2.1 Forecasting Objective and Importance

Forecasting blood demand is a foundational component of effective blood supply chain (BSC) management. Given the perishability of blood components—particularly platelets with their short shelf life—and the reliance on voluntary donations, any mismatch between demand and supply can result in critical shortages, wastage, or both. The goal of demand forecasting is to anticipate future needs accurately enough to support proactive planning across collection, storage, and distribution stages.

Inaccurate forecasts can exacerbate the Bullwhip Effect[2], leading to inflated collection, resource misallocation, and inventory instability (Liu et al. 2022, Alghamdi 2023). In contrast, precise forecasting enables more efficient workforce scheduling, optimized inventory levels, and better coordination between donation centers and hospitals.

In Algeria, the challenge is compounded by data sparsity, regional variability in demand, and underdeveloped information systems. These conditions necessitate forecasting models that are not only accurate but also robust to uncertainty and scalable to data-constrained environments.

This section lays the foundation for selecting and evaluating forecasting models—both traditional time-series and modern machine learning approaches—that are suitable for addressing these operational complexities.

3.2.2 Time-Series Forecasting Models

Time-series forecasting methods are widely used in operational research and healthcare analytics for predicting future values based on historical patterns. In the context of blood supply chains (BSC), these models serve as a valuable foundation due to their mathematical rigor, interpretability, and relatively low computational requirements.

While more advanced machine learning models may offer flexibility and higher predictive performance in certain contexts, time-series models remain a reliable choice—especially in data-constrained environments like those found in many regional Algerian healthcare settings. This section presents the classical forecasting models used in this study, serving as a statistical benchmark for later comparison with ML-based models.

3.2.3 Autoregressive (AutoReg) Model

The Autoregressive AutoReg (AR) model is a type of time series forecasting model that uses past values of the series to predict future values. In this model, the current value of the series is assumed to be a function of its past values. The AR model of order p, denoted by AR(p), can be expressed as:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + \epsilon_t \tag{3.1}$$

where y_t is the time series variable at time t, β_i ($i = 0, 1, 2, \cdots, p$) are the coefficients, ϵ_t is the error term at time t, and p is the order of the AR model.

[2] The Bullwhip Effect refers to the amplification of demand fluctuations as one moves upstream in the supply chain.

The goal of the AR model is to estimate the coefficients θ_i from the historical data and use them to predict the future values of the time series. The estimation of the coefficients is usually done using a method called maximum likelihood estimation.

In the context of blood demand forecasting in supply chain management, the AR model can be used to predict the future demand of blood based on its past demand. The order of the model p can be chosen based on the autocorrelation plot and partial autocorrelation plot of the time series. The predicted demand can then be compared with the actual demand to evaluate the accuracy of the model.

3.2.4 Autoregressive Moving Average (ARMA) Model

The most common time series models in statistics are ARMA models (Autoreg Ressive Moving Average). Given a time series X_t, the Autoregressive Moving Average (ARMA) model is a method to forecasts the future values of this series. The model is composed of two parts: AR and MA, autoregressive and a moving average, respectively. The model is generally denoted ARMA (p,q), where p is the order of the AR part and q is the order of the MA part. An autoregressive model of order p, abbreviated AR(p), is written:

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \tag{3.2}$$

where $\varphi_1, \ldots, \varphi_p$ are the model parameters, c is a constant, and ε_t is white noise. The constant is often omitted in the literature, with the process then being said to be centred.

Additional constraints on the parameters are necessary to guarantee stationarity. An autoregressive moving-average model of orders (p, q) is a discrete temporal process $(Y_t, t \in N)$ verifying:

$$Y_t = \varepsilon_t + \sum_{i=1}^{p} \varphi_i Y_{t-i} + \sum_{i=1}^{q} \vartheta_i \varepsilon_{t-i} \tag{3.3}$$

where φ_i and ϑ_i are the model parameters and the error terms ε_i.

The ARMA model assumes that the current value of the time series is a linear combination of its past values and the past errors, with the addition of some random noise. ARMA models are widely used in time series forecasting, as they can capture both short-term and long-term dependencies in the data.

3.2.5 Autoregressive Integrated Moving Average (ARIMA) Model

The ARIMA model is made popular by Box–Jenkins models. The acronym ”ARIMA” stands for ”AutoRegressive Integrated Moving Average”, with the ”I” representing ”Integrated”. This integration involves the differentiation of the time series data, which aims to eliminate the trends in the data to make it stationary.

The correction of a non-stationarity in terms of variance can be carried out by transformations of the logarithmic type (if the variance increases with time) or conversely exponential. These transformations must be carried out before differentiation. The Auto Regressive Moving Average (ARIMA) equation can be written as:

$$(1 - \sum_{i=1}^{p} \phi_i L^i)(1 - L)^d yt = c + (1 + \sum_{i=1}^{q} \vartheta_i L^i)\epsilon_t \tag{3.4}$$

where:

- L is the lag operator, such that $L^i y_t = y_{t-i}$
- y_t is the time series at time t
- p is the order of the autoregressive part (AR) of the model
- ϕ_i is the coefficient of the ith lag term in the AR part
- d is the degree of differencing used to make the series stationary
- q is the order of the moving average part (MA) of the model
- ϑ_i is the coefficient of the ith lag term in the MA part
- ε_t is the error term at time t
- c is a constant term

The equation describes how the current value of the time series y_t depends on its own past values (up to p lags) and on the past errors (up to q lags). The $(1 - L)^d$ term represents the differencing operation applied to make the series stationary. The coefficients ϕ_i and ϑ_i are estimated from the data using maximum likelihood estimation. The ARIMA model is therefore a combination of this differentiation process and the classical ARMA process.

ARIMA models have some limitations, such as:

1. Limited forecasting horizon: ARIMA models are not suitable for forecasting long-term trends, as they only model the next step in the series.
2. Sensitive to outliers: ARIMA models are sensitive to outliers and extreme values in the time series data, which can affect the accuracy of the forecasts.
3. Limited ability to capture seasonality: While ARIMA models can handle seasonal trends to some extent, they may not capture more complex seasonal patterns in the data.
4. Data requirements: ARIMA models require a sufficient amount of historical data to be accurate, which can be a limitation for newer or smaller data sets.
5. Model selection: Selecting the appropriate ARIMA model can be a complex and time-consuming process, requiring knowledge of statistical techniques and careful evaluation of model performance.

3.2.6 Seasonal ARIMA Model (SARIMA)

The Seasonal ARIMA (SARIMA) model is an extension of the ARIMA model that takes into account seasonal patterns in the time series data. It is suitable for time series data that exhibit regular patterns at fixed intervals, such as monthly or quarterly data. The Seasonal ARIMA (SARIMA) model is often used in forecasting applications, such as predicting monthly or quarterly sales figures for a business.

The (SARIMA model includes additional terms to account for the seasonal patterns, such as seasonal differences, seasonal autoregressive terms, and seasonal moving average terms. The model is typically denoted as SARIMA(p,d,q)(P,D,Q)m, where m represents the number of time periods in a season.

The SARIMA model can be expressed in the following equation:

$$\begin{aligned}&(1 - \phi_1 B - \phi_2 B^2 - \cdot\cdot\cdot - \phi_p B^p)(1 - \Phi_1 B^m - \Phi_2 B^{2m} - \cdot\cdot\cdot - \Phi_P B^{Pm})\nabla^d \nabla_s^D Y_t \\ &= (1 + \vartheta_1 B + \vartheta_2 B^2 + \cdot\cdot\cdot + \vartheta_q B^q)(1 + \Theta_1 B^m + \Theta_2 B^{2m} + \cdot\cdot\cdot + \Theta_Q B^{Qm})\epsilon_t\end{aligned} \tag{3.5}$$

where Y_t is the time series data, B is the backshift operator, ϕ_i and ϑ_i are the autoregressive and moving average coefficients at non-seasonal time periods, Φ_i and Θ_i are the autoregressive and moving average coefficients at seasonal time periods, p and q are the orders of the non-seasonal autoregressive and moving average terms, P and Q are the orders of the seasonal autoregressive and moving average terms, d and D are the orders of non-seasonal and seasonal differencing, and ϵ_t is the error term.

The SARIMA model is a strong tool to model and forecast seasonal time series. Nonetheless, the model may be complex and computationally prohibitive while estimating the parameters of the model, particularly on large data sets. Besides, in using a SARIMA model, it is presumed that seasonal patterns are fixed and not subject to variability over time, which need not necessarily be true in reality.

3.2.7 Seasonal Exponential Smoothing Model (SESM)

The Seasonal Exponential Smoothing model, sometimes called Holt-Winters' method, is a very popular time series forecasting technique that is applied in supply chain management, including blood demand forecasting. This model is an extension of the Exponential Smoothing model; it is used to forecast time series data that has a trend but without seasonality.

The Seasonal Exponential Smoothing model is used for time series data displaying both trend and seasonality. It incorporates seasonality in data and allows the computation of the forecast for more than one period ahead. The model is based on three components: level, trend, and seasonality. In the equation, the level component represents the average value of the time series, the trend component represents the overall direction of the time series, and the seasonality component represents repetitive patterns that occur over time. The equations for Seasonal Exponential Smoothing are as follows:

Level equation:

$$L_t = \alpha(y_t - S_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \tag{3.6}$$

Trend equation:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \tag{3.7}$$

Seasonal equation:

$$S_t = \gamma(y_t - L_t) + (1 - \gamma)S_{t-m} \tag{3.8}$$

Forecast equation:

$$\hat{y}_{t+h} = L_t + hT_t + S_{t-m+h_k} \tag{3.9}$$

where y_t is the actual value of the time series at time t, L_t is the level component at time t, T_t is the trend component at time t, and S_t is the seasonal component at time t. m is the number of seasons in a year, α, β, and γ are smoothing parameters, and h is the number of periods ahead for which the forecast is being made.

Compared to the ARIMA model, the Seasonal Exponential Smoothing (SESM) model is easier to understand and implement. It is also computationally less expensive and requires fewer parameters. However, it may not be suitable for time series data with complex patterns or long-term trends.

3.2.8 Multiplicative Holt–Winters Model

The Multiplicative Holt-Winters Model is a time series forecasting method that is used to forecast data with a seasonal component. This model is an extension of the Exponential Smoothing model and is also known as the Triple Exponential Smoothing model. It considers three components: level, trend, and seasonality.

The model uses a weighted average of past observations to predict future values, with the weights decreasing exponentially as the observations get older. The weights are determined by three smoothing parameters: alpha, beta, and gamma, which control the smoothing of the level, trend, and seasonality components, respectively.

The equation for the Multiplicative Holt-Winters Model is:

$$\hat{y}_{t+m|t} = (L_t + mT_t)S_{t-m+1+(m-1)mod m} \tag{3.10}$$

where $\hat{y}_{t+m|t}$ is the forecast for time $t + m$ given data up to time t, L_t is the level component at time t, T_t is the trend component at time t, and S_t is the seasonality component at time t. The notation $m - 1\ mod\ m$ means the remainder when $(m - 1)$ is divided by m.

The level component is updated as follows:

$$L_t = \alpha(y_t/S_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \tag{3.11}$$

where α is the smoothing parameter for the level component, and y_t is the observed value at time t.

The trend component is updated as follows:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \tag{3.12}$$

where β is the smoothing parameter for the trend component.

The seasonality component is updated as follows:

$$S_t = \gamma(y_t/(L_{t-1} + T_{t-1})) + (1 - \gamma)S_{t-m} \tag{3.13}$$

where γ is the smoothing parameter for the seasonality component.

The values of the smoothing parameters α, β, and γ are typically determined by minimizing the mean squared error between the predicted and observed values on a training set.

3.2.8.1 Concluding Remarks

The time-series models presented in this section offer interpretable and statistically robust tools for short- to medium-term forecasting of blood demand. Their structured assumptions and ease of implementation make them ideal as baseline models, particularly in contexts where data is limited or operational environments are well understood.

However, the performance of these models may degrade in the presence of non-linearity, irregular seasonality, or shifting trends—all of which are common in healthcare demand settings. The next section explores machine learning-based forecasting models, which are designed to overcome such limitations by learning patterns directly from data without requiring strong assumptions about the underlying process.

3.2.9 Machine Learning Models for Forecasting

In addition to classical time-series techniques, this study employs several supervised machine learning models for forecasting blood demand. These models are designed to automatically learn complex, nonlinear relationships in the data and adapt to dynamic patterns that may not be captured by traditional approaches.

The models selected for this study include Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), Support Vector Regression (SVR), and Linear Regression (LR). Each of these models is described below in terms of their architecture, forecasting capability, assumptions, and limitations.

As shown in Table 3.1, ANN and RNN models generally perform well across most evaluation criteria, including accuracy, tolerance to noise, and ability to handle interdependent attributes. These characteristics make them suitable for capturing the nonlinear and often noisy nature of healthcare-related time-series data such as blood demand.

Support Vector Regression (SVR), while strong in accuracy and classification speed, is more prone to overfitting and less tolerant to noise. Linear Regression (LR), although simple and computationally efficient, exhibits limitations in modeling complexity and robustness, making it better suited as a baseline benchmark.

The following subsections present a detailed overview of each selected model, including their internal architecture, training mechanism, and forecasting applications within the blood donation context.

	ANN	**RNN**	**LR**	**SVR**
Accuracy in general	strong	medium	weak	strong
Learning speed	medium	weak	strong	strong
Speed of classification	medium	strong	medium	strong
Tolerance to highly interdependent attributes	strong	strong	weak	weak
Tolerance to noise	strong	strong	medium	weak
Dealing with danger of overfitting	strong	strong	weak	high
Model parameter handling	strong	strong	weak	strong

Table 3.1 Comparing the performance of ML algorithms in time series forecasting

3.2.9.1 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) are inspired by the structure and function of biological neural systems and are widely used in predictive modeling tasks. In the context of blood demand forecasting, ANN models are particularly advantageous due to their ability to capture nonlinear relationships, tolerate noisy data, and learn complex temporal patterns from historical records.

Neural networks represent a function estimation technique that is driven by data and which, in the case of complex functions, can be considered as a reliable solution. Artificial neural networks, or ANNs, are networks of interconnected processing elements called neurons. They work by processing input signals and producing corresponding output signals. These neurons are put together into multiple layers that create a network structure. Each neuron takes input signals from the previous layer, does some computation using weights and bias. (Figure 3.1).

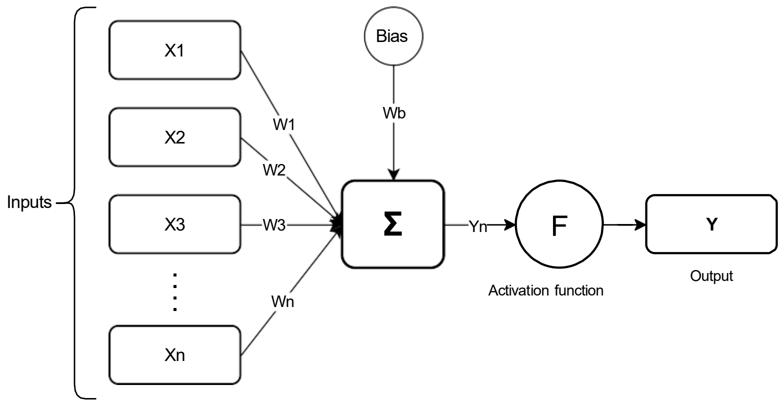


Fig. 3.1 ANN consists of interconnected neurons with assigned weights that aggregate to pass to an activation function.

The computation is based on a mathematical function that simulates the behavior of biological neurons. The output from the computation goes to the next layer of neurons until the final output is produced. Neural networks are non-hypothetical, meaning that they learn from the data rather than being hard-coded with a predefined set of rules. They can be trained using a variety of learning algorithms, such as backpropagation (Figure 3.2), to adjust the weights and biases of the neurons and enhance the accuracy of predictions. Neural networks have been applied in all manners of fields: image and speech recognition, natural language processing, time series prediction, and so on Kurani et al. (2023).

In creating an artificial neural network, the inputs have to be decided upon and then tried across different networks until the best is found. For our analysis, we used a common type of artificial neural network: layered neural networks. These are architectures with multiple layers of neurons, where each layer has a set number of neurons, named n.

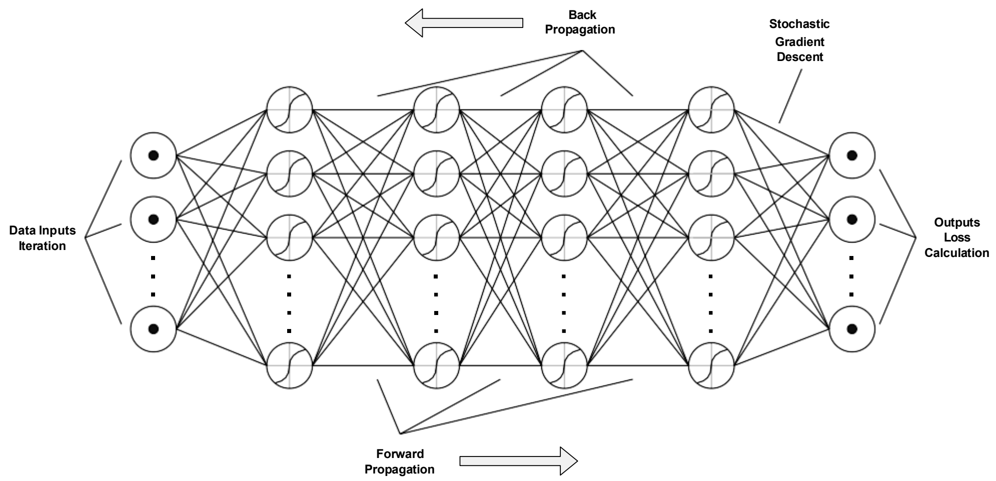


Fig. 3.2 Back Propagation in Neural Network

Furthermore, there is an activation function, f, which takes the input of each neuron to produce an output, a. It also has a bias term, b, a constant added to the input of all neurons, and weight vectors, w, which characterize the strength of the connections.

Remarks. Given their flexibility and learning capacity, ANN models are well-suited for forecasting blood demand in dynamic and uncertain environments. Their ability to generalize from historical data with complex, nonlinear characteristics makes them a strong candidate in comparison to linear statistical models. However, they require careful tuning and sufficient data to avoid overfitting and computational inefficiency.

3.2.9.2 Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNNs) are a class of neural networks particularly suited for modeling sequential and time-dependent data. Unlike feedforward networks, RNNs incorporate temporal feedback by maintaining internal memory, which allows them to capture dependencies across time steps. This makes them highly effective for time-series forecasting tasks such as blood demand prediction, where historical values influence future outcomes.

In an RNN architecture, the network comprises an input layer, a hidden (context) layer with recurrent connections, and an output layer. At each time step t, the input x_t and the previous hidden state h_{t-1} are used to compute the current hidden state h_t:

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{3.14}$$

$$y_t = W_{hy}h_t + b_y \tag{3.15}$$

where: - W_{xh}, W_{hh}, W_{hy} are the weight matrices, - b_h, b_y are the bias vectors, - f is a nonlinear activation function (e.g., tanh or sigmoid).

Figure 3.3 illustrates a typical RNN architecture.

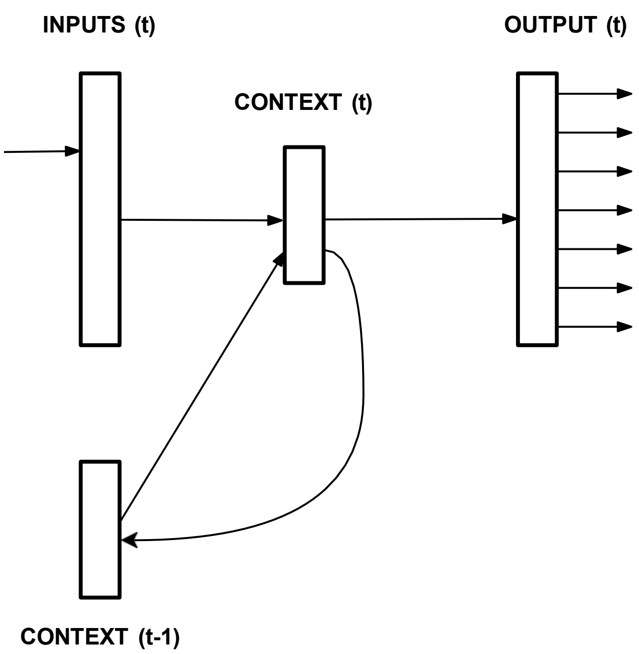


Fig. 3.3 Basic Recurrent Neural Network architecture

RNNs have been applied in a wide range of applications, including natural language processing, energy consumption prediction, stock price forecasting, and blood demand forecasting. Their strength lies in their ability to process input sequences of variable length and to model long-range temporal dependencies.

However, RNNs are not without limitations. They are prone to vanishing or exploding gradients, especially with long sequences. These issues can hinder learning and degrade forecast accuracy. Advanced variants such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) have been developed to address these problems, but their use was beyond the scope of this study.

Remarks. RNNs provide a powerful framework for modeling time-dependent behavior in complex datasets. For blood demand forecasting, they offer the ability to track trends, seasonality, and irregular patterns more effectively than traditional models. Nonetheless, their training complexity and sensitivity to initialization require careful tuning and appropriate data preprocessing.

