

ملخص:

عملنا المنجز في إطار هذه المذكرة يتعلق بتصميم برنامج ذكي يسمح بمتابعة عملية التصنيف الموضوعي لنصوص مكتوبة بلغات متعددة اعتمادا على مجموعة من الأصناف المعرّفة مسبقا (سياسة، اقتصاد، فن ، رياضة) ، وهذا من خلال استخدام خوارزمية التصنيف المعروفة باسم شجرة القرار. **الكلمات المفتاحية :** التصنيف الآلي للنصوص، المعالجة الآلية للغات المختلفة ،خوارزمية التصنيف شجرة القرار.

Abstract:

Our work realized in this memory consists in carrying out of an intelligent system which permit to follow step by step the process of thematic categorization of multilingual texts in a set of categories which are predefined in the beginning (Politic, Economic, Art and Sport) by using of one of the most known algorithms in the field of classification which is the decision tree.

Key words: Supervised classification, Training Algorithms, Multilanguage TALN, Decision tree.

Résumé:

Notre travail réalisé dans le cadre de ce mémoire consiste à réaliser un système intelligent permettant de suivre pas à pas le processus de catégorisation thématique des textes multilingues dans un ensemble de catégories prédéfinies au départ (Politique, Economie, Art et Sport) en utilisant l'un des algorithmes de classification les plus connus dans le domaine qui est l'arbre de decision.

Mots clés: Classification supervisée, Algorithme d'apprentissage, TALN multilingue, Arbres de décision.