

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
ÉCOLE SUPÉRIEURE EN INFORMATIQUE
-08 MAI 1945- SIDI BEL ABBES



MÉMOIRE DE FIN D'ÉTUDE

En vue de l'obtention du diplôme d'ingénieur d'état

Filière : **Informatique**

Spécialité : **Système d'Information et Web (SIW)**

Ooredoo Rayek : Un system d'analyse des sentiments des clients
de l'opérateur téléphonique Ooredoo

Présenté par :

- Mr Bouchiha Djelloul
- Mr Zendagui Iyas

Soutenu le : **23/07/2019.**

Devant le jury composé de :

- | | | |
|-------------------------------|--------------------------------|--------------|
| - Mr KECHAR Mohamed | Maître de conférences classe B | Président |
| - Mr BENSLIMANE Sidi Mohammed | Professeur | Encadreur |
| - Mme KLOUCHE Badia | Maître Assistante Classe A | Examinatrice |
| - Mme BENNABI Sakina Rim | Maître Assistante Classe A | Examinatrice |

Année Universitaire : 2018 / 2019

Résumé

L'analyse des sentiments (AS) est un domaine de recherche permanent qui se situe au carrefour de nombreux domaines, tels que la fouille des données, le traitement du langage naturel et l'apprentissage automatique. Il concerne l'extraction automatique des opinions exprimées dans un texte donné. En raison de ses vastes applications, de nombreuses études ont été menées dans le domaine de l'Opinion Mining, en particulier dans les textes en arabe algérien dialectal, tandis que d'autres langues utilisées, telles que l'arabe, le français et l'anglais, ont fait l'objet d'une attention amoindrie. Les opérateurs sont exigeants et prêtent une grande importance à l'opinion de leur clientèle, qui utilise un langage arabe dialectal dans les différents réseaux sociaux, tout en introduisant les sentiments dans leur plateforme respective et leur enjeu stratégique.

A l'instar des autres opérateurs téléphoniques, Ooredoo, dans sa nouvelle stratégie, se déploie à conquérir de nouveaux clients, à ajouter à sa clientèle habituelle, pour exploiter leurs opinions à travers une analyse des sentiments. C'est à cet égard, que s'inscrit ce travail, dont l'objet est de mettre en place un système chez Ooredoo, qui permet de collecter, stocker, exploiter et classer les données textuelles exprimées par ses clients.

Notre approche d'analyse des sentiments est basée sur des algorithmes d'apprentissage machine (Approche supervisée), il a pour but d'établir des règles de comportement à partir d'une base de données. La base de données est en principe un ensemble de couples entrées/sorties. Le but est d'apprendre à prédire pour toute nouvelle entrée X , la sortie Y (Positive, Negative ou Neutre). Les algorithmes et techniques utilisés pour la classification sont la machine à vecteurs de support (SVM) et Naïve Bayes (NB).

Nous présentons une approche hybride pour la translittération de l'arabizi algérien. Nous avons élaboré un ensemble de règles permettant le passage de l'arabizi vers l'arabe. À partir de ces règles nous générons un ensemble de candidats pour la translittération de chaque mot en arabizi vers l'arabe, et un parmi ces candidats sera ensuite identifié et extrait comme le meilleur candidat.

Mots Clés : Analyse des sentiments, apprentissage automatique, classification du texte, extraction d'opinion et des sentiments, arabe dialectal, Arabizi.

Abstract

Sentiment Analysis (SA) is an ongoing area of research at the crossroads of many areas, such as data mining, natural language processing and machine learning. It concerns the automatic extraction of opinions expressed in a given text. Due to its wide-ranging applications, many studies have been conducted in the field of Opinion Mining, particularly in texts in Algerian Arabic dialect, while other languages used, such as Arabic, French and Spanish. English has received less attention. Operators are demanding and attach great importance to the opinion of their customers, who use Arabic dialect language in the various social networks, while introducing the feelings in their respective platform and their strategic stake.

Like the other telephone operators, Ooredoo, in its new strategy, deploys to conquer new customers, to add to its usual clientele, to exploit their opinions through an analysis of feelings.

It is in this respect, that this work, whose object is to set up a system at Ooredoo, which allows to collect, store, exploit and classify the textual data expressed by its customers.

Our approach of sentiment analysis is based on machine learning algorithms (supervised approach), it aims to establish rules of behavior from a database. The database is in principle a set of input / output pairs. The goal is to learn to predict for any new input X, the output Y (Positive, Negative or Neutral).

We present a hybrid approach for the transliteration of Algerian arabizi. We have developed a set of rules allowing the transition from arabizi to Arabic. From these rules we generate a set of candidates for the transliteration of each arabizi word to Arabic, and one of these candidates will then be identified and extracted as the best candidate.

Key Words : Sentiment Analysis, Machine Learning, Text Classification, Opinion and Sentiment Extraction, Dialectal Arabic, Arabizi.