3.2.9.3 Linear Regression (LR)

Linear Regression (LR) is one of the most fundamental models in statistical forecasting. It assumes a linear relationship between a dependent variable y and one or more independent variables $x_1, x_2, \ldots, x_k$. In its simplest form, a univariate linear regression model is expressed as:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \tag{3.16}$$

For multivariate regression:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \tag{3.17}$$

where: - y_i is the response variable for observation i, - x_{ij} are the explanatory variables, - β_0 is the intercept, - β_j are the regression coefficients, - ϵ_i is the error term, assumed to be normally distributed with zero mean.

In forecasting applications, LR fits a linear model to historical data and uses it to project future values. Its parameters are typically estimated using the Ordinary Least Squares (OLS) method.

Application to Time-Series Forecasting. While LR is computationally efficient and easy to interpret, its applicability to time-series forecasting is limited by several assumptions:

- **Linearity:** It presumes a constant linear relationship between predictors and the target variable over time.
- **Independence:** LR assumes that the residuals are uncorrelated, which is often violated in time-series data.
- **Stationarity:** In dynamic systems like blood demand forecasting, patterns often exhibit trends and seasonality—violating LR assumptions.

Despite these limitations, LR can serve as a useful benchmark model to evaluate the performance of more advanced forecasting methods.

Remarks. Linear Regression provides a simple, interpretable baseline for time-series forecasting. However, due to its limited ability to model nonlinear patterns, time dependencies, and adaptive behaviors, it is typically outperformed by models such as ANN and RNN in real-world BSC applications. Its main utility lies in comparison and baseline validation.

3.2.9.4 Support Vector Regression (SVR)

Support Vector Regression (SVR) is a supervised machine learning algorithm derived from Support Vector Machines (SVM), originally designed for classification tasks. SVR adapts the same principles for predicting continuous outputs by fitting a function that approximates the target variable within a margin of tolerance, defined by the parameter ε.

The regression function is typically linear in the feature space and has the form:

$$f(x) = w^T x + b \tag{3.18}$$

where w is the weight vector and b is a bias term. SVR aims to minimize model complexity and the deviation from the true outputs simultaneously.

The optimization objective is formulated as:

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C \sum_{i=1} (\xi_i + \xi_i^*) \tag{3.19}$$

Subject to:

$$\begin{aligned} y_i - (w^T x_i + b) &\leq \varepsilon + \xi_i \\ (w^T x_i + b) - y_i &\leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad \forall i &= 1, \ldots, n \end{aligned} \tag{3.20}$$

Here, ξ_i and ξ_i^* are slack variables allowing for deviations beyond ε, and C is a regularization parameter that controls the trade-off between flatness of the function and tolerance to errors.

SVR supports a variety of kernel functions—such as linear, polynomial, and radial basis function (RBF)—to model nonlinear relationships when needed. The kernel choice affects the model's flexibility and computational complexity.

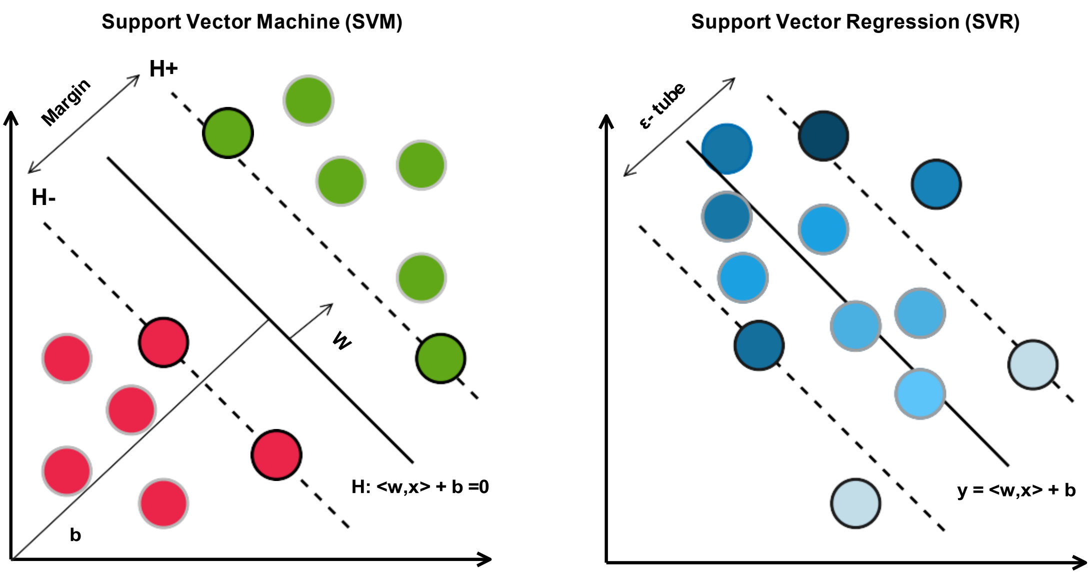


Fig. 3.4 Illustration of SVR vs. SVM

Application to Blood Demand Forecasting. SVR is a suitable choice for forecasting blood demand due to its ability to model complex relationships in moderate-sized datasets. However, several limitations should be considered:

- **Stationarity assumption:** SVR assumes a stable data distribution over time, which may not hold in healthcare contexts with shifting demand.
- **Hyperparameter sensitivity:** Its performance depends heavily on kernel type, regularization (C), and the ε margin.
- **Computational intensity:** Especially for non-linear kernels, training may become computationally expensive with large time series.
- **Sensitivity to noise and outliers:** SVR seeks to fit as many data points as possible within the margin, which can reduce robustness in noisy datasets.

Remarks. SVR offers a theoretically robust approach to regression and performs well in clean, moderately sized, and stable environments. However, its sensitivity to data anomalies and computational overhead makes it less suitable for high-frequency or volatile blood demand patterns. In this study, SVR is evaluated comparatively to ANN and RNN to assess its viability in real-world BSC forecasting.

3.2.10 Comparison of Forecasting Models

This section compares the performance of the forecasting models discussed in the previous sections, including both traditional time-series methods and machine learning models. The comparison is based on several key criteria relevant to blood demand forecasting, particularly in the context of the Algerian blood supply chain.

The models are evaluated across the following dimensions:

- **Accuracy:** The ability of the model to predict future demand with minimal error.
- **Noise Tolerance:** The model's robustness in the presence of noisy or incomplete data.
- **Computational Efficiency:** The time and resources required to train and implement the model.
- **Scalability:** The ability of the model to handle large datasets and adapt to growing data volumes.
- **Model Interpretability:** How easily the model's predictions and decision-making process can be understood and explained.

Table 3.2 summarizes the comparative evaluation of each model based on these criteria.

Table 3.2 Comparative Evaluation of Forecasting Models

Model	Accuracy	Noise Tolerance	Efficiency	Scalability	Interpretability
AR/ARMA	Medium	Low	High	High	High
ARIMA	Medium	Low	Medium	Medium	High
SARIMA	Medium-High	Medium	Medium	Medium	Medium
Holt-Winters	Medium	Medium	High	High	Low
SVR	High	Low	Low	Medium	Medium
LR	Low	Low	High	High	High
ANN	High	High	Medium	High	Low
RNN	High	High	Medium	Medium	Low

From this comparison, we can see that:

- **ANN and RNN** offer the highest accuracy and are highly adaptable to complex data patterns, but they come with increased computational demands and lower interpretability.
- **SVR** performs well in terms of accuracy and scalability, but it is less tolerant to noise and requires careful parameter tuning.
- **Traditional models** like ARIMA, SARIMA, and Holt-Winters offer simplicity and interpretability but may not perform as well on complex, nonlinear data.
- **Linear Regression (LR)**, while simple and efficient, is generally less effective on more complex datasets and nonlinear relationships.

Conclusion. Given the nonlinear nature of blood demand patterns and the challenges posed by noisy and incomplete data, machine learning models, particularly ANN and RNN, were selected for the proposed platform. These models are capable of learning complex relationships directly from the data, offering the best performance in terms of accuracy and adaptability for the Algerian blood supply system.

3.3 Blood Donor Classification

Blood donor classification plays a critical role in managing blood donations efficiently. By classifying donors into categories such as ”eligible”, ”ineligible”, or ”likely to donate”, healthcare providers can optimize the use of available resources, target communication efforts, and ensure the right blood products are available when needed. This section explores the use of machine learning models to predict donor behavior and classify them based on features such as donation history, demographics, and health status.

In this study, we employ several machine learning algorithms, including Decision Trees (DT), Random Forest (RF), Na¨ıve Bayes (NB), and K-Nearest Neighbors (KNN), to classify blood donors. These models are selected based on their ability to handle various feature types, interpretability, and scalability to large datasets. The goal is to identify the most effective classifier for predicting donor return behavior in the Algerian context.

3.3.1 Problem Description and Features

Blood donor classification involves predicting whether a donor will return for future donations based on historical data and personal characteristics. Key features used in this classification include:

- **Demographic features:** Age, gender, and location.
- **Donation history:** Frequency, last donation date, and type of donation (e.g., whole blood, plasma).
- **Eligibility criteria:** Health screening results, weight, hemoglobin levels, etc.

These features are essential in understanding donor behavior and predicting future donations. However, they are often subject to missing data, imbalances in class distribution, and non-linear relationships. Therefore, machine learning models are particularly well-suited for this task, as they can learn complex patterns from data without the need for explicit rules or assumptions.

3.3.1.1 Decision Trees (DT)

Decision Trees (DT) are a supervised learning algorithm widely used for classification and regression tasks. The algorithm works by recursively splitting the data based on the feature that provides the most information gain or minimizes impurity at each decision node. These splits create a tree-like structure where each internal node represents a decision based on an attribute, each branch represents the outcome of that decision, and each leaf node represents a predicted class label or numerical value.

The primary goal of Decision Trees is to create subsets of data that are as homogeneous as possible in terms of the target variable. This splitting criterion is commonly measured using metrics like **Information Gain** (for classification tasks) or **Gini Impurity**.

Mathematically, a Decision Tree can be expressed as a function $T(x)$ that maps input features x to a predicted class label or numerical value:

$$T(x) : \mathsf{R}^d \rightarrow \mathsf{Y} \tag{3.21}$$

Where Y is either the set of class labels (for classification) or real values (for regression). The tree is constructed by recursively splitting the feature space, choosing the best feature at each step to minimize impurity.

Steps in Decision Tree Induction:

1. Start with the entire dataset at the root node.
2. For each node, select the best feature to split the data based on metrics such as Information Gain (for classification) or Gini Impurity.
3. Split the data into sub-nodes based on the chosen feature.
4. Repeat the process recursively for each sub-node until a stopping criterion is met (e.g., maximum depth or minimum number of samples per leaf).

The performance of Decision Trees depends heavily on the quality of the input features, and they are prone to **overfitting** if the tree becomes too deep or complex.

Application to Blood Donor Classification. In this study, Decision Trees are applied to classify blood donors into eligible and ineligible categories based on features such as donation history, demographics, and eligibility criteria. One of the key advantages of Decision Trees is their ability to handle both **numerical** and **categorical** data, making them particularly useful for datasets with mixed types of attributes.

Decision Trees are also easy to interpret, and the decision-making process can be visualized, which makes them an attractive choice for explaining model predictions. However, their performance can degrade when there are too many features or when the tree becomes too deep and overfits the training data.

Remarks. While Decision Trees are intuitive and effective in many classification tasks, they have limitations. They tend to overfit the training data, especially with complex models, and their performance can be sensitive to small changes in the dataset. To overcome these limitations, ensemble methods such as Random Forests (RF) are often employed, where multiple Decision Trees are trained and their predictions are aggregated to improve accuracy and generalization.

3.3.1.2 Na¨ıve Bayes classifier (NB)

The Na¨ıve Bayes classifier is a probabilistic model based on Bayes' theorem, with the simplifying assumption that the features in the dataset are conditionally independent given the class label. Despite this ”naive” assumption, Na¨ıve Bayes classifiers are widely used for classification tasks due to their simplicity, efficiency, and effectiveness, especially in high-dimensional data.

The Naïve Bayes classifier computes the posterior probability of a class y given input features $X = \{X_1, X_2, \ldots, X_n\}$ using Bayes' theorem:

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)} \tag{3.22}$$

Since $P(X)$ is constant for all classes, it is ignored in practice, simplifying the decision rule to:

$$\hat{y} = \arg\max_y P(y) \prod_{i=1} P(X_i|y) \tag{3.23}$$

Where: - $P(y)$ is the prior probability of class y, - $P(X_i|y)$ is the likelihood of feature X_i given the class y.

For continuous features, the likelihood $P(X_i|y)$ is often modeled using a Gaussian distribution:

$$P(X_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(X_i - \mu_y)^2}{2\sigma_y^2}\right) \tag{3.24}$$

where μ_y and σ_y^2 are the mean and variance of feature X_i for class y.

Application to Blood Donor Classification. In the context of blood donor classification, the Naïve Bayes classifier is used to predict the likelihood of a donor returning for future donations based on their demographic features, donation history, and eligibility criteria. Given its simplicity and efficiency, Naïve Bayes is often used as a baseline classifier, particularly in large datasets with many features.

The classifier performs particularly well when features are conditionally independent. However, in real-world blood donation data, dependencies between features (e.g., age and donation frequency) may lead to suboptimal performance. Despite this, Naïve Bayes remains a strong choice for many classification tasks, especially when data is noisy or incomplete.

Remarks. The Naïve Bayes classifier is a fast, efficient, and simple algorithm that can provide surprisingly good performance, particularly with large datasets or when data is sparse. However, the assumption of feature independence may limit its accuracy in situations where feature interactions are important. In this study, Naïve Bayes is used as a benchmark for evaluating more complex models such as Decision Trees and Random Forests.

3.3.1.3 Random Forest (RF)

Random Forest (RF) is an ensemble learning method that builds multiple decision trees and merges their predictions to improve accuracy and robustness. Unlike single decision trees, which can suffer from high variance and overfitting, Random Forest mitigates these issues by training each tree on a random subset of the training data and features.

The algorithm operates as follows:

1. Randomly select n samples from the training set with replacement (bootstrapping).
2. For each sample, randomly select m features.
3. Construct a decision tree based on the selected subset of features and samples.
4. Repeat steps 1–3 k times to build k decision trees.
5. To make predictions, aggregate the results from all k trees: for classification, use majority voting; for regression, average the predictions.

Mathematically, the Random Forest algorithm can be summarized as:

$$\hat{y} = \frac{1}{k} \sum_{i=1}^{k} T_i(x) \tag{3.25}$$

where $\hat{y}$ is the predicted output, $T_i(x)$ is the prediction from the i-th tree, and k is the total number of trees in the forest.

Application to Blood Donor Classification. Random Forests are particularly effective for blood donor classification due to their ability to handle large datasets with many features, their resistance to overfitting, and their capability to handle missing data. The model works by aggregating the results of many decision trees, each trained on a random subset of the data and features, which enhances the model's stability and accuracy.

In this study, Random Forest is used to classify blood donors based on various features, including demographic information, donation history, and health eligibility. The model's ability to handle both numerical and categorical features, and its capacity to provide feature importance scores, makes it a robust choice for this application.

Remarks. Random Forest provides high accuracy and is highly resilient to overfitting. It can also handle missing data and is scalable to large datasets, making it ideal for blood donor classification where missing or incomplete data is common. However, it sacrifices interpretability compared to individual decision trees. Despite this, its performance and robustness make it a strong contender for deployment in real-world blood supply chain management applications.

3.3.1.4 K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a simple yet effective supervised machine learning algorithm used for classification and regression tasks. The basic idea behind KNN is that a data point's classification (or predicted value) is determined by the majority class (or average value) of its K nearest neighbors in the feature space.

KNN works as follows:

1. For a given test point, calculate the distance between the test point and all the points in the training dataset (using metrics such as Euclidean distance).
2. Sort the distances in ascending order.
3. Select the top K closest points and assign the most frequent class label (for classification) or average of the target values (for regression) as the prediction.

The model is mathematically represented by the following:

$$\hat{y} = \frac{1}{K} \sum_{i=1}^{K} y_i$$

for regression or for classification:

$$\hat{y} = \operatorname{argmax}_y \sum_{i=1}^{K} \mathbf{1}(y_i = y)$$

where - y_i is the target value (or class label) for the i-th nearest neighbor, - K is the number of neighbors considered.

Application to Blood Donor Classification. In blood donor classification, KNN can be used to predict whether a donor will return for future donations based on features like age, donation history, and health status. The advantage of KNN lies in its simplicity and flexibility—it can handle both numerical and categorical data and does not require explicit training. However, KNN can be computationally expensive when dealing with large datasets, as it requires calculating distances to all training points.

KNN also tends to perform well when the data points are well-separated and the decision boundaries are relatively simple. However, it struggles with high-dimensional data and can suffer from the **curse of dimensionality** when the feature space becomes too large.

Remarks. KNN is a powerful classifier that is easy to implement and understand. It is well-suited for small-to-medium-sized datasets with clear, well-defined decision boundaries. However, its computational inefficiency, particularly with large datasets, and sensitivity to irrelevant features, make it less suitable for complex, high-dimensional blood donation datasets. In this study, KNN is employed as a simple baseline model to compare against more complex algorithms such as Random Forests and Neural Networks.

3.3.1.5 Logistic Regression (LR)

Logistic Regression (LR) is a statistical method used for binary classification tasks, where the goal is to predict the probability that a given input belongs to one of two possible categories. Unlike linear regression, which predicts a continuous outcome, LR outputs probabilities through a logistic function, making it suitable for classification tasks.

The model is expressed as:

$$P(y = 1|X) = \frac{1}{1 + \exp -(\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n)} \tag{3.26}$$

Where: - $P(y = 1|X)$ is the probability that the outcome variable y equals 1 given input features X, - β_0 is the intercept term, - $\beta_1, \ldots, \beta_n$ are the regression coefficients for the features $x_1, \ldots, x_n$.

The model is trained by finding the values of $\beta_0, \beta_1, \ldots, \beta_n$ that maximize the likelihood of observing the given training data. This is typically done using methods like **Maximum Likelihood Estimation (MLE)**.

Application to Blood Donor Classification. Logistic Regression is widely used for binary classification tasks and is particularly effective in situations where the relationship between the input features and the target variable is expected to be linear. In the context of blood donor classification, LR can predict whether a donor will return for future donations based on their previous donation history, demographic information, and eligibility criteria.

However, while LR is simple and interpretable, it has some limitations:

- **Linearity assumption:** LR assumes a linear relationship between the input features and the log-odds of the target variable, which may not hold in more complex datasets.
- **Sensitivity to outliers:** The model's performance can degrade if there are outliers in the data.
- **Limited ability to handle non-linear relationships:** If the relationships between features and outcomes are not linear, LR may not provide the best predictive performance compared to more complex models like Random Forests or Neural Networks.

Remarks. Despite its limitations, Logistic Regression is a computationally efficient and interpretable model, making it an excellent choice for baseline models in classification tasks. It is particularly useful when the problem involves understanding the impact of individual features on the outcome, such as predicting donor behavior based on demographic or historical factors.

3.3.2 Comparison of Classification Models

In this section, we compare the performance of six machine learning models used for blood donor classification: Artificial Neural Networks (ANN), Logistic Regression (LR), Support Vector Machines (SVM), Decision Trees (DT), Random Forest (RF), and Na¨ıve Bayes (NB).

The comparison is based on several key criteria, including:

- **Accuracy:** How well the model predicts the target class (eligible vs. non-eligible donor).
- **Noise Sensitivity:** The model's robustness to noisy or incomplete data.
- **Computational Efficiency:** The time and resources required to train and apply the model.
- **Interpretability:** The ease with which the model's predictions can be understood and explained.
- **Scalability:** The ability of the model to handle large datasets and adapt to growing data volumes.

Table 3.3 summarizes the comparative performance of each model based on these criteria.

Table 3.3 Comparison of Classification Models

Model	Accuracy	Noise Sensitivity	Efficiency	Interpretability	Scalability
ANN	High	Medium	Medium	Low	High
LR	Low	Low	High	High	High
SVM	High	Low	Medium	Medium	High
DT	Medium	High	Medium	High	High
RF	High	Medium	Medium	Low	High
NB	Medium	High	High	High	Medium

From this table, we can draw the following conclusions:

- **ANN** provides the highest accuracy and is well-suited for handling complex, non-linear relationships in the data. However, it requires significant computational resources and is difficult to interpret.
- **Logistic Regression (LR)**, while fast and simple, struggles with more complex datasets and tends to underperform in comparison to other models, especially with non-linear data.
- **Support Vector Machines (SVM)** offer high accuracy and are effective in capturing complex boundaries, but they are sensitive to noise and require careful tuning of hyperparameters.
- **Decision Trees (DT)** are easy to interpret and handle both categorical and numerical data. However, they tend to overfit and may not perform well on large or noisy datasets.
- **Random Forest (RF)** provides high accuracy, scalability, and resilience to overfitting. However, it is computationally expensive and less interpretable than DT.
- **Na¨ıve Bayes (NB)** is efficient and interpretable but assumes independence between features, which can limit its accuracy if features are correlated.

Based on the comparison, **Random Forest (RF)** and **Artificial Neural Networks (ANN)** stand out as the most powerful models for blood donor classification. These models offer the best balance of accuracy, noise tolerance, and scalability, making them suitable for real-world applications in the Algerian blood supply system.

However, due to their complexity, **ANN** requires more computational resources and is harder to interpret, making **Random Forest (RF)** a more practical choice for large-scale deployment. **Logistic Regression (LR)** and **Na¨ıve Bayes (NB)** serve as baseline models for comparison, providing insights into simpler, more interpretable approaches but with limitations in performance on complex datasets.

