Machine Learning-Based Model for Sentiment and Sarcasm Detection

Aya H. Allam[1], Hanya M. Abdallah[1], Eslam Amer[2], Hamada A. Nayel[1]
[1]Department of Computer Science
Faculty of Computers and Artificial Intelligence, Benha University
[2]Faculty of Computer Science, Misr International University

{aya.allam,hanya.abdelhak,hamada.ali}@fci.bu.edu.eg,
eslam.amer@miuegypt.edu.eg

Abstract

Within the last few years, the number of Arabic internet users and Arabic online content is in exponential growth. Dealing with Arabic datasets and the usage of non-explicit sentences to express an opinion are considered to be the major challenges in the field of natural language processing. Hence, sarcasm and sentiment analysis has gained a major interest from the research community, especially in this language. Automatic sarcasm detection and sentiment analysis can be applied using three approaches, namely supervised, unsupervised and hybrid approach. In this paper, a model based on a supervised machine learning algorithm called Support Vector Machine (SVM) has been used for this process. The proposed model has been evaluated using ArSarcasm-v2 dataset. The performance of the proposed model has been compared with other models submitted to sentiment analysis and sarcasm detection shared task.

1 Introduction

Today, social networks such as Facebook and Twitter are considered new forms of social interaction, knowledge exchange, and collaboration. Social networking sites allow users to exchange ideas, post updates, reviews, and comments while sharing their broader interests. Du to the extreme usage of social networks, it is crucial to apply related analysis tasks automatically. These tasks ranges over sarcasm detection, offensive language detection(Nayel and L, 2019; Nayel, 2020), author profiling (Nayel, 2019), irony detection (Nayel et al., 2019) and sentiment analysis. Sentiment analysis is precious as it allows us to capture an overview of social network users' opinions towards specific topics, products, or services. Sentiment analysis refers to the use of Natural Language Processing (NLP) and Machine Learning (ML) to identify and extract subjective information in a piece of writing. Sarcasm is defined as an indirect way to express an opinion as the written is literally not the intended meaning (Wilson, 2006). It is considered a widespread phenomenon in social media, which is difficult to analyse, not just automatically but also for humans (Maynard and Greenwood, 2014). Although It has a significant effect on sentiment, it is ignored in social media analysis because it is considered too difficult to handle. So, sarcasm is still considered a challenge in sentiment analysis.

Due to the importance of sentiment analysis, many research works have been developed in this research field. However, most of these studies have focused on English and other Indo-European languages (Lunando and Purwarianti, 2013). Actually, few studies have been conducted on sarcasm and sentiment analysis in morphologically rich languages such as Arabic (Boudad et al., 2018). Three main variations of the Arabic language exist classical Arabic, which is the language of the Qur'an (Islam's Holy Book); Modern Standard Arabic (MSA), and dialectical Arabic. MSA is an Arabic language form used greatly in formal speeches and writings. Dialectical Arabic refers to all oral varieties spoken in daily communication. These vary from one Arab country to another and from one region of the same country to another.

The work in Arabic Sarcasm Sentiment analysis using a variety of Arabic dialects started in 2008. Besides, a set of research work has been rapidly conducted such as (Abdul-Mageed et al., 2011; Al-Ghadhban et al., 2017; Alsayat and Elmitwally, 2020; Karoui et al., 2017). The authors use different ML approaches such as supervised, unsupervised, and hybrid approaches to detect sarcasm and sentiment. These approaches carry out the classification task as it classifies text into either sarcasm or not. However, more work is still needed to achieve better accuracy.

Proceedings of the Sixth Arabic Natural Language Processing Workshop, pages 386–389
Kyiv, Ukraine (Virtual), April 19, 2021.

In this paper, we present a model to detect sarcastic and sentiment Arabic tweets from real ones using SVM. The model is developed in the framework of the shared task on sarcasm and sentiment detection in Arabic (Abu Farha et al., 2021). The rest of this paper is organized as follows. Section 2 reviews the dataset, section 3 gives a detailed description of how the systems were built and trained. In Section 4, the system is evaluated, and the results are discussed. The paper is concluded in Section 6 after a general discussion is introduced in section 5.

