

الجمهورية الشعبية الديمقراطية الجزائرية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي .08 ماي 1945. بسيدي بلعباس
École Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès



THESIS

To obtain the diploma of **Master**
Field: **Computer Science**
Specialty: **Ingénierie des Systèmes Informatiques (ISI)**

Theme

**Sentiment analysis in Algerian dialects considering
Code-switching**

Presented by:
ZELLAGUI Iskander

Submission Date: **July, 2023**
In front of the jury composed of:

Mrs. KLOUCHE Badia	President
Mrs. TAOULI Amina	Examiner
Mrs. ALIANE Hassina	Supervisor
Mr. KHALDI Belkacem	Co-Supervisor
Mr. ALIANE Ahmed Amine	Co-Supervisor

Academic Year : 2022/2023

Acknowledgement

First and before all, Alhamdulillah, all praises to Allah, for helping, guiding and giving us the patient and motivation to complete this humble work.

We offer our sincerest gratitude and thanks to our supervisor Dr ALIANE Hassina, the Head of Natural Language Processing and Digital Content team, Dr KHALDI Belkacem, and ALIANE Amine, for helping us during this journey, and guiding us through this study with their knowledge, support, patience, and motivation.

It was a great privilege and honor to work under their guidance.

This work could not have been achieved without them.

We would also like to offer our thanks to the jury members, for spending their precious time reviewing our work.

We would like to express our sincere thanks and appreciation to both our parents and siblings, for their support and encouragement throughout our academic journey. We would like to extend our thanks to our friends, especially BENSALAH Amdjed, NASRI Mohammed, and BOUCHOUKA Hidaya, for the great help they have assisted us with.

Last but not least, we would like to thank all the professors and staff of ESI-SBA for all their efforts.

Abstract

Code-switching is the act of using two or more languages in the same text due to cultural or contextual reasons.

The code-switching property of text is an essential step to sentiment analysis, that piqued the interest of many researchers. Thus, the study of sentiment analysis when code-switching is involved have gained increased popularity, especially with the continuous rise of social media users and content that has more and more become integrated with our daily lives.

The recent approaches used to deal with code-switching mostly depend on lexicons, and ready-to-use corpora, which makes it different and more difficult for the Arabic language, and its dialects, that use large and varied dialects and lexicons.

In our research, we aim to address the problem of Arabic sentiment analysis when code-switching is involved, and more precisely in the Algerian dialect. Many techniques have been used in previous researches including machine learning, and deep learning ones, in addition to getting aid form various NLP tools that were used to ease up the task.

Our main objective is comparing various researches that has tried to deal with this challenge until today, by proposing different approaches to solve the problem at hand both in English and Arabic Languages using machine learning and deep learning. And after, we give a summary and compare between these methods.

Keywords: Sentiment Analysis, Social Media, Deep learning, Algerian Dialects, Code-switching, Arabic language.