Thus, for the purposes of this research, **Random Forest (RF)** is selected as the primary model for implementation, with **ANN** considered for tasks where high accuracy is paramount and computational resources are available.

3.4 Summary and Justification of Model Selection

This chapter has explored a range of forecasting and classification models used for blood donation management. In particular, we examined traditional time-series models (ARIMA, SARIMA, Holt-Winters) and machine learning models (Artificial Neural Networks, Logistic Regression, Support Vector Regression, Decision Trees, Random Forest, and Na¨ıve Bayes).

Each of these models was evaluated across several criteria, including predictive accuracy, noise tolerance, computational efficiency, interpretability, and scalability. From this evaluation, it became clear that while traditional time-series models offer simplicity and interpretability, they often fail to capture complex, nonlinear relationships in the data. On the other hand, machine learning models such as **Random Forest (RF)** and **Artificial Neural Networks (ANN)** demonstrated superior performance in terms of accuracy, adaptability, and robustness, particularly in capturing the nonlinear patterns inherent in blood donation data.

- Random Forest (RF): This model was selected as the primary model for its robustness, high accuracy, and ability to handle large datasets with minimal overfitting. Its ability to process both categorical and numerical features, combined with its resistance to noise and outliers, makes it an ideal candidate for blood donor classification and blood demand forecasting in the Algerian context.
- **Artificial Neural Networks (ANN)**: Although ANN is computationally more intensive and harder to interpret, its capacity to model complex, non-linear relationships made it a strong contender for improving the accuracy of blood demand predictions. ANN is particularly useful when large amounts of data are available, and its ability to generalize makes it well-suited for dynamic healthcare environments.

- **Logistic Regression (LR)**: Despite its limitations in handling complex, non-linear data, LR was used as a baseline model due to its simplicity, speed, and interpretability. It serves as an effective starting point for classification tasks, particularly when the relationships between features and outcomes are expected to be linear.
- **Support Vector Regression (SVR)**: SVR was selected for its ability to model both linear and non-linear data. While it is sensitive to hyperparameters and noise, it performed well in cases where the relationships between features and the target variable were not easily captured by traditional models.

Given these strengths and weaknesses, **Random Forest (RF)** was ultimately selected as the primary model for implementation. The model's flexibility, coupled with its strong predictive performance and scalability, makes it well-suited for deployment in real-world blood donation management systems. **ANN** is considered a complementary model, especially for scenarios requiring higher accuracy, where computational resources are available.

By combining these models into a hybrid system, we aim to leverage the strengths of each model to improve the accuracy and efficiency of blood supply chain management.

3.5 Conclusion

In this chapter, we explored the application of Artificial Intelligence (AI) and Machine Learning (ML) techniques in blood donation management, with a particular focus on blood demand forecasting and donor classification. After reviewing both traditional time-series models and machine learning algorithms, we selected Random Forest (RF) and Artificial Neural Networks (ANN) as the primary models due to their robustness, accuracy, and adaptability to complex, nonlinear patterns in the data. Logistic Regression (LR) and Support Vector Regression (SVR) were also evaluated as baseline models for comparison.

These models were chosen not only for their predictive accuracy but also for their scalability and suitability for deployment in real-world healthcare environments, particularly in the context of the Algerian blood supply system. The following chapters will detail the i**mplementation of these models into an intelligent decision support platform** to address operational challenges in the blood supply chain, including real-time forecasting, donor classification, and scheduling.

The next chapter will delve into the **Research Methodology** where we will outline the approach, data collection methods, model implementation strategies, and evaluation techniques used to develop and test the proposed intelligent platform.

CHAPTER 4

Research Methodology and System Design

In order to develop a robust and intelligent system for managing the blood supply chain in Algeria, a comprehensive methodological framework was required—one that combines both structured project design and data-driven decision-making processes. This chapter presents the methodology followed throughout the research, justifying the selected approaches and detailing each stage from problem formulation to the deployment of AI/ML-based tools. The methodological design integrates macro- and micro-level perspectives, combining theoretical analysis, practical constraints, and real-world data experimentation.

The chapter is structured as follows: it begins by comparing several established macro-level research methodologies, leading to the selection of the CRISP-DM approach as the most appropriate for the study's objectives. Next, the micro-methodological strategy is introduced, emphasizing the role of data-driven techniques. Finally, the chapter details the actual implementation steps taken, including data collection, preprocessing, model building, and evaluation processes that enabled the construction of the intelligent decision support platform.

4.1 Introduction

The success of any data-driven system depends not only on the strength of the algorithms used but also on the soundness of the methodology that guides its development. In this research, the objective is to optimize the blood supply chain in Algeria using artificial intelligence and machine learning techniques. Achieving this goal requires a rigorous methodological framework that ensures clarity, traceability, and adaptability throughout the entire system design process.

To address these needs, the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology was selected as the foundation of this research. It provides a structured yet flexible approach suitable for iterative development, especially when working with real-world healthcare data that is often noisy, incomplete, and subject to change.

In addition to the macro-level methodology, this study also adopts complementary micro-model strategies to guide the development of specific analytical components, particularly those driven by machine learning.

The overall methodology followed in this research is structured around two major components:

1. **Data lifecycle:** This includes data collection, validation, processing, and statistical evaluation.
2. **AI/ML-based decision support tool:** This comprises forecasting blood demand, classifying donors, and optimizing donor scheduling.

The rest of this chapter details the methodological foundations, system architecture, data preparation processes, and the design and deployment of intelligent decision-support modules.

4.2 Methodological Foundations

The selection of an appropriate methodological framework is a crucial step in ensuring the rigor, relevance, and success of applied research—especially when dealing with complex, data-intensive systems such as those found in healthcare logistics. This section explores the foundational methodologies considered for this study, both at the macro level (project-wide structure) and micro level (model-specific design logic).

It begins by comparing several well-known macro-level research methodologies, assessing their suitability for guiding the design and development of a smart blood supply chain system. This is followed by a justification for adopting CRISP-DM as the principal framework. Finally, the section introduces complementary micro-model methodologies that inform the internal logic of specific AI/ML modules developed as part of this research.

4.2.1 Macro-model Methodologies

To identify the most suitable approach for guiding this research, a review of prominent macro-level methodologies was conducted. These include:

- **Scientific Research Methodology (SR)** – A theory-driven approach based on hypothesis formulation, data collection, experimentation, and validation. While suitable for basic research, it lacks the iterative structure and operational focus required in applied AI systems.

- **Operations Research (OR)** – A mathematically rigorous methodology centered on optimization models. Despite its relevance to logistics, it assumes static parameters and deterministic models, making it less flexible for dynamic, data-driven environments like blood supply chains.
- **Waterfall Model** – A linear and sequential software development model. It is appropriate for well-defined systems with fixed requirements, but it does not support iterative refinement or ongoing data learning, which are essential in this study.
- **CRISP-DM (Cross-Industry Standard Process for Data Mining)** – A flexible and widely used data mining methodology that supports iterative development, continuous feedback, and close integration between business goals and technical modeling.

Figure 4.1 illustrates the structural differences among these approaches.

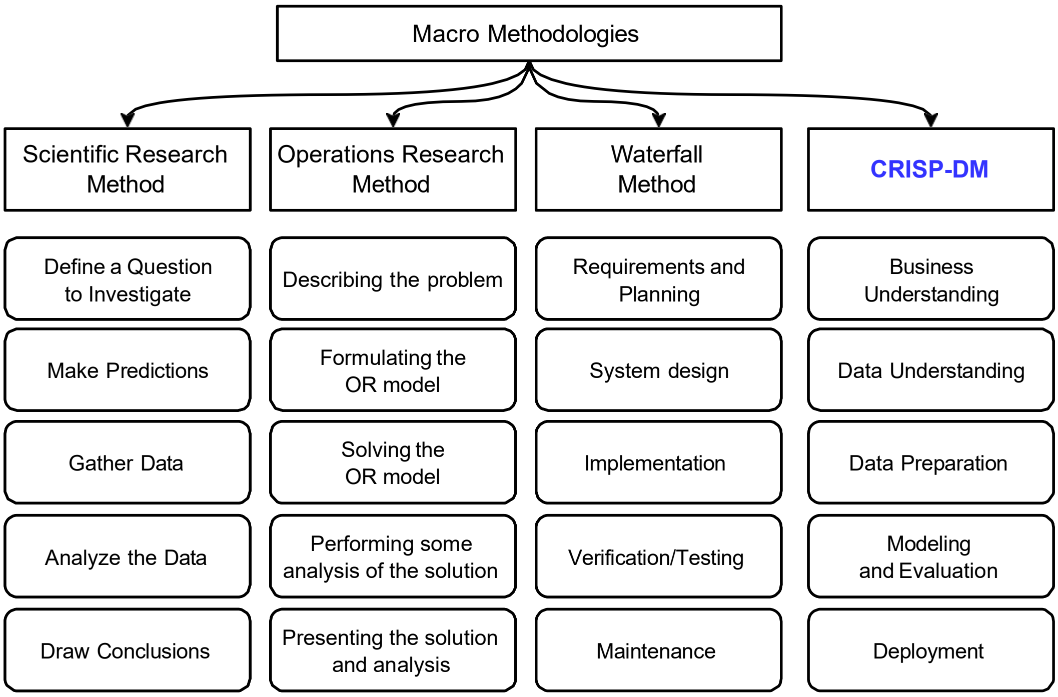


Fig. 4.1 Comparison of Macro-level Research Methodologies

This comparison highlights the strengths and limitations of each methodology and provides the rationale for adopting CRISP-DM in the context of this research.

4.2.2 Justification for Using CRISP-DM

After evaluating several established research methodologies, the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework was selected as the most suitable for this research. This choice was based on both the nature of the problem and the characteristics of the data involved.

Unlike traditional linear approaches such as the Waterfall model or rigid mathematical frameworks like Operations Research, CRISP-DM provides a flexible, iterative process that supports continuous refinement—an essential feature when working with dynamic, real-world healthcare data.

CRISP-DM is particularly well-suited to this project for the following reasons:

- **Iterative Development:** The cyclic structure allows revisiting earlier stages (e.g., data understanding, modeling) in light of new findings.
- **Business Alignment:** The methodology emphasizes clear definition of business goals and ensures that modeling efforts directly support them.
- **Data-Centric Focus:** It integrates data understanding and preparation as core phases, which aligns with the challenges of processing healthcare and behavioral data.
- **Modular Structure:** The process can be adapted to accommodate multiple modeling components, such as forecasting, classification, and optimization.

These advantages made CRISP-DM the logical foundation for structuring the development of the intelligent platform proposed in this study.

4.2.3 Micro-model Methodologies

While CRISP-DM provides a solid macro-level structure, the development of individual analytical modules—such as forecasting, classification, and scheduling—requires more specific strategies. To this end, the study incorporates micro-model methodologies that guide how data-driven solutions are designed and refined.

Three general categories of micro-methodological approaches are considered in the context of this research:

- **Exploration and Discovery:** Used when the problem is not well-defined and requires iterative investigation, prototyping, and pattern recognition. This strategy is typical in early-stage data exploration.
- **Solutions Independent of Data:** Based on domain knowledge, heuristics, or expert rules. These approaches do not rely heavily on data but are driven by predefined logic or rules. They are more suited for deterministic environments.
- **Solutions Dependent on Data:** These are fully data-driven approaches that extract insights and patterns from historical datasets to inform predictions and decisions. Machine learning models and statistical forecasting techniques fall under this category.

Given the availability of large-scale donation data and the goal of developing predictive and adaptive tools, this study primarily adopts the ”Solutions Dependent on Data” strategy to guide the design of the platform’s AI components.

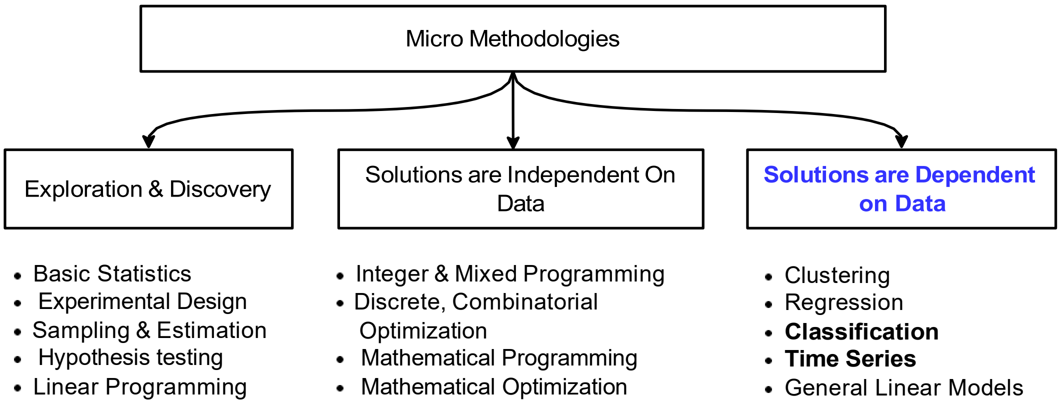


Fig. 4.2 Micro-model Methodologies

4.3 System Architecture

To operationalize the CRISP-DM methodology and integrate both data processing and AI-driven components, a layered system architecture was developed. This architecture enables the end-to-end flow from raw data acquisition to intelligent decision-making and visualization for healthcare staff.

The system is structured across five core layers:

- **Data Collection Layer:** Interfaces with blood donation centers, mobile collection units, and hospital systems to capture data related to donors, donations, demand, and blood inventory levels.
- **Validation and Storage Layer:** Performs automated data cleaning, consistency checks, and formatting, before securely storing the validated records in a centralized database.
- **Decision Support Layer:** Hosts the machine learning modules for demand forecasting, donor classification, and donor scheduling optimization. These models operate on real-time and historical data.
- **Application Layer:** Provides dashboards, notifications, and visual analytics to support healthcare staff in daily operations and strategic planning.
- **Monitoring and Feedback Layer:** Tracks user activity, system outputs, and key performance indicators to allow continuous improvement of the models and workflows.

The interaction among these layers enables real-time, adaptive decision-making aligned with operational constraints in blood collection environments. Figure 4.5 illustrates the overall architecture of the platform and the interaction between its components.

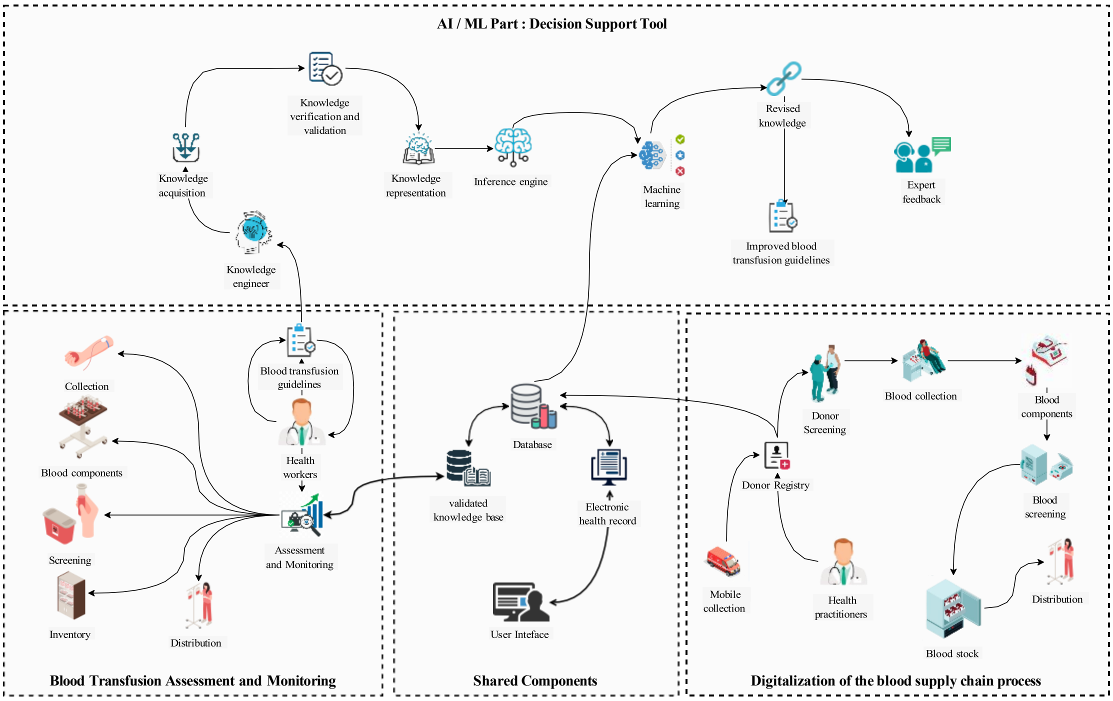


Fig. 4.3 System Architecture

4.4 Research Methodology Design

The practical application of the CRISP-DM methodology in this research is structured around two interconnected components: (1) a comprehensive data lifecycle, and (2) an AI/ML-based decision support tool. This dual structure ensures that all system components are developed based on validated data and continuously refined through feedback loops.

1. **Data Lifecycle:** This component includes the acquisition, processing, validation, and statistical evaluation of operational and behavioral data collected from national blood services and partner hospitals. It forms the foundation upon which all analytical models are built.

2. **AI/ML-based Decision Support Tool:** This component integrates predictive analytics into the blood supply chain workflow through three key modules:

    - **Blood Demand Forecasting Model:** Predicts short- and medium-term blood needs using time series analysis and machine learning.
    - **Donor Classification Model:** Identifies return donors and estimates donation likelihood based on behavioral and demographic patterns.
    - **Donor Scheduling Optimization Module:** Matches donors to collection schedules using AI-based prioritization and availability modeling.

This design allows the system to maintain high adaptability, accuracy, and operational relevance, especially in a healthcare environment characterized by uncertainty and resource constraints.

Figure 4.4 shows the architecture of the research design and how data and AI modules are integrated into the overall platform.

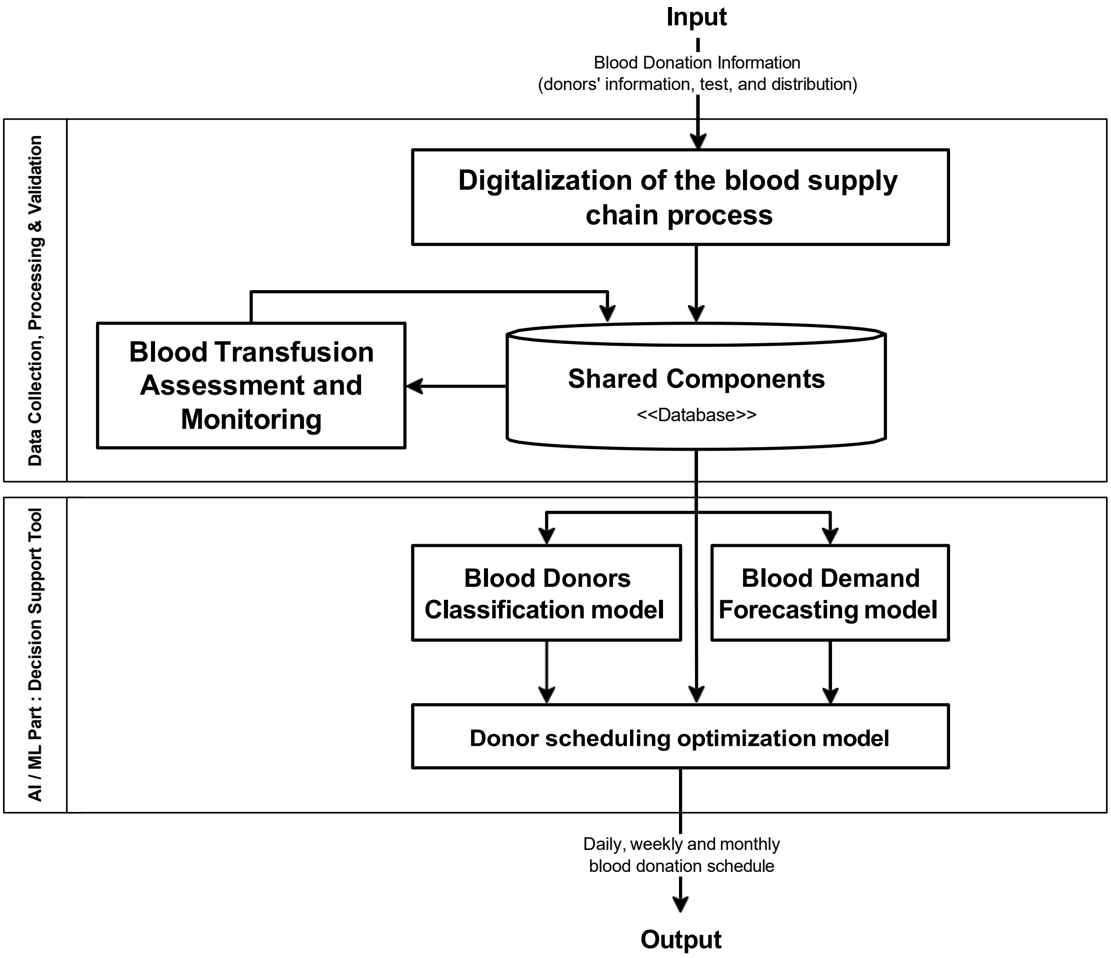


Fig. 4.4 Design Overview of Proposed Solution

4.5 Data Collection, Processing and Validation

Data quality is a fundamental factor that impacts the accuracy of predictions in blood component requests. In this study, the accuracy of the forecasting models depends on both the quality of the model and the data used for training. While the former was thoroughly evaluated in previous sections, this section focuses on the challenges and solutions encountered in ensuring the quality of the data.