2 Dataset

For both tasks, ArSarcasm-v2 (Abu Farha et al., 2021) a dataset has been distributed to all participants. It contains a labeled training set with 12,548 tweets and a blind test set contains 3,000 unlabelled tweets. Table 1 contains a statistics of the training set for both tasks. In addition to the two labels for both tasks, dialect information about a tweet is also given.

3 System

The proposed system, as described in Figure 1, combines three main phases: preprocessing phase, the feature extraction phase, and the classification phase.

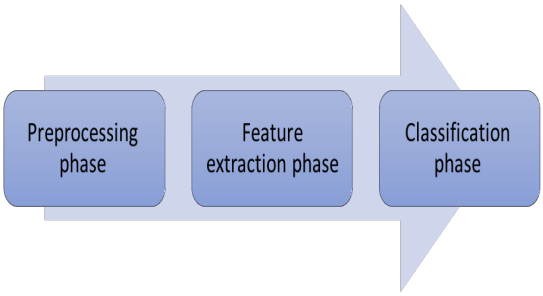

Figure 1: Proposed system structure

3.1 Preprocessing Phase

In social media communications, users use slang words instead of formal words such as emojis, numbers, etc. Hence, it is a must to clean and prepare data for the feature extraction process. At first, We changed the encoding of the dataset file to UTF-8 to recognize Arabic characters. Then, We preprocessed the tweets by removing the hashtags, emojis, numbers, and URLs. This process reduces the dimensionality of the problem low. Hence, it speeds up the classification process and makes it more accessible.

3.2 Feature Extraction Phase

In this phase, unigram or bigram can be used as features for the Arabic tweets. Several methods can be used to determine the significance of each feature in the whole text, such as Term Frequency (TF) or and Term Frequency Inverse Document Frequency (TF/IDF).

For term weights, we used TF/IDF calculated as follows:-

$$t_j = tf_j * \log\left(\frac{N+1}{df_j+1}\right)$$

where tf_j is the total number of occurrences of term t_j in the given tweet, df_j is the number of tweets in which the token/term t_j occurs and N is the total number of tweets.

3.3 Classification Phase

Sentiment Analysis can be classified into three levels: document level, sentence level and aspect level. In this phase, the document level is adopted where the whole piece of writing is dealt with as one unit and is assigned to a particular class label. SVM is a linear classifier that have been designed binary classification problem (Theodoridis and Koutroumbas, 2009). It can be adjusted for multi class classification problems using one-vs-one or one-vs-many strategies. SVM is one of the strongest and accurate algorithms used with Arabic text (Nayel, 2020, 2019; Nayel et al., 2019).

4 Experiments and Results

4.1 Evaluation Metric

The performance evaluation of sarcasm and sentiment detection shared task is measured as the accuracy of the system in addition to class-wise accuracy which is calculated using Precision (**P**), Recall (**R**) and **F1** measure [1]. For each class, **P** is the number of tweets correctly classified over the total number of tweets that system classified as same class. **R** is the number of tweets correctly classified over the actual number of tweets of the class. **F1** measure is the harmonic mean of **P** and **R**, which can be calculate as follow:-

$$F1 = \frac{2 * P * R}{P + R}$$

[1] http://www.nltk.org/_modules/nltk/metrics/scores.html

Task	Class	No. of Tweets	Percentage (%)
Sarcasm Detection	TRUE	2,168	17.28%
	FALSE	10,380	82.72%
Sentiment Analysis	POS	2,180	17.37%
	NEG	4,621	36.83%
	NEU	5,747	45.80%

Table 1: Statistics of training set for both tasks

The official metric of the sarcasm detection task is F1-score of the sarcastic class (F1-sarcastic), while macro average of the F-score of the positive and negative classes (F-PN) will be the official metric of the sentiment detection task.

