



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 بعباس
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

**Sarcasm Detection in Arabic Language Twitter Posts with
Deep Learning Approaches.**

Presented by :
Touhami Wided Ahlem
Taleb Douaa

Submission Date : **4th July , 2022** In front of the jury composed of :

Mr. KESKES Nabil
Mme. ALIANE Hassina (CERIST)
Mr. KHALDI Belkacem
Mr. ALIANE Ahmed Amine (CERIST)
Mme. Klouche Badia

President
Supervisor
Supervisor
Co-Supervisor
Examiner

Academic Year : 2021/2022

Abstract

Sarcasm is a subtle form of language in which individuals say the opposite of what is implied. It is generally used to express negative sentiments that can easily cause a serious emotional impact on the recipient like depression and low self-esteem. Sarcasm is a crucial step to sentiment analysis, that presents a real challenge even for humans. Therefore, Automatic Sarcasm detection becomes a necessity owing to the increase in popularity of social networking sites and has come very helpfully to improving sentiment analysis, being proved their direct relationship recently. The existing methods for automatic sarcasm detection rely primarily on lexical and linguistic cues. Therefore, the situation is not different for those who use the Arabic language, since there is a big variety and contrast in dialects and lexical. In the case of our study, we aim to address the difficult task of sarcasm detection, numerous techniques including Machine learning, Deep learning, and NLP tools were used to identify and reduce this problem on social media platforms.

The objective of this work is to compare a set of researches in this task up to the moment that proposes approaches for Sarcasm detection in both English and Arabic Language Twitter Posts with Machine Learning and Deep Learning. Then we give a summary of some methods and a comparative table of them.

Keywords: Sarcasm Detection, Sentiment Analysis , Natural Language Processing, Machine learning, Deep learning, Transformers.

Résumé

Le sarcasme est une forme subtile de langage dans laquelle les individus disent le contraire de ce qui est sous-entendu. Il est généralement utilisé pour exprimer des sentiments négatifs qui peuvent facilement avoir un impact émotionnel grave sur le destinataire, comme la dépression et une faible estime de soi. Le sarcasme est une étape cruciale de l'analyse des sentiments, qui présente un véritable défi même pour les humains. Par conséquent, la détection automatique du sarcasme devient une nécessité en raison de l'augmentation de la popularité des sites de réseaux sociaux et s'est avérée très utile pour améliorer l'analyse des sentiments, ayant récemment prouvé leur relation directe. Les méthodes existantes de détection automatique du sarcasme reposent principalement sur des indices lexicaux et linguistiques. Par conséquent, la situation n'est pas différente pour ceux qui utilisent la langue arabe, car il existe une grande variété et un contraste dans les dialectes et le lexique. Dans le cas de notre étude, nous visons à aborder la tâche difficile de détection du sarcasme, de nombreuses techniques, notamment l'apprentissage automatique, l'apprentissage en profondeur et les outils NLP, ont été utilisées pour identifier et réduire ce problème sur les plateformes de médias sociaux.

L'objectif de ce travail est de comparer un ensemble de recherches dans cette tâche jusqu'à présent qui proposent des approches pour la détection du sarcasme dans les publications Twitter en anglais et en arabe avec l'apprentissage automatique et l'apprentissage en profondeur. Ensuite, nous donnons un résumé de certaines méthodes et un tableau comparatif de celles-ci.

Mot Clé: Détection de sarcasme, analyse des sentiments, traitement du langage naturel, apprentissage automatique, apprentissage en profondeur, Transformeurs.