To overcome the challenges posed by inconsistencies, missing values, and non-standard formats, a digital platform was proposed to automate the data collection process and digitize all stages of the blood supply chain. This platform, shown in The interaction among these layers enables real-time, adaptive decision-making aligned with operational constraints in blood collection environments. 4.5, aims to streamline data acquisition, making it more efficient, reliable, and readily available for processing and modeling.

The primary objectives of the platform are:

- To automate the collection of data across the entire blood supply chain.
- To ensure timely and accurate data entry for use in the AI/ML decision support models.
- To improve the overall efficiency and accuracy of blood donation and collection operations.

The platform's main role is to ensure that the data fed into the predictive models is of high quality, which directly influences the accuracy of the predictions for blood component requirements.

This system aims to transform the blood supply chain process by introducing digital tools that enhance both operational effectiveness and model accuracy. With this platform, healthcare organizations can improve decision-making processes and better plan for future blood supply needs.

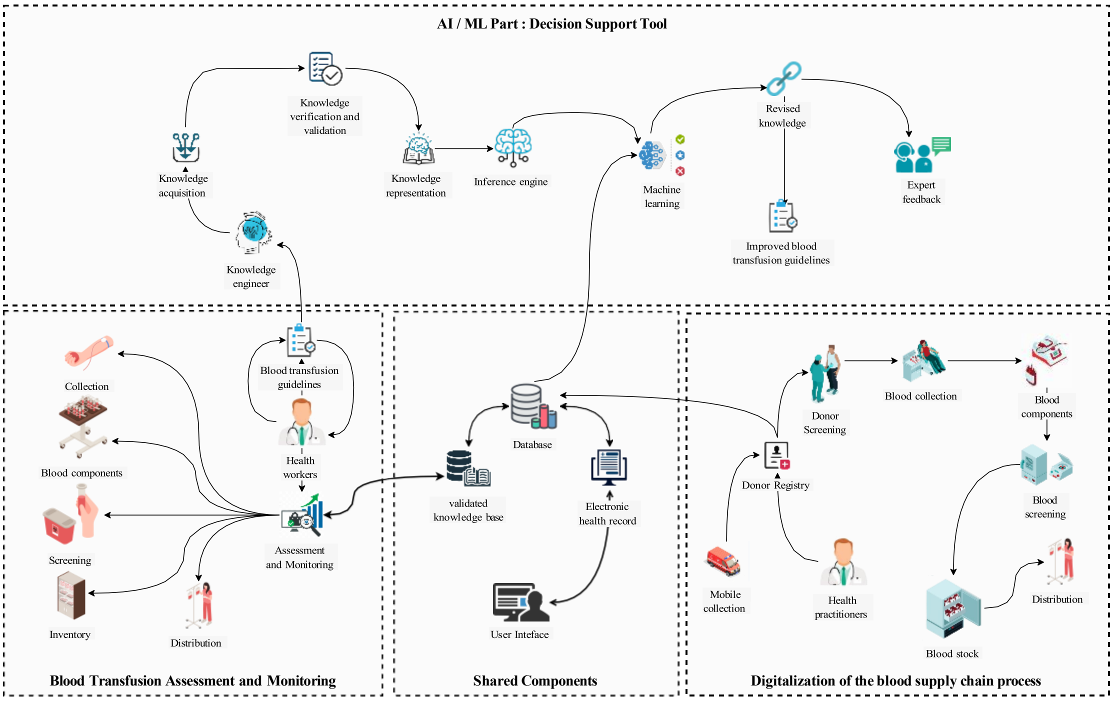


Fig. 4.5 System Architecture

4.5.0.1 Exploratory Analysis and Statistical Evaluation

To ensure that the collected data was suitable for building reliable predictive models, several exploratory and statistical analyses were performed. These analyses focused on assessing the stationarity, normality, and autocorrelation characteristics of the blood demand time series. The findings directly informed the choice of forecasting models and preprocessing strategies.

For a deeper understanding of the data, we analyze the data using standard economic analysis techniques to evaluate its normal state. We perform a static test and evaluate one of the Box-Jenkins methodology hypotheses for data normality. Table 4.1 summarizes several common economic tests and procedures that were utilized in this study. We use a table description because it provides a clear description of each test and is easy to understand.

Table 4.1 Econometric tests and procedures.

Tests	Description	Purpose of Its Use
AIC, BIC, LR	The Akaike Information Criterion is abbreviated as AIC, whereas the Bayesian Information Criterion is abbreviated BIC. The AIC/BIC is explicitly designed for model selection and is used to select the best-fit model. We choose the model with the highest AIC/BIC function and/or lowest error. STATA computes the AIC and BIC results using the log Likelihood Ratio (LR), described by Hamilton.	This function determines the best order of p and q for a given ARMA model.
Invertibility	The stationary test focuses on the data's autoregressive representation, whereas the invertibility test focuses on the data's moving average representation. This test determines if a process can be described as a function of previous lag values plus an error term.	This function determines the stability percentage of the moving average (MA stationary).
ADF	The t-test is used to calculate the mean and variance. The null hypothesis confirms that there is no stability. Rejection of the null leads to the conclusion that the data is not fixed.	We can forecast the dependent variable in a significant period when a process is steady.
Jarque–Bera	This is used since maximum likelihood and Chi-square assess whether the provided probability distribution fits a normal distribution. As a result, Kurtosis (the measure of "peakedness") and Skewness (the measure of asymmetry) are assessed for this test.	This test determines whether a process's probability distribution is similar to the normal distribution.

Stationarity Assessment: The Augmented Dickey–Fuller (ADF) test was conducted to determine whether the total blood demand time series exhibited stationarity. The results confirmed the stationarity of the data (see Table 4.2):

Table 4.2 Augmented Dickey–Fuller test for Total Blood Demand

	T-Statistics	*p*-Value	Critical Value (1%)	Critical Value (5%)
TBD	−8.99	6.91×10^{-15}	−3.43	−2.86

Normality Check: To evaluate whether the blood demand data followed a normal distribution, the Jarque–Bera test was applied (see Table 4.3):

Table 4.3 Jarque–Bera test for normality of data.

	Jarque Bera Test Statistics	*p*-Value	Estimated (Skewness)	Estimated (Kurtosis)
TBD	41.54	9.51	0.38	3.32

The results indicated that the data is approximately normally distributed and suitable for standard time series forecasting techniques.

Autocorrelation Analysis: Autocorrelation and partial autocorrelation functions (ACF and PACF) were used to assess temporal dependencies within the time series. Figure 4.6 shows the ACF and PACF plots for total blood demand. The plots revealed significant lag structures, confirming that past demand values could be used effectively in modeling future demand.

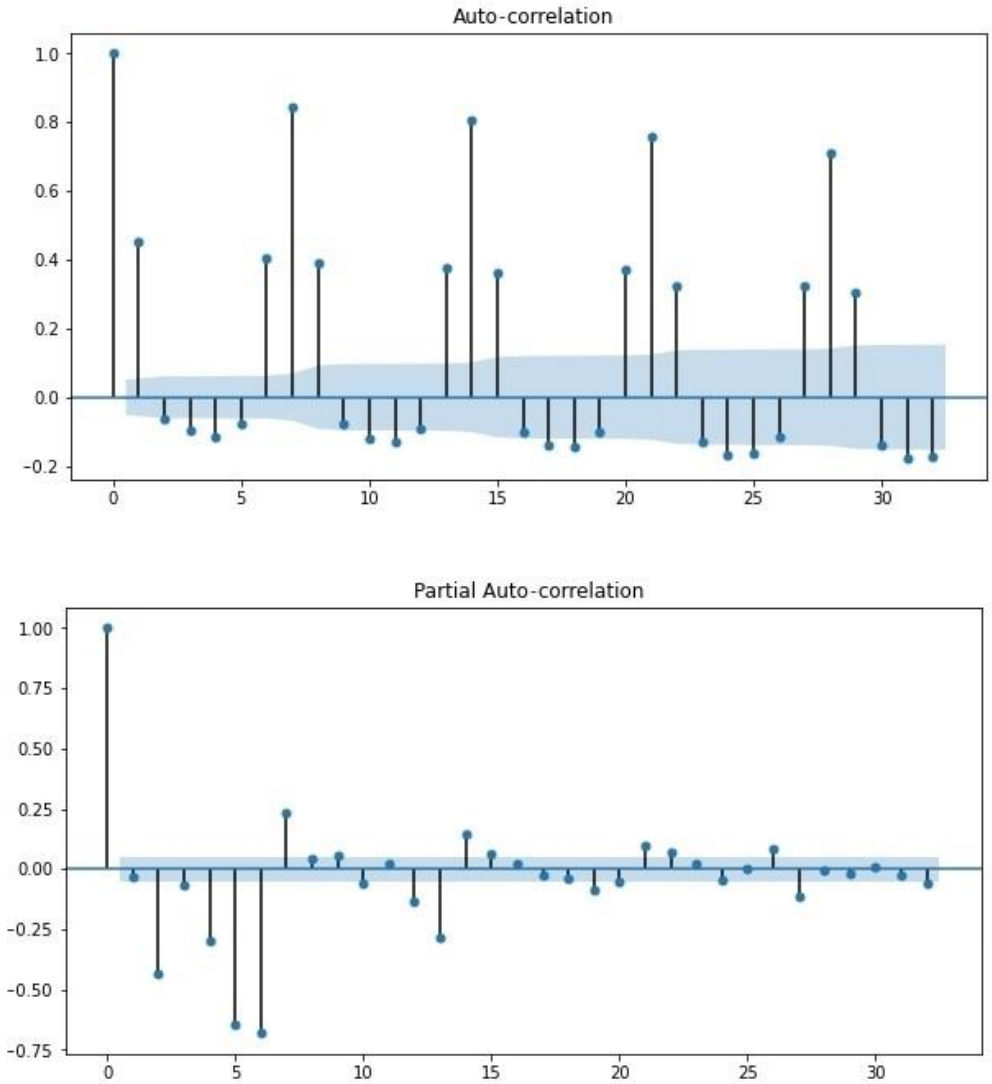


Fig. 4.6 Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for total blood demand data

4.5.0.2 Summary of Exploratory Analysis and Statistical Evaluation

The exploratory analysis and statistical evaluation of the blood demand data provided essential insights into the underlying characteristics of the dataset. The stationarity test confirmed the stability of the data, which is crucial for time series forecasting. The normality test indicated that the blood demand data follows a distribution close to normal, ensuring the appropriateness of standard forecasting models.

Moreover, the autocorrelation and partial autocorrelation analysis revealed significant lag dependencies in the time series, which were critical in informing the selection of forecasting models such as ARIMA. These findings demonstrate the importance of thorough exploratory data analysis in preparing the data for predictive modeling and highlight the robustness of the dataset for future modeling tasks.

By understanding the statistical properties of the data, the study lays a solid foundation for the development of accurate and reliable forecasting models in the subsequent stages of the research.

4.5.1 AI/ML-based Decision Support Tool

The AI/ML-based decision support tool is the core of this study's approach, designed to predict blood demand, classify donors, and optimize donation scheduling. This tool is composed of three key models that interact with each other to ensure a balanced and efficient blood supply system.

4.5.1.1 Business/Problem Understanding

The problem understanding process consisted of two stages: a field study and a theoretical study. In the field study, meetings were held with healthcare officials in Algeria to study the blood supply chain operations with the help of specialized medical staff. This allowed us to gain a thorough understanding of the core problems, which we identified as three main issues:

1. Uncertainty in blood demand.
2. Imbalance between blood collection and distribution.
3. The interrelated issues of blood shortages and wastage.

In the theoretical study, a review of the literature was conducted, highlighting the need for data-driven techniques to improve blood supply chain management. Resource planning, supply-demand balancing, and donor behavior were identified as key areas requiring further research.

The central question that guides this research is: **How can historical data be leveraged to improve decision-making processes and resolve the issues of uncertainty in blood demand and imbalance between blood collection and distribution?**

The solution to this problem is the AI-based decision support tool, which includes three fundamental models: blood demand forecasting, blood donor classification, and donation scheduling optimization (as shown in Figure 4.7).

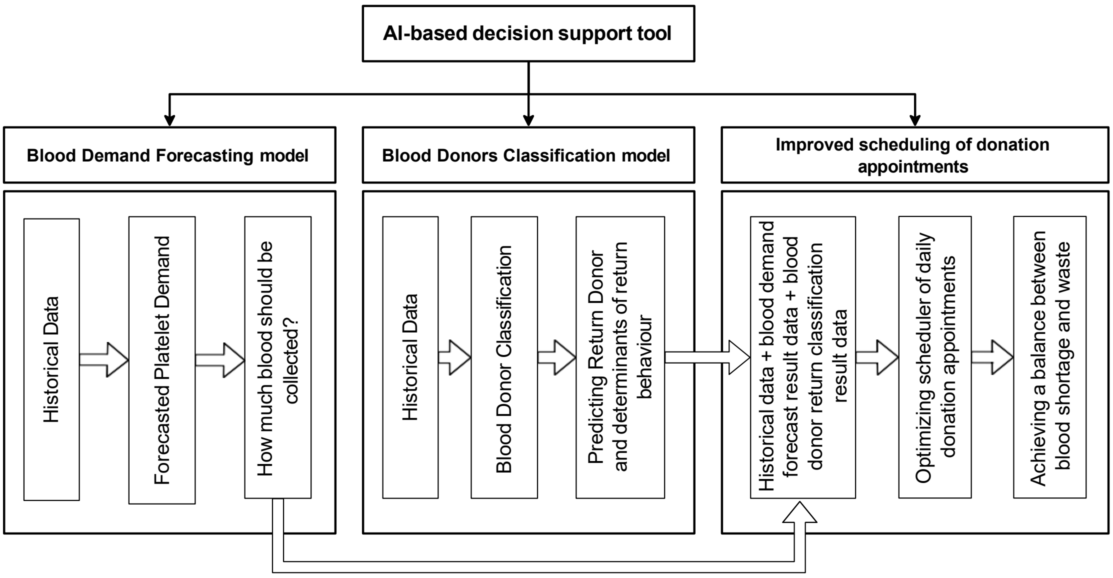


Fig. 4.7 The AI/ML-based decision support tool: An overview of three main components

1. **Blood Demand Forecasting Model:** The prediction of blood component demand is a crucial aspect of the blood supply chain. Accurate demand forecasting allows blood banks to maintain a proper balance between supply and demand, reducing wastage and meeting patient needs efficiently.

2. **Blood Donor Classification Model:** This model classifies blood donors into two groups: those who are likely to return for future donations and those who are not. The classification is based on donor characteristics such as eligibility, donation history, and frequency. In this study, six machine learning algorithms are used for classification: Artificial Neural Networks, Logistic Regression, Support Vector Machines, Decision Trees, Random Forests, Naive Bayes, and K-Nearest Neighbors.

3. **Improved Scheduling of Donation Appointments:** The scheduling model optimizes the allocation of donation appointments to maximize efficiency. By predicting blood demand and classifying donors, the model helps reduce deferrals and wastage, ensuring that blood collection aligns with demand. It also considers donor eligibility and the shelf life of blood components.

4.5.1.2 Data Understanding

For our study, we utilized data collected from the Algerian National Blood Agency (National Blood Agency ANS), several blood centers, and hospitals in the Algiers region. The data was gathered over a four-year period, from 2017 to 2021. Algiers was chosen as the site for data collection for two primary reasons: first, the region has a well-structured blood supply chain, making it ideal for data collection; second, data from other regions were limited, and many blood facilities did not have data for the specified study period.

The collected data was divided into two main categories: blood supply chain data for specific time intervals and donor-related data. These two types of data provide a comprehensive view of the blood supply process, from donation behavior to inventory management.

Blood Supply Data: This data includes information on blood demand and supply trends over time. The supply statistics were collected on a weekly basis, and the data includes quantities of blood donated, blood type, and dates of collection. Table 4.4 shows the data analytics for the blood supply across different years. It is evident that the average blood supply remained stable during most of the study period, except for the drastic decrease observed in 2020 and 2021 due to the COVID-19 pandemic.

Table 4.4 2017–2020 weekly blood supply statistics.

Year	Day	Average	Max.	Standard deviation	Coef. variation (%)
2017	FRIDAY	29.92	47	8.58	28.66
	MONDAY	289.67	359	80.25	27.70
	SATURDAY	34.75	65	10.56	30.39
	SUNDAY	293.67	356	60.63	20.65
	THURSDAY	284.08	337	59.51	20.95
	TUESDAY	288.04	346	79.90	27.74
	WEDNESDAY	293.77	349	71.09	24.20
2018	FRIDAY	34.94	46	7.73	22.12
	MONDAY	288.89	356	87.83	30.40
	SATURDAY	38.08	54	7.79	20.45
	SUNDAY	282.49	358	88.49	31.33
	THURSDAY	269.11	343	75.46	28.04
	TUESDAY	298.11	353	62.64	21.01
	WEDNESDAY	288.00	352	73.19	25.41
2019	FRIDAY	19.06	29	5.71	29.97
	MONDAY	288.11	357	86.86	30.15
	SATURDAY	27.17	49	9.01	33.15
	SUNDAY	286.50	344	93.11	32.50
	THURSDAY	264.53	331	86.45	32.68
	TUESDAY	276.38	349	96.81	35.03
	WEDNESDAY	278.42	340	98.27	35.29
2020	FRIDAY	10.38	31	6.88	66.26
	MONDAY	198.02	358	112.37	56.75
	SATURDAY	14.40	27	7.78	54.03
	SUNDAY	184.25	351	116.87	63.43
	THURSDAY	180.25	337	102.07	56.63
	TUESDAY	204.94	403	107.79	52.60
	WEDNESDAY	197.83	351	108.89	55.04

Donor Data: The donor data includes attributes such as age, gender, blood type, donation history, and donation frequency. Table 4.5 shows the key attributes for blood donors, along with their descriptions.

Table 4.5 Dataset attributes of blood donor.

Attribute	Type	Explanation
ID	Integer	Donor ID
Blood Type	Categorical	$(A^+, A^-, B^+, B^-, O^+, O^-, AB^+, AB^-)$
Gender	Categorical	Male (M), Female (F)
Age	Categorical	Range from 18 to 66
Recency	Discrete	Months since last donation
Frequency	Discrete	Total number of donations
Monetary	Discrete	Total blood donated
Time	Discrete	Months since first donation
Rejected	Discrete	Total number of rejected donations
Class	Binary	1 if return donor, 0 if non-return donor

Blood Donation Trends: Figure 4.8 presents a graphical representation of daily whole blood donations from 2017 to 2021. This chart visualizes the trends in blood donations, providing insights into the patterns of donation volume over the years.

The figure clearly shows how donation trends have fluctuated, with noticeable peaks during certain periods (e.g., public health campaigns, seasonal variations) and significant drops in others, particularly during the COVID-19 pandemic. These patterns are essential for forecasting future blood demand and optimizing collection strategies.

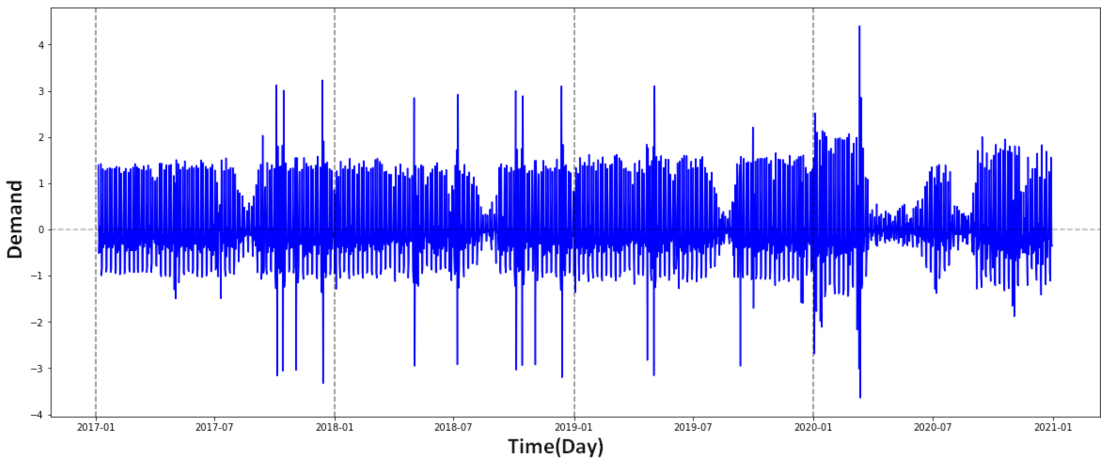


Fig. 4.8 Daily whole blood donations, 2017—2021

4.5.1.3 Modeling and Evaluation

The choice of machine learning and statistical models plays a crucial role in ensuring the accuracy and effectiveness of the blood demand forecasting system. In this study, several models were considered for both forecasting blood demand and classifying donors. The models were evaluated based on their ability to predict future blood needs accurately and to classify donors efficiently.

The following performance metrics were used to evaluate the models:

- **Mean Absolute Error (MAE):** Measures the average magnitude of errors in a set of predictions, without considering their direction.
- **Mean Squared Error (MSE):** Similar to MAE but squares the errors, giving larger errors more weight.
- **Mean Absolute Percentage Error (MAPE):** Represents the accuracy as a percentage, making it easier to interpret the relative performance across different models.

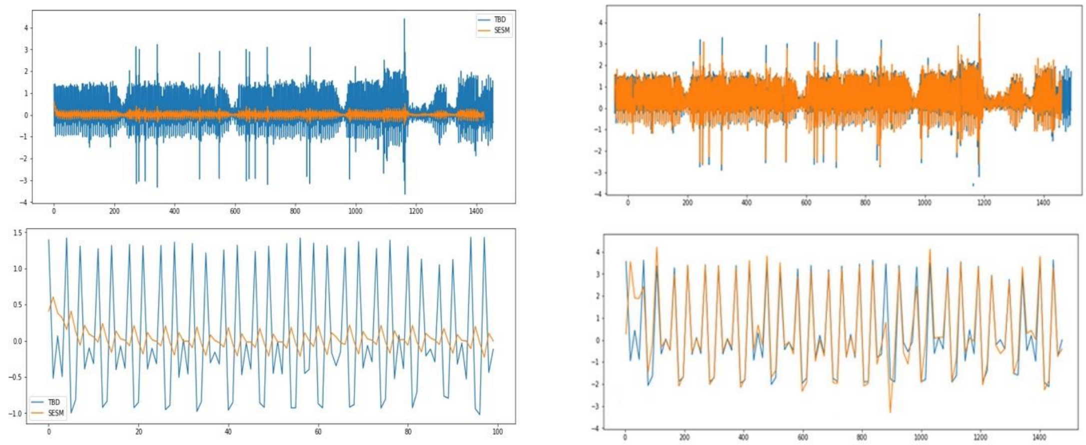
Table 4.6 summarizes the three evaluation metrics used in this research.

Table 4.6 The three evaluation methods of forecasting models performed used.

Method	Equation	Definition
Mean Absolute Error	$\frac{1}{n}\sum_{i=1}^{n} \lvert e_i \rvert$	It calculates the average significance of forecast errors, with all individual errors being given equal weight.
Mean Squared Error	$\frac{1}{n}\sum_{i=1}^{n} e_i^2$	It assesses the significance of forecast inaccuracies, with larger errors penalized more severely because of squaring.
Mean Absolute Percentage Error	$\sum_{i=1} \frac{e_i}{x_i} \times 100$	It indicates the relative value of forecasting errors as a percentage.

Model Training and Improvement: Figure 4.9 demonstrates the improvement in model performance during the training phase. The graphs show how the model's learning process evolves over time, with performance improving as it adapts to the training data. These improvements can be visually tracked by observing the reduction in errors (such as Mean Squared Error or Mean Absolute Error) across training iterations.

The model's ability to generalize better with each cycle is a key indicator of its effectiveness in making accurate predictions. This type of graph is essential for monitoring the model's progress, ensuring that overfitting or underfitting does not occur, and providing insights into necessary adjustments for hyperparameters.



Low Training/Test Accuracy

High Training/Test Accuracy

Fig. 4.9 Improvement in Model Learning Demonstrated by Graphs

4.5.1.4 Model Deployment

Once the model has been trained and validated, it is deployed in a real-world environment to evaluate its practical effectiveness. An experimental study was conducted at one of the blood donation centers in Algiers over a nine-month period, where the most accurate models were deployed to assess their performance in predicting blood demand and optimizing donation schedules.

During deployment, the following steps were taken:

- **Blood Demand Forecasting:** The forecasting model was used to predict future demand for blood components. These predictions helped blood supply managers plan and allocate resources efficiently, reducing both shortages and excess stock.
- **Donor Classification:** The donor classification model was used to categorize donors into various groups, including regular donors, first-time donors, and those at risk of deferral. This information was vital for targeted outreach programs.
- **Scheduling Optimization:** The scheduling optimization model was used to assign donation appointments based on predicted demand, donor availability, and eligibility criteria. This ensured that blood collection aligned with the forecasted demand, minimizing waste and maximizing efficiency.

Figure 4.10 illustrates how the model's predictions evolved over the deployment period, showing a consistent improvement in accuracy as more data was incorporated into the system.

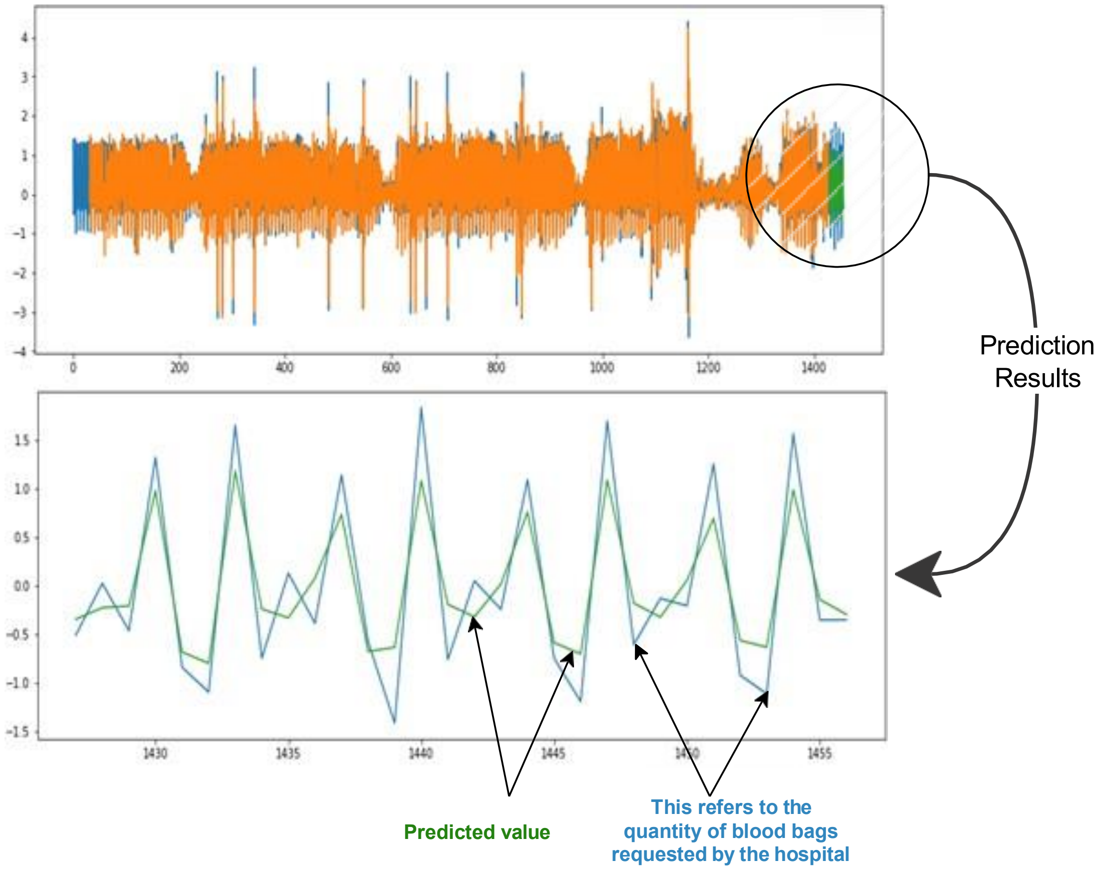


Fig. 4.10 Improvement in Model Learning Demonstrated by Graphs

4.6 Conclusion

In conclusion, the proposed solution for blood supply chain management involves improving all stages of the process, from data collection to predictive modeling and scheduling optimization. The system developed in this research leverages data-driven techniques to address key challenges faced by blood supply chain operations, ensuring a more efficient, accurate, and adaptive system.

The key outcomes of the study are as follows:

1. Predicting the number of unbooked donors on any given day helps regulate blood collection, reduce donor waiting time, and improve health workers' working conditions.

2. Reducing the number of canceled appointments by using the donor classification model, which helps schedule appropriate donors based on the expected blood demand and available blood components.

3. Improving the quality of blood collected by identifying qualified donors and ensuring their eligibility through a detailed scheduling system, reducing unqualified donations and increasing the overall efficiency of the blood supply process.

4. Minimizing blood wastage through optimized scheduling, ensuring that the supply of blood components is better aligned with actual demand.

5. Enhancing the overall blood supply chain by reducing uncertainty, over-collection, and under-collection, which leads to better management of resources and the timely availability of blood components.

Figure 4.11 illustrates the flow of the proposed system, summarizing the main steps of the blood supply chain process, from donor classification and demand forecasting to scheduling optimization and real-time monitoring.

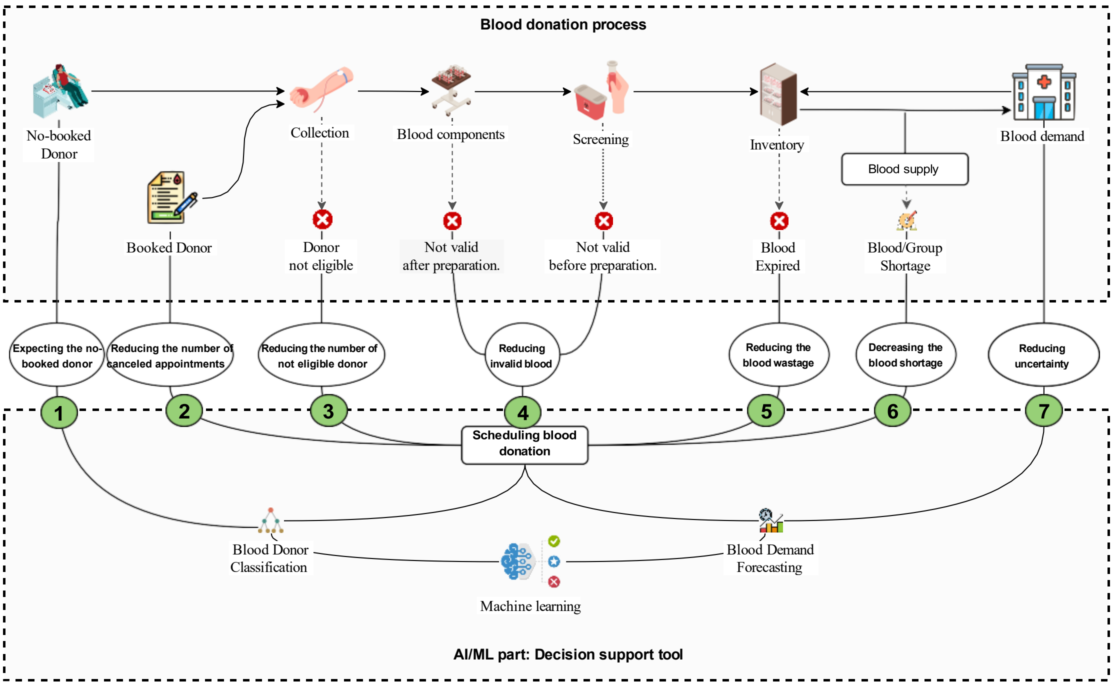


Fig. 4.11 Flow diagram of the proposed system

In the next chapter, we will explore the deployment of the proposed AI/ML-based blood supply chain management system in a real-world environment. The experimental results will be presented, demonstrating the effectiveness of the system in optimizing blood donation schedules, forecasting demand, and improving donor classification. This chapter will also discuss the challenges faced during deployment and the insights gained from the system's operation.

Following the deployment details, a comprehensive analysis of the results will be conducted, highlighting the impact of the system on blood supply chain efficiency and the potential for future enhancements. The discussion will also address the limitations of the current system and suggest possible areas for further research and improvement.

By understanding the deployment and real-world performance of the system, we aim to provide valuable insights into its practical application and its ability to transform blood supply chain management practices.

4.6.1 Feature Engineering Strategy

Based on the statistical diagnostics above, the data was standardized, seasonally adjusted, and encoded to maximize compatibility with the selected forecasting and classification algorithms. Categorical variables such as donation type and location were one-hot encoded, while continuous features were normalized to improve model convergence.

CHAPTER 5

System Deployment, Results, and Discussion

5.1 Introduction

This chapter presents the practical implementation and evaluation of the proposed Smart Platform for Blood Management in Healthcare, designed to enhance the efficiency and reliability of the blood supply chain through the use of digital infrastructure and AI/ML technologies. Unlike the previous chapter, which focused on the design and methodology behind the system, this chapter emphasizes real-world deployment, system functionalities, and measurable outcomes obtained during experimentation. The chapter is structured around three key axes:

- A detailed overview of how the platform was deployed and integrated into a real blood donation center in Algiers.
- A functional walkthrough of the main modules of the system, highlighting their role in daily operations and strategic planning.
- An analysis of the experimental results, including forecasting accuracy, classification performance, and the overall impact of the system on blood management indicators.

Throughout the chapter, screenshots from the platform interface and technical diagrams (use case, activity, sequence diagrams) are incorporated to illustrate workflows and system interactions.

5.2 System Deployment Overview

Fig. 5.1 Super admin interface for managing user access and permissions.

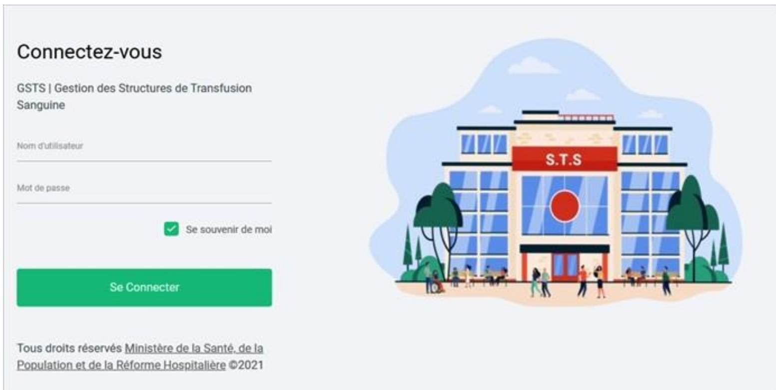


Fig. 5.2 System login interface enabling secure access for authorized users.

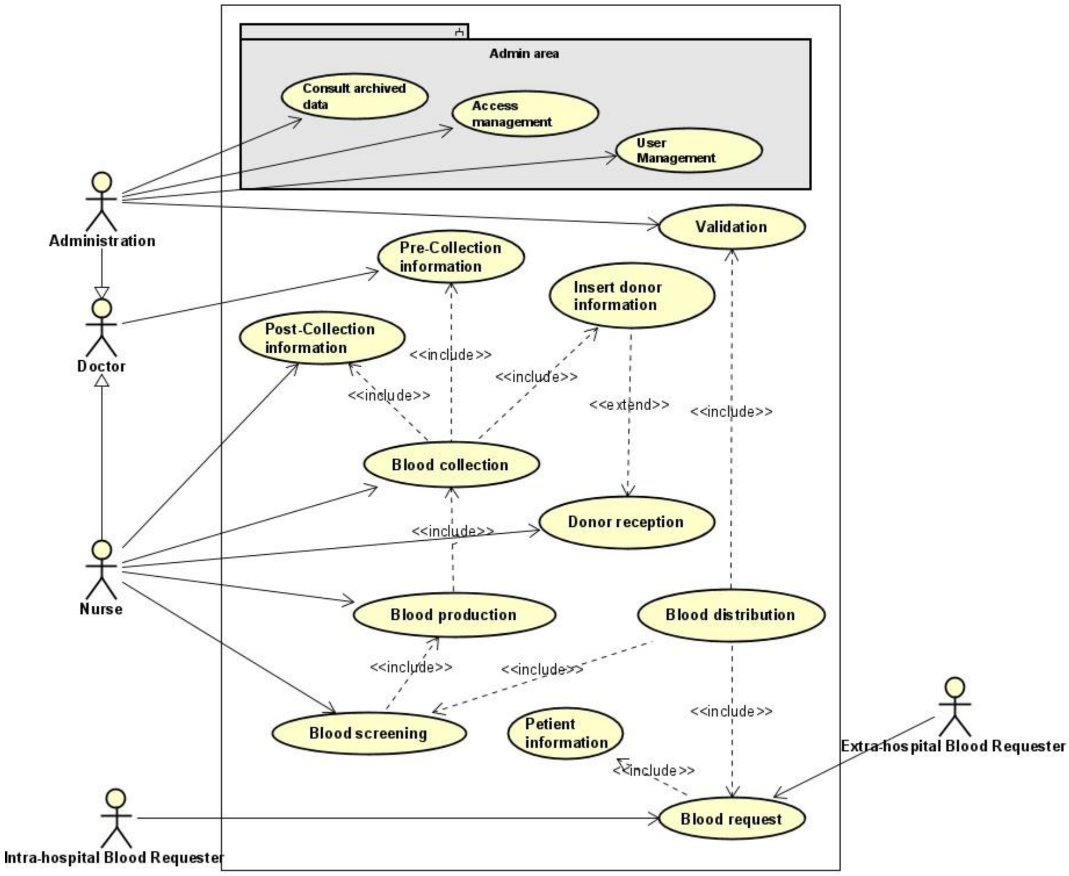


Fig. 5.3 Use Case Diagram showing coordination between hospital actors: administrators, doctors, nurses, and intra/extra-hospital blood requesters.

This section describes the real-world implementation process of the proposed platform within an operational blood donation environment. It covers the deployment context, the phases followed during installation and integration, the architectural structure supporting the platform, the workflow for data handling, and the challenges encountered throughout the deployment process. Each subsection highlights key aspects that shaped the platform's adaptation to practical constraints.

5.2.1 System Deployment Context

The proposed platform was deployed within one of the largest blood donation centers in Algiers, which operates both fixed and mobile collection units. This center was selected due to its extensive donor network, diversified blood product demands, and operational complexity, making it an ideal environment to assess the platform's real-world effectiveness.

The deployment context included a combination of centralized blood collection sites, hospital blood banks, and remote mobile donation units distributed across multiple urban and semi-urban regions. Each site presented distinct challenges in terms of infrastructure readiness, data availability, staff digital literacy, and workflow standardization.

The primary goal was to seamlessly integrate the platform into existing operations without disrupting ongoing blood collection and distribution activities. To achieve this, a phased and flexible deployment strategy was adopted, taking into account the heterogeneity of collection sites, the varying technological capabilities, and the need for minimal downtime during the transition.

This context set the foundation for evaluating the platform's scalability, adaptability, and robustness under real operational constraints.

5.2.1.1 Deployment Environment

The deployment site exhibited the following characteristics:

- **High donor turnout:** The center consistently handled a large number of donors on a daily basis, including peaks during national campaigns and special events. This variability provided an ideal test case for evaluating the platform's forecasting and scheduling modules.
- **Complex inventory management:** The blood bank operated both fixed-site facilities and mobile collection units across multiple regions, requiring dynamic inventory monitoring, distribution planning, and rapid response to demand fluctuations.
- **Data-driven operational reliance:** The center heavily depended on predictive analytics for demand planning, donor targeting, and blood component stock management. Its reliance on real-time data made it a suitable environment for validating the platform's decision support capabilities.

This environment provided the necessary operational challenges and data diversity to rigorously assess the performance and scalability of the platform.

5.2.2 Deployment Phases

The deployment of the platform was conducted through a structured, phased approach designed to minimize operational disruptions and ensure a smooth transition. Each phase built progressively on the previous one, allowing continuous learning, adaptation, and optimization. Figure 5.4 illustrates the main phases and their logical progression.

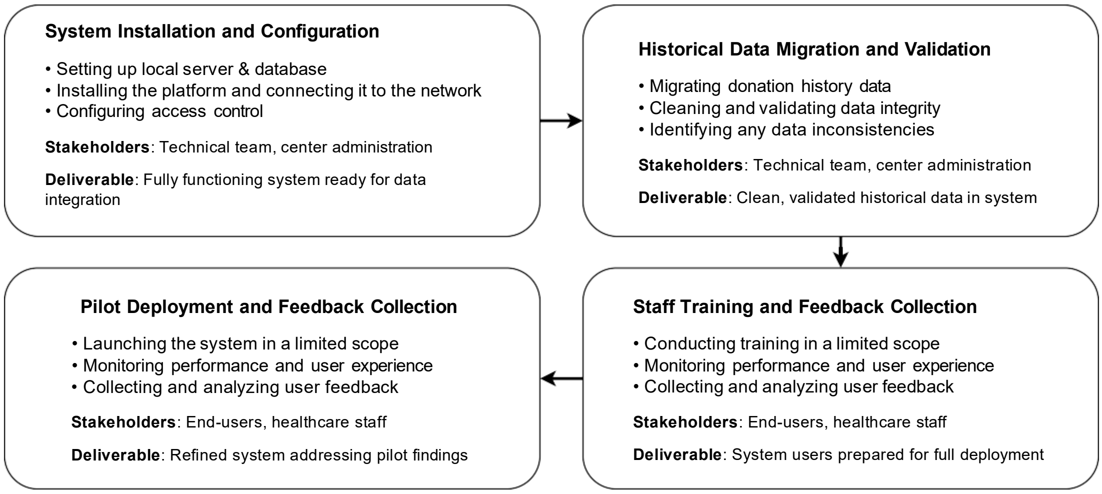


Fig. 5.4 Deployment Phases of the Platform

The major phases of the deployment process are summarized below:

- **System Installation:** Installation of the platform's core components on local servers and integration with hospital networks. Basic system configurations, security setups, and access controls were established during this phase.
- **Historical Data Migration:** Migration of legacy data from previous blood collection and inventory management systems into the new platform. This included cleaning, validating, and standardizing historical records to ensure data integrity.
- **Staff Training and Orientation:** Conducting structured training sessions for healthcare professionals, technicians, and administrative staff. The focus was on practical usage, common workflows, and handling exceptions or system errors.
- **Pilot Deployment:** Launching the platform at a limited scale within select blood collection units. This phase tested real-time operational integration, collected user feedback, and helped identify minor issues or required adjustments.
- **Full Rollout:** After successful pilot testing and refinement, the platform was fully deployed across all operational units of the center, including mobile collection teams. Ongoing monitoring and technical support accompanied the rollout to ensure stability and responsiveness.

5.2.3 Deployment Challenges and Solutions

Despite the structured deployment strategy, several challenges were encountered during the implementation of the platform in a real-world setting. These issues spanned technical, organizational, and human factors, each requiring tailored solutions to ensure the success of the rollout.

- **Technical Constraints:** Some sites suffered from limited network connectivity, outdated hardware, or lack of integration with legacy systems. These issues were mitigated by introducing offline data capture modules, upgrading essential infrastructure, and developing custom synchronization scripts.
- **Data Migration Difficulties:** Historical data were inconsistently formatted and often incomplete. To address this, a multi-step migration process was implemented, including data preprocessing, validation scripts, and fallback manual verification where necessary.
- **Resistance to Change:** Some staff members expressed hesitation in using the new system due to limited digital skills or preference for manual methods. A series of targeted workshops and practical training sessions were conducted, followed by on-site support and iterative UI improvements based on user feedback.
- **Support and Maintenance:** During the first months post-deployment, regular support requests were logged, mostly related to unfamiliar workflows or minor system errors. A ticketing system was introduced to track and resolve issues efficiently, supported by a local helpdesk and remote technical team.

By addressing these challenges early in the deployment, the platform was successfully stabilized, and user confidence improved progressively, paving the way for its full operational adoption.

Table 5.1 Summary of Deployment Challenges and Adopted Solutions

Challenge	**Solution**
Limited network infrastructure and hardware constraints	Deployed offline data modules, upgraded critical systems, and created lightweight versions of the platform for low-resource settings.
Inconsistent and incomplete historical data	Implemented multi-stage migration with data preprocessing, validation scripts, and manual correction support.
Resistance to change among staff	Delivered hands-on training workshops, integrated feedback into UI design, and provided continuous on-site support.
Increased support needs during early use	Introduced a ticketing system and helpdesk, backed by a remote support team to address issues efficiently.

5.3 System Architecture Overview

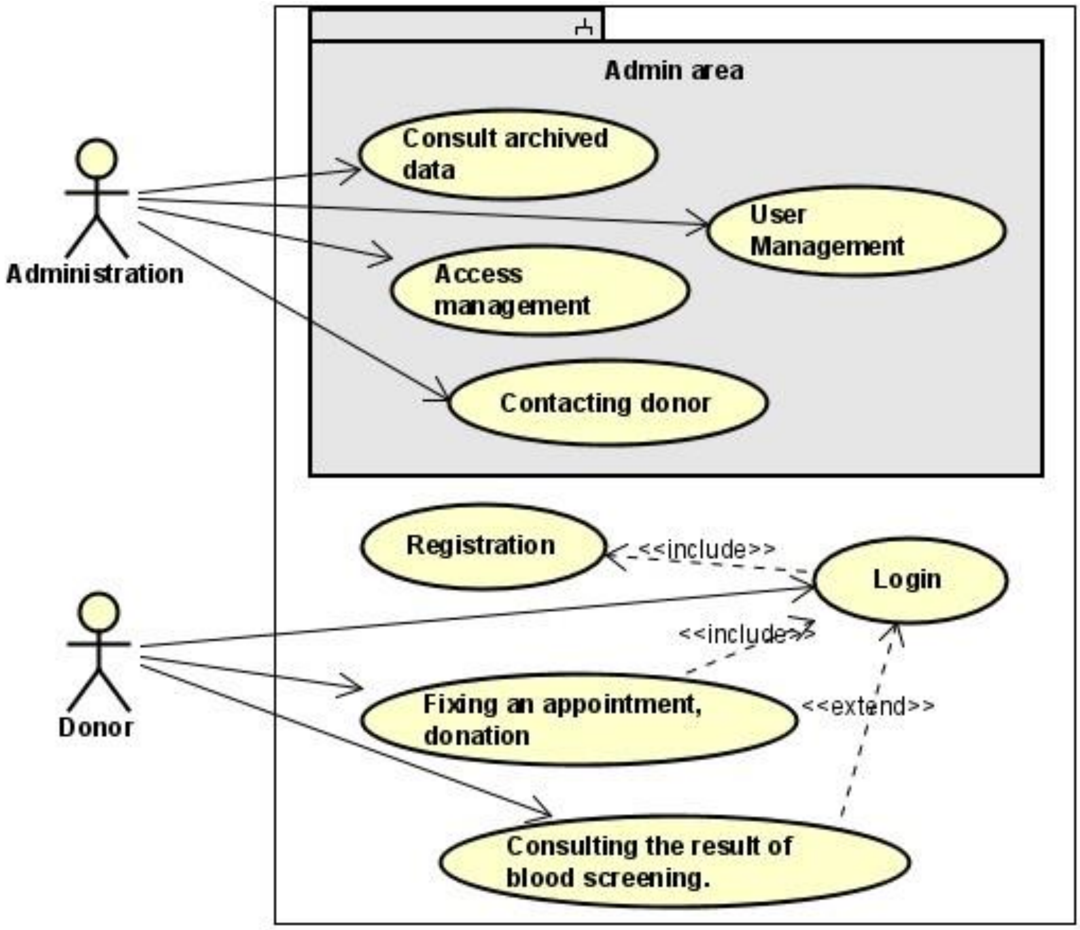


Fig. 5.5 Use Case Diagram illustrating interactions between donors and the digital platform.

The architecture of the proposed smart blood management platform is designed to support a comprehensive and integrated approach to blood collection, processing, inventory management, and distribution. It is structured in multiple layers, each of which handles a specific set of functionalities, from real-time data collection to decision-making and reporting.

This layered architecture ensures flexibility, scalability, and seamless integration with existing operational systems, while also supporting future upgrades or additions. By dividing the system into distinct layers, the platform achieves a high level of modularity, allowing each component to function independently while still working cohesively as part of the larger ecosystem.

In the following sections, we provide a detailed overview of the key layers of the system architecture, followed by an in-depth explanation of the data collection, validation processes, and the decision support tools embedded within the platform.

5.3.1 Data Collection and Validation Workflow

The data collection process within the platform covers various sources, including blood donors, blood banks, hospitals, and mobile donation units. Data is captured using digital forms, automated systems, and external database integrations.

Once collected, the data undergoes a validation and cleaning phase. This phase ensures that data are complete, correctly formatted, free of duplications, and standardized across all entities. Critical validation steps include checking donor eligibility criteria, verifying blood type compatibility, and auditing inventory records.

Figure 5.6 illustrates the workflow from initial data collection to validation and integration into the platform database.

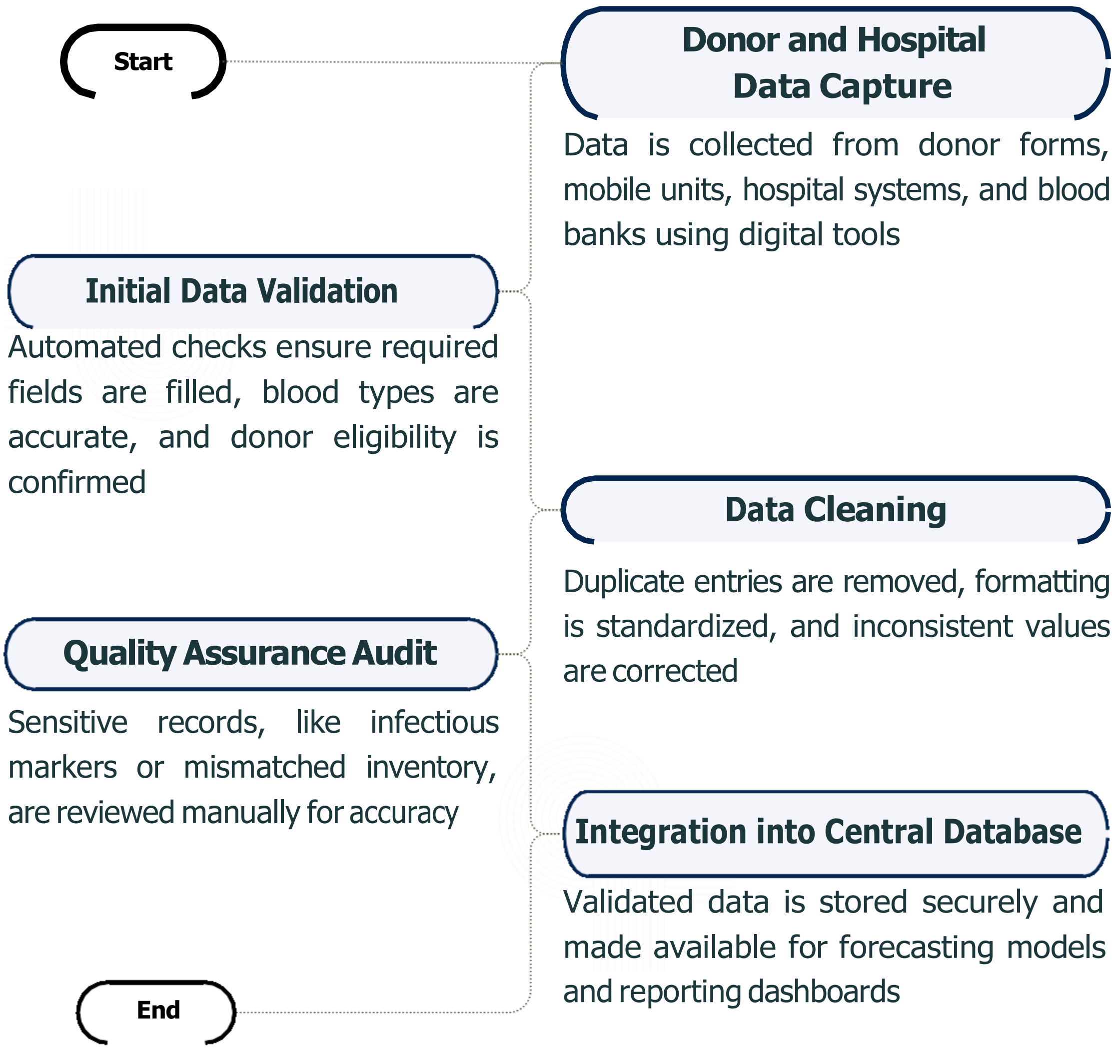


Fig. 5.6 Data Collection and Validation Workflow

The main steps in the data workflow are:

- **Data Collection:** Data is collected through digital forms, mobile applications, and hospital databases. This process ensures that all relevant information, including donor eligibility, blood type, and donation history, is captured in real-time.
- **Initial Validation:** Basic validation checks are performed to ensure that all mandatory fields are filled, and the data is correctly formatted. This step also ensures that donor eligibility and blood type information match predefined criteria.
- **Data Cleaning:** Duplicate data entries are identified and merged or removed. Inconsistent values are corrected, and missing data is supplemented using predefined rules or manual intervention when necessary.
- **Quality Assurance:** The cleaned data undergoes a final quality assurance step, where critical data points such as infectious disease markers and inventory levels are audited manually to ensure that they meet operational standards.
- **Data Integration:** Validated and cleaned data is integrated into the platform's central database, ready for further processing by decision support models and reporting systems.

5.3.2 System Layers Overview

The proposed platform was designed based on a multi-layered architecture to ensure modularity, scalability, and efficient management of the blood supply chain processes. This section presents an overview of the key layers composing the system Figure 5.7:

- **Data Collection Layer:** Responsible for gathering donor, blood, and hospital information through digital forms, mobile units, and integration with hospital databases.
- **Data Processing and Validation Layer:** Ensures that collected data are cleaned, validated, and standardized before being used in the system.
- **Decision Support Layer:** Hosts AI/ML models for blood demand forecasting, donor classification, and donor scheduling optimization.
- **Application Layer:** Provides user interfaces for healthcare professionals, staff members, and administrators to interact with the system functionalities.
- **Monitoring and Reporting Layer:** Generates dashboards, KPIs, and statistical reports to assist in operational decision-making and strategic planning.

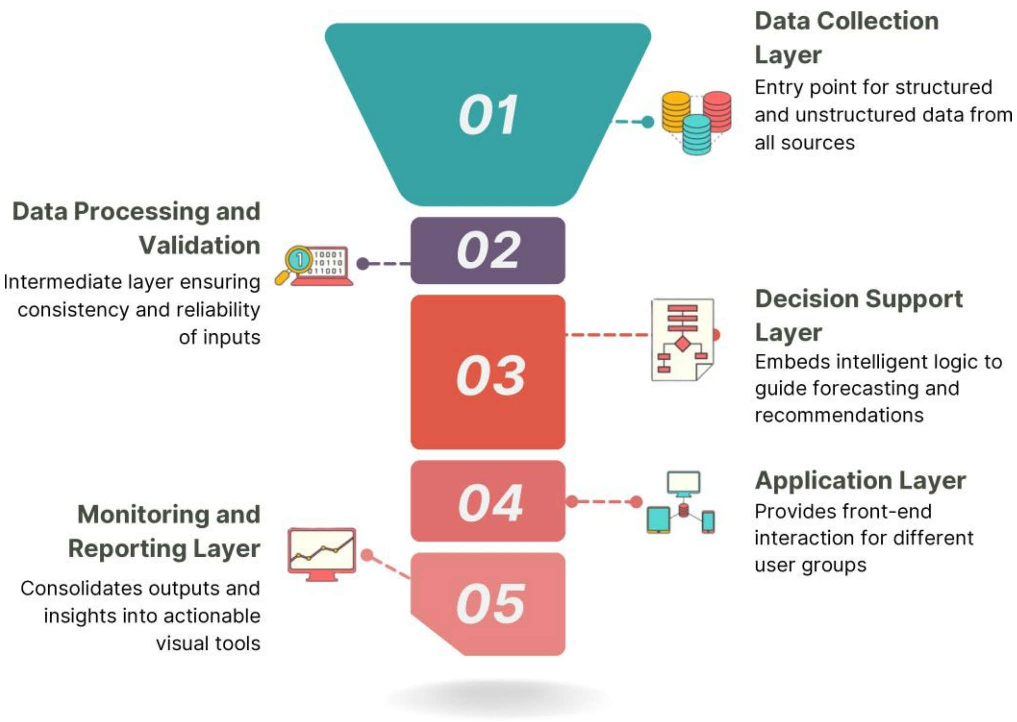


Fig. 5.7 Overview of the System Architecture: Functional Layers

This layered structure allows for flexibility and easy maintenance, while enabling seamless integration of new components and algorithms as needed.

5.4 Smart Platform Modules Walkthrough

The platform is designed to manage the entire blood supply chain, with several modules that handle different stages of the process, from blood collection to distribution. In this section, we will walk through the core modules of the platform, starting from the initial steps of blood collection, to the prediction of demand, classification of donors, and the optimization of donor scheduling. Each module is key to ensuring efficient operations and reducing waste, while improving the overall blood management system.

5.4.1 Blood Collection and Processing Management

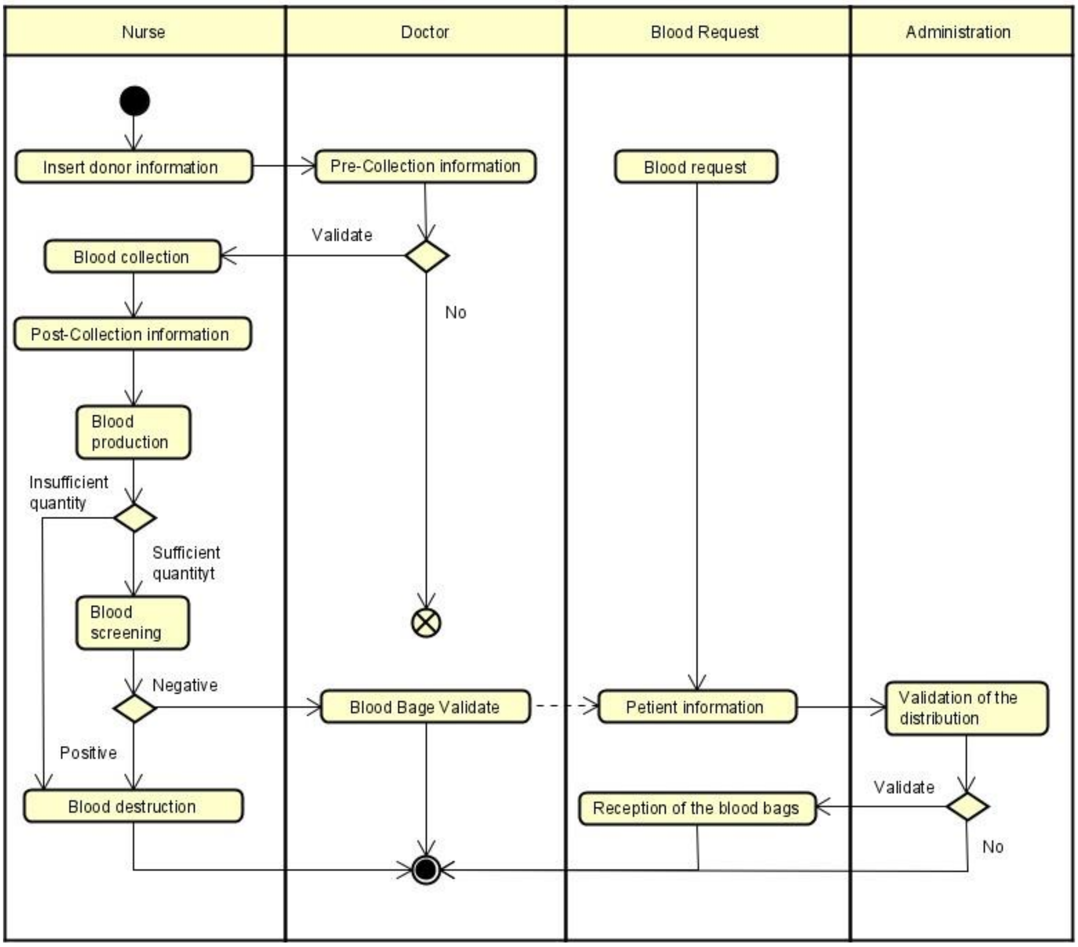


Fig. 5.8 Activity Diagram for the end-to-end blood donation workflow, from donor registration to final distribution.

The Blood Collection and Processing Management module is crucial for the platform's operational functionality, ensuring the smooth flow from donor registration, through blood collection, to blood processing and storage. This module guarantees full traceability, enhances quality control, and boosts operational efficiency within blood collection centers. By automating various stages of the process, it minimizes manual errors, reduces waiting times for donors, and optimizes blood product handling from collection to distribution.

5.4.1.1 Donor Registration and Pre-screening

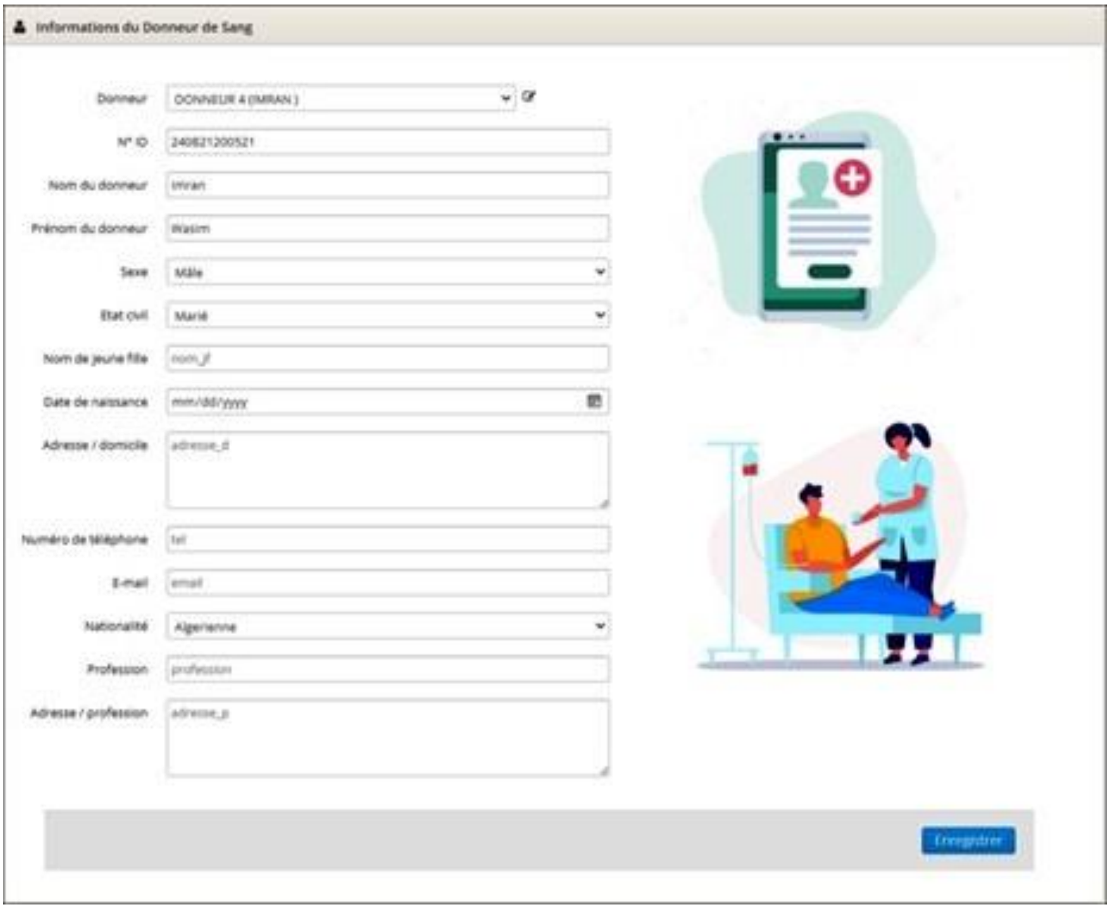


Fig. 5.9 Donor information entry interface used during registration and pre-screening.

The donor registration and pre-screening process begins when an individual expresses interest in donating blood at a fixed or mobile donation unit. Donors fall into three main categories: regular, occasional, and voluntary donors for specific patients. Based on this classification, the registration workflow adapts accordingly.

- **Regular donors** can log into the system, update personal information, and directly schedule appointments.
- **Occasional donors** and **voluntary donors** fill out a one-time form and may link their donation to a specific patient.
- **Donor Registration and Pre-screening:** Upon arrival at the blood center or mobile unit, donors are registered through a digital interface. Medical eligibility checks (age, weight, hemoglobin levels, medical history) are recorded and validated automatically.
- **Blood Collection Management:**

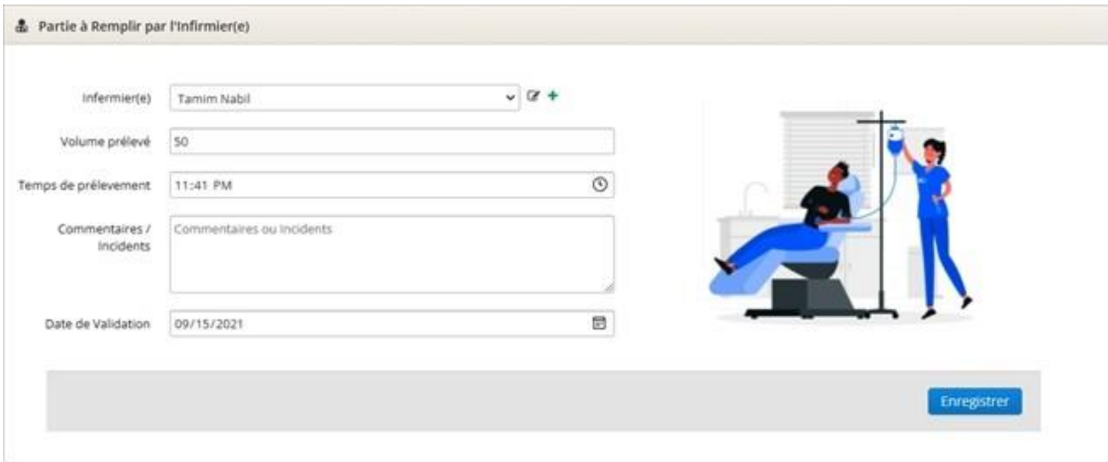


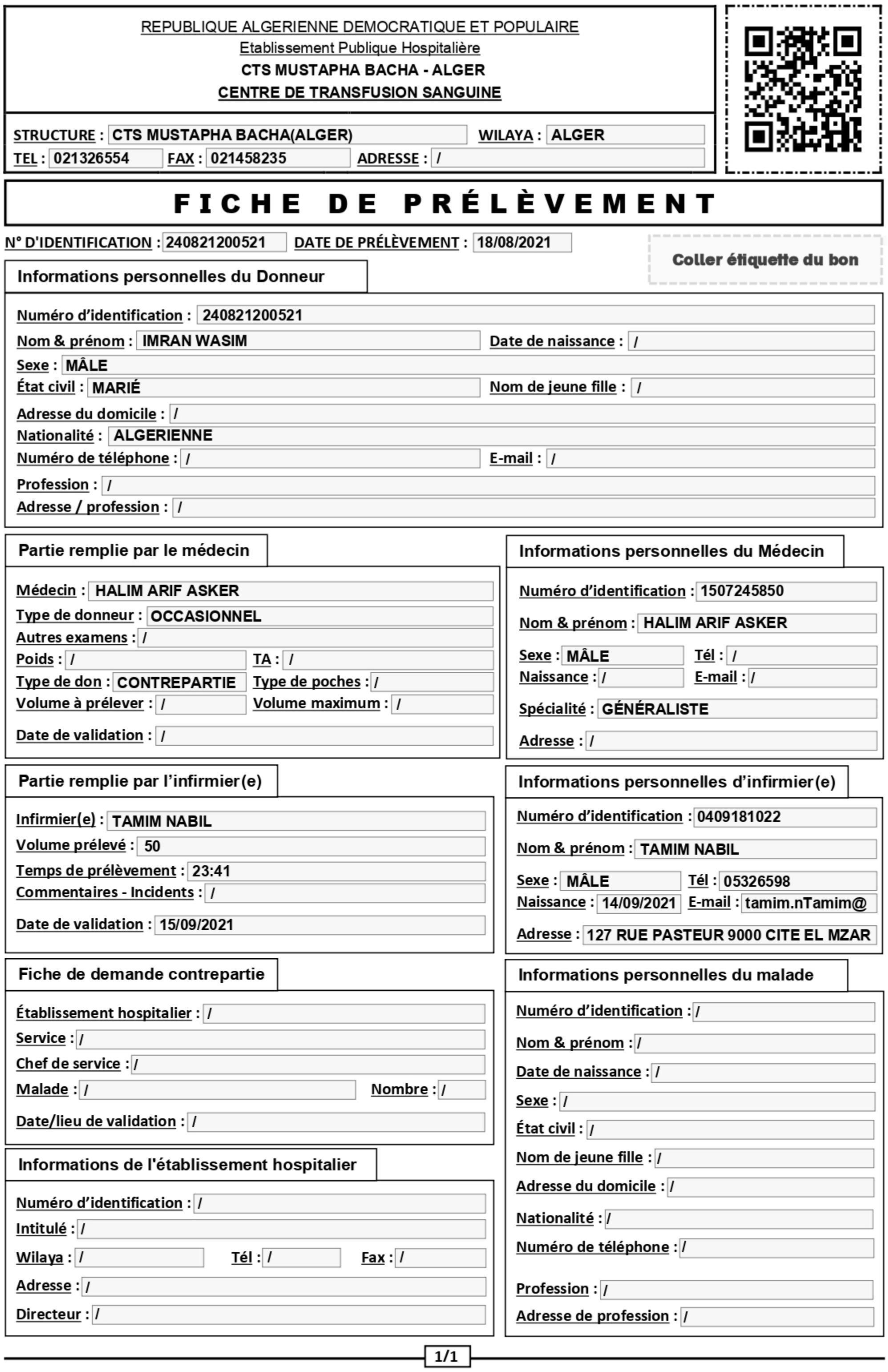
Fig. 5.10 Nurse input interface for documenting blood collection and any incidents.

Fig. 5.11 Collection management screen showing ID generation and status tracking across phases.

Blood collection procedures (whole blood, platelet apheresis, plasma donation) are recorded electronically. The system generates collection IDs and links donations to the donor profile for full traceability.

- **Blood Testing and Processing:** Collected blood units are sent to the laboratory for testing against infectious diseases (HIV, Hepatitis B and C, Syphilis, etc.). Results are digitally entered into the platform, linking test outcomes to each blood unit.

- **Component Production:**



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Etablissement Publique Hospitalière
CTS MUSTAPHA BACHA - ALGER
CENTRE DE TRANSFUSION SANGUINE

STRUCTURE : CTS MUSTAPHA BACHA(ALGER) WILAYA : ALGER
TEL : 021326554 FAX : 021458235 ADRESSE : /

FICHE DE PRÉLÈVEMENT

N° D'IDENTIFICATION : 240821200521 DATE DE PRÉLÈVEMENT : 18/08/2021

Coller étiquette du bon

Informations personnelles du Donneur

Numéro d'identification : 240821200521
Nom & prénom : IMRAN WASIM Date de naissance : /
Sexe : MÂLE
État civil : MARIÉ Nom de jeune fille : /
Adresse du domicile : /
Nationalité : ALGERIENNE
Numéro de téléphone : / E-mail : /
Profession : /
Adresse / profession : /

Partie remplie par le médecin

Médecin : HALIM ARIF ASKER
Type de donneur : OCCASIONNEL
Autres examens : /
Poids : / TA : /
Type de don : CONTREPARTIE Type de poches : /
Volume à prélever : / Volume maximum : /
Date de validation : /

Informations personnelles du Médecin

Numéro d'identification : 1507245850
Nom & prénom : HALIM ARIF ASKER
Sexe : MÂLE Tél : /
Naissance : / E-mail : /
Spécialité : GÉNÉRALISTE
Adresse : /

Partie remplie par l'infirmier(e)

Infirmier(e) : TAMIM NABIL
Volume prélevé : 50
Temps de prélèvement : 23:41
Commentaires - Incidents : /
Date de validation : 15/09/2021

Informations personnelles d'infirmier(e)

Numéro d'identification : 0409181022
Nom & prénom : TAMIM NABIL
Sexe : MÂLE Tél : 05326598
Naissance : 14/09/2021 E-mail : tamim.nTamim@
Adresse : 127 RUE PASTEUR 9000 CITE EL MZAR

Fiche de demande contrepartie

Établissement hospitalier : /
Service : /
Chef de service : /
Malade : / Nombre : /
Date/lieu de validation : /

Informations personnelles du malade

Numéro d'identification : /
Nom & prénom : /
Date de naissance : /
Sexe : /
État civil : /
Nom de jeune fille : /
Adresse du domicile : /
Nationalité : /
Numéro de téléphone : /
Profession : /
Adresse de profession : /

Informations de l'établissement hospitalier

Numéro d'identification : /
Intitulé : /
Wilaya : / Tél : / Fax : /
Adresse : /
Directeur : /

1/1

Fig. 5.12 Auto-generated collection summary form consolidating validated donation data.

Qualified blood units are processed into components (Red Blood Cells, Platelets, Plasma). The system records each separated component, assigning unique identifiers for inventory tracking.

- **Storage and Inventory Integration:**

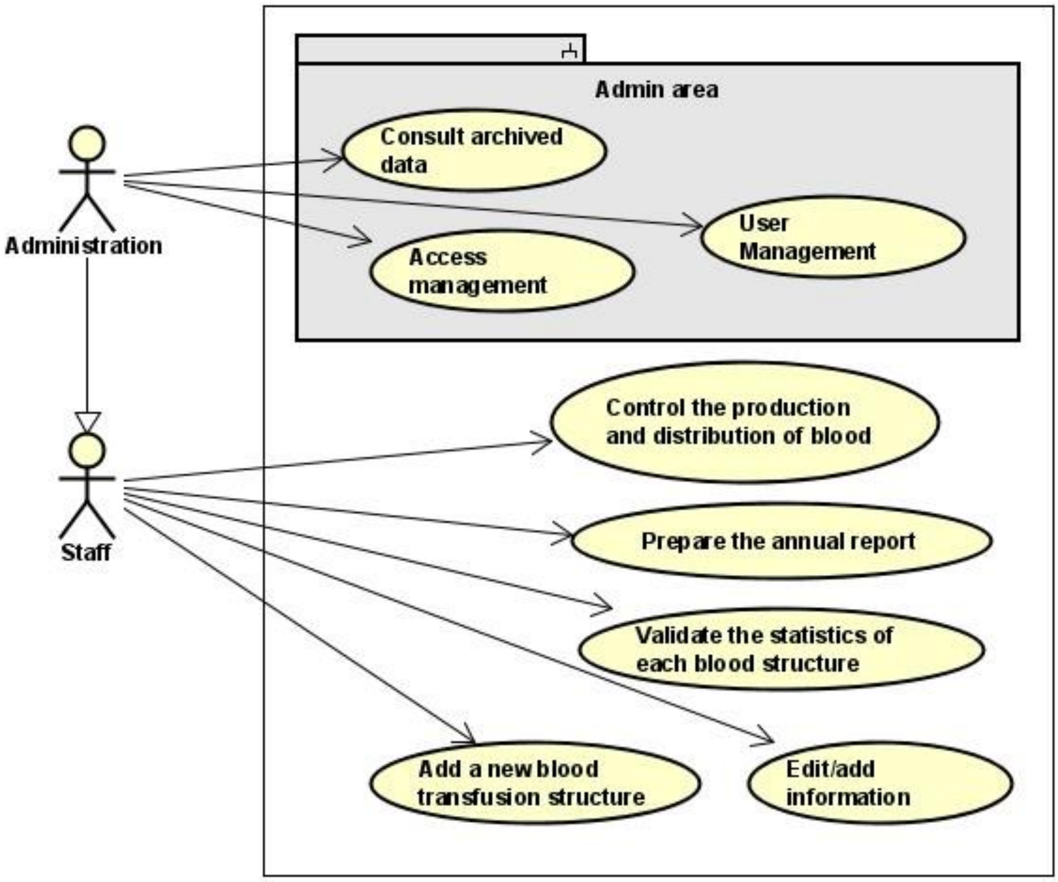


Fig. 5.13 Use Case Diagram for managing blood units across preparation, qualification, and distribution phases.

Blood components are stored under monitored conditions (temperature, expiration dates). The system updates the inventory in real-time, tagging units based on expiry risk (First-In, First-Out principle).

5.4.2 Data Monitoring and Reporting Unit

The Data Monitoring and Reporting module ensures continuous tracking and reporting of key metrics across all participating blood centers. It provides valuable insights into performance, blood usage trends, and regional disparities in donations. This module enables transparency, aids in decision-making, and supports strategic planning at both local and national levels.

- **Data Collection and Aggregation:** The platform aggregates operational data from blood banks and mobile units across the national territory, including donation counts, blood types, deferrals, wastage rates, etc.

- **Periodic Reporting Engine:** Every three months, the platform automatically generates comprehensive reports summarizing blood supply chain performance, including donation trends and wastage analysis.
- **Validation and Data Integrity Checks:** The system automatically performs validation checks to ensure data consistency and correct any discrepancies before report generation.
- **Reporting Recipients and Usage:** Completed reports are forwarded to the National Blood Agency for strategic planning and operational monitoring, supporting decisions on blood distribution and donor engagement.

5.4.3 Blood Donor Classification Module

The Blood Donor Classification module leverages machine learning techniques to predict donor behavior, specifically the likelihood that a donor will return for future donations. This module is designed to help blood centers identify and prioritize reliable donors, ensuring optimal blood collection and reducing uncertainties in donor availability.

- **Purpose and Approach:** The module uses machine learning to classify donors into ”likely to return” and ”unlikely to return” based on their donation history and eligibility factors.
- **Data Inputs and Features:** The system uses demographic data, donation history, and eligibility features to predict donor behavior.
- **Machine Learning Models Used:** Various machine learning algorithms, including Artificial Neural Networks, Logistic Regression, and Random Forests, were tested to predict donor retention.
- **D. Classification Output and Usage:** The system assigns return likelihood scores to each donor, enabling better scheduling and targeted outreach campaigns.

5.4.4 Blood Demand Forecasting Module

The Blood Demand Forecasting module is vital for predicting future blood needs based on historical data and various influencing factors. This module helps blood centers plan collection activities and manage inventories effectively, ensuring that supply aligns with actual demand.

- **Forecasting Objective and Importance:** This module forecasts future blood demand by analyzing historical data, seasonal trends, and regional disparities.

- **Data Preparation and Time Series Structure:** Models are trained using time series data, including blood collection records and seasonal patterns.
- **Forecasting Models Used:** Various models, including ARIMA and machine learning approaches, were compared for forecasting accuracy.
- **Forecast Outputs and Usage:** The forecast outputs help plan future blood donation campaigns, manage inventories, and support policy decisions.

5.4.5 Donor Scheduling Optimization Module

The Donor Scheduling Optimization module helps maximize the efficiency of donor appointments by aligning collection efforts with forecasted demand and donor behavior predictions. This module ensures that resources are optimally utilized and that appointment cancellations and deferrals are minimized.

5.4.5.1 Objective and Approach

By using donor classification and blood demand predictions, the module optimizes scheduling to meet collection goals.

Figure 5.12 - Donor Scheduling Optimization Workflow Diagram, showing how data inputs are processed to optimize scheduling.

5.4.5.2 Optimization Inputs

Includes forecasted demand, donor classifications, and available slots at collection centers.

5.4.5.3 Optimization Model

Uses optimization techniques like integer programming to assign slots and balance appointments across collection units.

5.4.5.4 Output and System Interface

The module generates a detailed schedule that staff can review and adjust through an interactive interface.

5.5 Results and Experimentation

This section presents the experimental evaluation of the proposed Smart Platform for Blood Management in Healthcare. The performance of the forecasting and classification modules is assessed, along with the outcomes of donor scheduling optimization.

We also show how the system contributes to operational and strategic decision-making in the blood supply chain.

The experiments were based on real operational data collected over a 10-month period (March to December 2021) from a major blood donation center in Algiers. This data includes historical donation records, inventory movements, and testing results, which were used to train, validate, and evaluate the AI/ML models integrated into the platform.

5.5.1 Experimental Setup

The experimental framework consisted of the following steps:

- Data extraction and preprocessing from historical records covering donation, testing, and inventory management.
- Training and evaluation of forecasting and classification models using cross-validation techniques.
- Integration and testing of the scheduling optimization engine.
- KPI monitoring before and after system deployment.

provides an overview of the experimental workflow, from data acquisition to performance analysis.

5.5.2 Blood Demand Forecasting Results

To assess the forecasting module, several models were tested on monthly blood collection data segmented by blood type:

- SARIMA
- Holt-Winters
- Support Vector Regression (SVR)
- Random Forest
- Artificial Neural Networks (ANN)

Models were evaluated using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).

5.5.3 Blood Donor Classification Results

Classification models were evaluated using historical donor data to predict return likelihood:

- ANN, Logistic Regression, SVM, Decision Trees
- Random Forest, Naive Bayes, K-Nearest Neighbors

5.5.4 Donor Scheduling Optimization Results

The effectiveness of the scheduling module was evaluated using pre- and post-deployment comparisons of key KPIs:

- Reduction in deferrals
- Decrease in appointment cancellations
- Increase in attended donations vs scheduled

5.5.5 System Screenshots

To better illustrate the platform in operation, this section includes selected screenshots from the deployed system:

5.6 The Impact of the Proposed Solution

Following the deployment of the proposed smart platform, a measurable positive impact was observed across all phases of the blood supply chain. Improvements were recorded in blood collection efficiency, testing and screening outcomes, inventory accuracy, and regional distribution effectiveness. This section presents key performance insights derived from real-world data collected during the platform's trial deployment.

5.6.1 Blood Collection Reports

5.6.1.1 Overall Collection Volume

The system facilitated more targeted blood drives and improved planning, resulting in an increase in the volume of blood collected.

5.6.1.2 Donor Demographics Analysis

The platform enabled detailed demographic profiling of donors, supporting more effective outreach campaigns. illustrates the distribution of donations by age group.

5.6.2 Blood Testing Reports

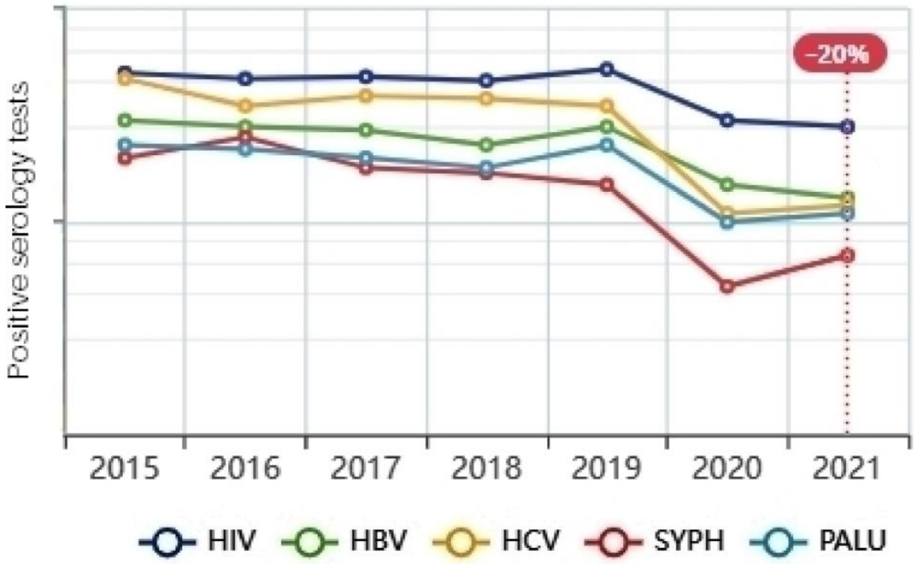


Fig. 5.14 Positive serology results by infection type over the years.

5.6.2.1 Positive Infectious Serology Rates

After implementing the donor classification and scheduling modules, the rate of positive serology results dropped.

5.6.3 Inventory and Distribution Reports

Fig. 5.15 Annual blood stock metrics and waste reduction trends from 2015 to 2021.

5.6.3.1 Blood Inventory Statistics

Real-time monitoring and predictive planning reduced overstocking and expired units. Table **??** compares inventory KPIs before and after the platform launch.

5.6.3.2 Distribution Efficiency

Improvements in forecasting and inventory visibility enabled better inter-center redistribution. illustrates the reduction in emergency redistribution events.

5.6.3.3 Validated Outcomes and Observations

The collected performance metrics and field feedback validate the practical benefits of the proposed platform across multiple dimensions of the blood supply chain. Improvements were consistently observed in operational agility, data-driven planning, and service responsiveness in participating centers.

Key outcomes included reduced appointment cancellations, improved inventory turnover, more accurate forecasting, and enhanced coordination between collection, testing, and distribution units. Furthermore, the platform facilitated better engagement with repeat donors and allowed decision-makers to respond proactively to regional demand fluctuations.

These findings not only demonstrate the system's effectiveness but also highlight its scalability and adaptability to different operational contexts. The integration of AI and digital tools into the blood supply workflow proved instrumental in supporting both routine operations and strategic health planning.

5.7 Discussion

The deployment and evaluation of the Smart Platform for Blood Management revealed critical insights into the operational and technological challenges of managing a national blood supply chain. This section interprets the results obtained, discusses the practical benefits and feedback from users, highlights implementation challenges, and outlines key lessons and future recommendations.

5.7.1 Interpretation of Key Results

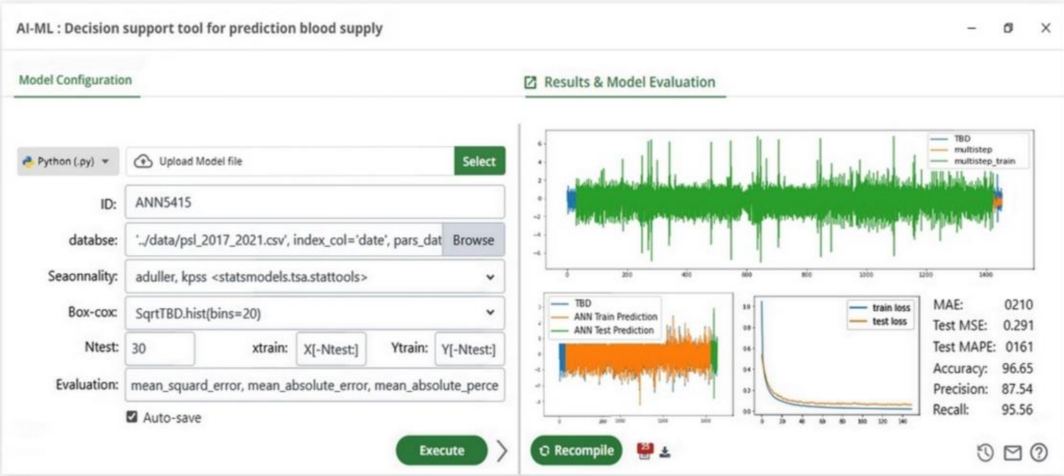


Fig. 5.16 Forecasting module evaluation interface and predictive accuracy output.

The experimentation phase showed substantial improvements in blood supply chain performance:

- **Forecasting Accuracy:** Time series models, particularly SARIMA and Holt-Winters, demonstrated strong predictive capabilities, improving planning and inventory control.
- **Donor Behavior Prediction:** Machine learning models, especially Random Forest and ANN, successfully classified donors into return likelihood categories, helping optimize engagement strategies.
- **Scheduling Optimization:** The intelligent scheduling module significantly reduced appointment cancellations and donor deferrals, enhancing operational efficiency.
- **Operational Efficiency:** Tangible improvements were recorded in blood collection volumes, serology screening outcomes, inventory turnover, and regional distribution logistics.

The positive correlation between AI-driven predictions and operational performance validates the system's design and its relevance in practical healthcare environments.

5.7.2 User Experience and Feedback

Feedback from system users provided additional insights into the platform's usability and adoption challenges:

Positive Aspects:

- Intuitive and user-friendly interfaces.
- Reduced administrative workload due to automated forecasting and reporting.
- Improved responsiveness to demand fluctuations.

Challenges Reported:

- Initial resistance to digital transformation by some staff.
- Learning curve related to advanced data entry and validation tools.
- Occasional data inconsistencies requiring manual correction during the early phases of deployment.

These insights reinforce the importance of ongoing user support and digital literacy development.

5.7.3 Implementation Challenges

Despite its overall success, the platform's deployment revealed several technical and organizational challenges:

- **Legacy Data Integration:** Inconsistent or incomplete historical data complicated the migration and required multiple preprocessing cycles.
- **Connectivity Issues:** Intermittent network availability, especially in remote centers, occasionally disrupted real-time synchronization.
- **User Training Requirements:** Continuous training was necessary to ensure system usability and to onboard staff with limited IT experience.

5.7.4 Lessons Learned and Recommendations

Several lessons emerged from the development and deployment process:

- **Data Quality is Essential:** AI model performance depends heavily on clean, validated data. Robust validation protocols should be in place from day one.
- **User-Centric Design:** Interfaces must be clear, efficient, and tailored to user roles, with a focus on minimizing cognitive load and streamlining workflows.
- **Incremental Rollout is Effective:** Gradual deployment through pilot programs allows time to identify and address technical and human factors before full-scale implementation.
- **Ongoing Training and Support:** Dedicated user support and regular training help maintain confidence, especially in hybrid or paper-to-digital transitions.

- **Platform Modularity Supports Scalability:** Modular system architecture facilitates future expansion, integration with other health platforms, and adaptation to policy changes.

These recommendations should guide future implementations and provide a framework for scaling the platform nationally or in other contexts.

5.8 Key Findings and Insights

This chapter presented the practical deployment, experimentation, and evaluation of the Smart Platform for Blood Management. By implementing the platform in a major Algerian blood center, we were able to assess its real-world impact on healthcare operations and logistics.

The chapter outlined the system's architecture and core modules, including data digitization, AI-driven forecasting, donor behavior prediction, automated scheduling, and integrated performance dashboards. The experimental analysis revealed substantial operational benefits:

- Notable improvements in forecasting accuracy and donor retention rates.
- Lower blood wastage and improved inventory management through predictive scheduling.
- More transparent and timely reporting that enabled quicker strategic interventions.
- Stronger coordination between regional centers, optimizing blood collection and redistribution.

To quantitatively summarize these results, Table 5.2 compares key performance indicators before and after the deployment of the platform:

Table 5.2 Summary of Key Performance Indicators Before and After Platform Deployment

KPI	Before	After	Change
Forecasting Accuracy (MAPE)	17.6%	8.3%	↑ 52.8%
Donor Return Rate	38%	57%	↑ 50.0%
Appointment Cancellation Rate	22%	9%	↓ 59.0%
Expired Blood Units	6.5% of stock	2.1% of stock	↓ 67.7%
Inventory Stockout Events (monthly avg.)	4.2	1.0	↓ 76.2%
Positive Serology Rate	2.3%	1.4%	↓ 39.1%
Average Report Compilation Time	5 days	< 1 day	↓ 80%

These findings support the conclusion that the proposed platform significantly improves blood supply chain performance and enhances healthcare logistics in complex operational settings.

The discussion of deployment challenges highlighted the importance of high-quality data, ongoing user support, and phased implementation. Taken together, the evidence confirms that digital transformation in blood management is both feasible and impactful when guided by a strategic, data-driven approach.

The next and final chapter concludes the thesis by summarizing the key contributions, reflecting on the initial research goals, and suggesting future work directions.

5.9 Conclusion

The integration of AI-powered modules significantly enhanced the efficiency, accuracy, and reliability of the blood donation process. The forecasting, classification, and scheduling components worked in tandem to streamline operations, optimize donor engagement, and support informed decision-making. The next chapter summarizes the main findings of this research and outlines future work to expand and refine the smart platform in diverse healthcare settings.

CHAPTER 6

Conclusion and Future Work

This chapter concludes the research work presented in this thesis. It builds upon the implementation and evaluation discussed in previous chapters and summarizes the main findings and contributions of the proposed Smart Platform for Blood Management in Healthcare.

The chapter revisits the research objectives, reflects on how they were addressed, outlines key insights, and discusses both limitations and directions for future research. These reflections also emphasize the broader implications of integrating AI/ML in healthcare logistics, especially in the context of blood supply management.

6.1 Restatement of Research Objectives

The initial objective of this research was to address critical challenges in Algeria's blood supply chain—namely demand uncertainty, blood wastage, and inefficiencies in collection and distribution.

To this end, the platform was designed to:

- Digitize the full blood supply chain, from donation to distribution.
- Forecast blood demand using time series models.
- Classify donor behavior using machine learning techniques.
- Optimize donor appointment scheduling.
- Support national planning through real-time dashboards and automated reporting.

These objectives aimed to produce both academic contributions and practical improvements in healthcare logistics.

6.2 Summary of Key Findings

The research led to several important findings:

- **Forecasting Accuracy:** Time series models (SARIMA, Holt-Winters) enabled precise forecasting, improving planning and reducing wastage.
- **Donor Classification:** Machine learning models (Random Forest, ANN) effectively predicted return behavior, supporting engagement strategies.
- **Scheduling Optimization:** Appointment cancellations and deferrals were significantly reduced via intelligent scheduling tools.
- **Operational Improvements:** Centers observed improved collection volumes, lower positive serology rates, and enhanced inventory turnover.
- **Strategic Visibility:** Dashboards and reporting tools supported faster and more informed national-level planning.

These outcomes validate the platform's design and the hypothesis that AI/ML can enhance the performance and resilience of blood supply chains.

6.3 Research Contributions

This work contributed:

- **A Full-Stack Smart Platform:** Covering the entire operational and analytical cycle of the blood supply chain.
- **Novel Use of AI/ML:** Applied time series and classification models to blood forecasting and donor behavior—still underexplored in current literature.
- **Optimized Scheduling:** Introduced a data-driven model to manage donor appointments more efficiently.
- **Strategic Decision Support:** Real-time monitoring and reporting enhanced visibility for healthcare planners.
- **Digital Health Advancement:** Contributed to digital transformation in Algerian public health logistics.

6.4 Research Limitations

Several limitations were encountered:

- **Data Quality:** Incomplete records limited training data and model reliability.
- **Limited Scope:** Deployment was limited to one center, requiring broader testing for generalization.
- **Infrastructure Issues:** Some remote areas lacked the connectivity needed for seamless synchronization.
- **User Resistance:** Initial hesitance toward digital tools required significant training efforts.
- **Model Generalizability:** AI models were trained on local data; performance may vary in other demographics.

6.5 Recommendations for Future Work

Future research directions include:

- Nationwide deployment and evaluation across diverse regions.
- Integration with hospital information systems (HIS) and transfusion databases.
- Use of advanced models (e.g., Long short-term memory (LSTM)) for deeper predictive capability.
- Inclusion of contextual variables (holidays, weather, epidemics) in forecasting models.
- Development of a donor mobile app for scheduling and engagement.
- Real-time analytics and alert systems for proactive management.
- Improved data privacy and cybersecurity compliance.

6.6 Final Remarks

This thesis demonstrated the practical potential of integrating artificial intelligence into national blood supply chains through the development of a smart donor management platform. Despite facing challenges related to infrastructure and user adaptation, the system showed tangible improvements in forecasting, planning, and reporting.

These findings offer value both academically and practically. They also serve as a foundation for future research and innovation in digital health systems, particularly in resource-constrained environments. With continued development, platforms like this can play a critical role in building more resilient and responsive healthcare infrastructures. Moving forward, the continued refinement and scaling of such platforms may serve as a blueprint for broader AI adoption in public health across developing countries.

REFERENCES

Adarang, H., Bozorgi-Amiri, A., Khalili-Damghani, K. and Tavakkoli-Moghaddam, R. (2020), 'A robust bi-objective location-routing model for providing emergency medical services', *Journal of Humanitarian Logistics and Supply Chain Management* **10**(3), 285–319.

Agac, G., Baki, B., Ar, I. M. and Kahraman, H. T. (2023), 'A supply chain network design for blood and its products using genetic algorithm: A case study of turkey', *Journal of Industrial and Management Optimization* **19**(7), 5407–5446.

Aghsami, A., Abazari, S. R., Bakhshi, A., Yazdani, M. A., Jolai, S. and Jolai, F. (2023), 'A meta-heuristic optimization for a novel mathematical model for minimizing costs and maximizing donor satisfaction in blood supply chains with finite capacity queueing systems', *Healthcare Analytics* p. 100136.

Ahmadimanesh, M., Safabakhsh, H. R. and Sadeghi, S. (2023), 'Designing an optimal model of blood logistics management with the possibility of return in the three-level blood transfusion network', *BMC Health Services Research* **23**(1), 1304.

Alabdulkarim, L., Khorshid, M. and Raahemi, B. (2019), 'Decision support system for blood bank operations: a systematic review', *Journal of Medical Systems* **43**(5), 124.

Alghamdi, S. Y. (2023), 'A review of blood delivery for sustainable supply chain management (bscm)', *Sustainability* **15**(3), 2757.

Ayer, T., Zhang, C., Zeng, C., White III, C. C., Joseph, V. R., Deck, M., Lee, K., Moroney, D. and Ozkaynak, Z. (2018), 'American red cross uses analytics-based methods to improve blood-collection operations', *Interfaces* **48**(1), 24–34.

Cagliano, A. C., Grimaldi, S. and Rafele, C. (2021), 'A structured approach to analyse logistics risks in the blood transfusion process', *Journal of Healthcare Risk Management* **41**(2), 18–30.

Cagliano, A. C., Grimaldi, S., Rafele, C. and Campanale, C. (2022), 'An enhanced framework for blood supply chain risk management', *Sustainable Futures* **4**, 100091.

Casal-Guisande, M., Ceide-Sandoval, L., Mosteiro-Añón, M., Torres-Durán, M., Cerqueiro-Pequeño, J., Bouza-Rodr´ıguez, J.-B., Fernández-Villar, A. and Comesaña-Campos, A. (2023), 'Design of an intelligent decision support system applied to the diagnosis of obstructive sleep apnea', *Diagnostics* **13**(11), 1854.

Davoodi, S. M. R. and Goli, A. (2019), 'An integrated disaster relief model based on covering tour using hybrid benders decomposition and variable neighborhood search: Application in the iranian context', *Computers and Industrial Engineering* **130**, 370–380.

El-Amine, H., Bish, E. K. and Bish, D. R. (2018), 'Robust postdonation blood screening under prevalence rate uncertainty', *Operations Research* **66**(1), 1–17.

Epifani, I., Lanzarone, E. and Guglielmi, A. (2023), 'Predicting donations and profiling donors in a blood collection center: a bayesian approach', *Flexible Services and Manufacturing Journal* pp. 1–24.

Ganesh, K. and Narendran, T. (2007), 'Clash: a heuristic to solve vehicle routing problems with delivery, pick-up and time windows', *International Journal of Services and Operations Management* **3**(4), 460–477.

Ghaderzadeh, M. and Asadi, F. (2024), 'Barriers and challenges in ai-based healthcare in developing countries: a scoping review', *BMC Medical Informatics and Decision Making* **24**(1), 21.
URL: *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11010755/*

Ghaderzadeh, M., Asadi, F. and Abolghasemi, H. (2023), 'Hybrid deep learning model for forecasting blood demand in pandemic disruptions', *Journal of Biomedical Informatics* **138**, 104338.

Hashim, S. A., Al-Madani, A. M., Al-Amri, S. M., Al-Ghamdi, A. M. and Nahla, B. S. B. (2014), 'Online blood donation reservation and management system in jeddah', *Life Science Journal* **11**(8), 60–65.

Hosseini, S. M. H., Behroozi, F. and Sana, S. S. (2023), 'Multi-objective optimization model for blood supply chain network design considering cost of shortage and substitution in disaster', *RAIRO-Operations Research* **57**(1), 59–85.

Kauten, C., Gupta, A., Qin, X. and Richey, G. (2021), 'Predicting blood donors using machine learning techniques', *Information Systems Frontiers* pp. 1–16.

Kauten, C., Gupta, A., Qin, X. and Richey, G. (2022), 'Predicting blood donors using machine learning techniques', *Information Systems Frontiers* pp. 1–16.

Kayvanfar, V., Elomri, A., Kerbache, L., Vandchali, H. R. and El Omri, A. (2024), ‘A review of decision support systems in the internet of things and supply chain and logistics using web content mining’, *Supply Chain Analytics* p. 100063.

Kurani, A., Doshi, P., Vakharia, A. and Shah, M. (2023), ‘A comprehensive comparative study of artificial neural network (ann) and support vector machines (svm) on stock forecasting’, *Annals of Data Science* **10**(1), 183–208.

Kwon, H. J., Park, S., Park, Y. H., Baik, S. M. and Park, D. J. (2024), ‘Development of blood demand prediction model using artificial intelligence based on national public big data’, *Digital Health* **10**, 20552076231224245.

Li, N., Chiang, F., Down, D. G. and Heddle, N. M. (2021), ‘A decision integration strategy for short-term demand forecasting and ordering for red blood cell components’, *Operations Research for Health Care* **29**, 100290.

Liu, Q., Tao, Z., Tse, Y. and Wang, C. (2022), ‘Stock market prediction with deep learning: The case of china’, *Finance Research Letters* **46**, 102209.

Mansur, A., Wangsa, I. D., Rizky, N. and Vanany, I. (2025), ‘An efficient blood supply chain network model with multiple echelons for managing outdated products’, *Healthcare Analytics* **7**, 100377.

Meneses, M., Marques, I. and Barbosa-Póvoa, A. (2023), ‘Blood inventory management: Ordering policies for hospital blood banks under uncertainty’, *International Transactions in Operational Research* **30**(1), 273–301.

Meneses, M., Santos, D. and Barbosa-Póvoa, A. (2023), ‘Modelling the blood supply chain’, *European Journal of Operational Research* **307**(2), 499–518.
URL: *https://www.sciencedirect.com/science/article/pii/S0377221722004611*

Nagurney, A. and Dutta, P. (2019), ‘Supply chain network competition among blood service organizations: a generalized nash equilibrium framework’, *Annals of Operations Research* **275**(2), 551–586.

NBA (2023), ‘National blood authority, overview: Blood systems’. Last checked on Nov 01, 2013.
URL: *https://www.blood.gov.au/blood-systems*

Paul, J. A. and Wang, X. J. (2019), ‘Robust location-allocation network design for earthquake preparedness’, *Transportation research part B: methodological* **119**, 139–155.

Pirabán, A., Guerrero, W. J. and Labadie, N. (2019), ‘Survey on blood supply chain management: Models and methods’, *Computers and Operations Research* **112**, 104756.

Rahiminia, M., Shahrabifarahani, S., Alipour-Vaezi, M., Aghsami, A. and Jolai, F.

(2025), ‘A novel data-driven patient and medical waste queueing-inventory system under pandemic: A real-life case study’, *International Journal of Production Research* **63**(2), 418–434.

Ramadhan, M. N. S., Amyus, A., Fajar, A. N., Sfenrianto, S., Kanz, A. F. and Mufaqih, M. S. (2019), Blood bank information system based on cloud computing in indonesia, *in* ‘Journal of Physics: Conference Series’, Vol. 1179, IOP Publishing, p. 012028.

Rezaei, A., Bani-Salameh, H. and Ahmadi, M. (2023), ‘Comparative study of classical and deep learning models for blood demand forecasting’, *Health Informatics Journal* **29**(4), 14604582231192011.

Rezaei Kallaj, M., Abolghasemian, M., Moradi Pirbalouti, S., Sabk Ara, M. and Pourghader Chobar, A. (2021), ‘Vehicle routing problem in relief supply under a crisis condition considering blood types’, *Mathematical Problems in Engineering* **2021**, 1–10.

Salem, F., Youssef, A. and El Sayed, T. (2024), ‘Bridging the ai divide: challenges and strategies for inclusive adoption in the global south’, *Humanities and Social Sciences Communications* **11**(1), 131.
URL: *https://www.nature.com/articles/s41599-024-03947-w*

Sarvestani, S. E., Hatam, N., Seif, M., Kasraian, L., Lari, F. S. and Bayati, M. (2022), ‘Forecasting blood demand for different blood groups in shiraz using auto regressive integrated moving average (arima) and artificial neural network (ann) and a hybrid approaches’, *Scientific Reports* **12**(1), 22031.

Shokouhifar, M. and Ranjbarimesan, M. (2022), ‘Multivariate time-series blood donation/demand forecasting for resilient supply chain management during covid-19 pandemic’, *Cleaner Logistics and Supply Chain* **5**, 100078.

Sulaiman, S., Hamid, A. A. K. A. and Yusri, N. A. N. (2015), ‘Development of a blood bank management system’, *Procedia-Social and Behavioral Sciences* **195**, 2008–2013.

Twumasi, C. and Twumasi, J. (2022), ‘Machine learning algorithms for forecasting and backcasting blood demand data with missing values and outliers: A study of tema general hospital of ghana’, *International Journal of Forecasting* **38**(3), 1258–1277.

Wong, S., Tang, H. and de Steiger, R. (2015), ‘Blood management in total hip replacement: an analysis of factors associated with allogenic blood transfusion’, *ANZ journal of surgery* **85**(6), 461–465.

Wu, H.-y., Li, Z.-g., Sun, X.-k., Bai, W.-m., Wang, A.-d., Ma, Y.-c., Diao, R.-h., Fan, E.-y., Zhao, F., Liu, Y.-q. et al. (2022), ‘Predicting willingness to donate blood based on machine learning: two blood donor recruitments during covid-19 outbreaks', *Scientific Reports* **12**(1), 19165.