4.2 Results

In developing phase, cross-validation approach with five folds has been used to train different classification algorithms for each subtask. These algorithms are: SVM, Linear Regression (LR), Naïve Bayes (NB), Complementary Naïve Bayes (CNB) and Stochastic Gradient Descent (SGD). The performance of all classifiers during development phase are given in Table 2.

	Sentiment	Sarcasm
SVM	**85.55**%	**84.22**%
CNB	83.47%	82.97%
NB	78.92%	78.10%
LR	65.89%	64.03%

Table 2: 5-fold Cross-Validation performance of different classifiers for both tasks

From Table 2 it is clear that SVM outperforms all other algorithms in both tasks.

4.2.1 Official Submission Results

The proposed model has been applied to the blind test set and the output of both tasks have been submitted to the shared task. Table 3 shows the detailed results of the proposed model for both subtasks.

	Sarcasm	Sentiment
Accuracy	0.7460	0.5980
Macro-F1	0.5457	0.5291
Precision	0.7048	0.5434
Recall	0.5602	0.5207
Oficial Metric	F1-sarcastic	F-PN
	0.5936	0.2440

Table 3: The official results of our submissions

5 Discussion

Sarcasm and sentiment detection in Arabic is a widely NLP studied problem. The supervised learning approach outperforms unsupervised and semi-supervised methods in this detection process. Note also that some supervised classifiers such as SVM have repeatedly been applied. The effectiveness of this classifier is the reason behind this choice. The same thing can be said about the used feature; n-grams are frequently chosen as features.

6 Conclusion

Sarcasm and sentiment detection is one of the most useful and challenging tasks. In this paper, we proposed SVM based classifier model to detect Arabic sarcasm and sentiment in Twitter. The results are produced from training the model using SVM classifier, one of the accurate classification algorithms. The model achieves moderate accuracies, which are relatively high, especially regarding Arabic text. The performance can be improved using different weighting scores such as word embeddings. Deep learning based models also can be used for the training purpose.

References

Muhammad Abdul-Mageed, Mona Diab, and Mohammed Korayem. 2011. Subjectivity and sentiment analysis of Modern Standard Arabic. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 587–591, Portland, Oregon, USA. Association for Computational Linguistics.

Ibrahim Abu Farha, Wajdi Zaghouani, and Walid Magdy. 2021. Overview of the WANLP 2021 Shared Task on Sarcasm and Sentiment Detection in Arabic. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*.

D. Al-Ghadhban, E. Alnkhilan, L. Tatwany, and M. Alrazgan. 2017. Arabic sarcasm detection in twitter. In *2017 International Conference on Engineering MIS (ICEMIS)*, pages 1–7.

Ahmed Alsayat and Nouh Elmitwally. 2020. A comprehensive study for arabic sentiment analysis (challenges and applications). *Egyptian Informatics Journal*, 21(1):7–12.

Naaima Boudad, Rdouan Faizi, Rachid Oulad Haj Thami, and Raddouane Chiheb. 2018. Sentiment analysis in arabic: A review of the literature. *Ain Shams Engineering Journal*, 9(4):2479–2490.

Jihen Karoui, Farah Banamara Zitoune, and Véronique Moriceau. 2017. Soukhria: Towards an irony detection system for arabic in social media. *Procedia Computer Science*, 117:161–168. Arabic Computational Linguistics.

E. Lunando and A. Purwarianti. 2013. Indonesian social media sentiment analysis with sarcasm detection. In *2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pages 195–198.

Diana Maynard and Mark Greenwood. 2014. Who cares about sarcastic tweets? investigating the impact of sarcasm on sentiment analysis. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 4238–4243, Reykjavik, Iceland. European Language Resources Association (ELRA).

Hamada Nayel. 2020. NAYEL at SemEval-2020 task 12: TF/IDF-based approach for automatic offensive language detection in Arabic tweets. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pages 2086–2089, Barcelona (online). International Committee for Computational Linguistics.

Hamada A. Nayel. 2019. NAYEL@APDA: Machine Learning Approach for Author Profiling and Deception Detection in Arabic Texts. In *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019*, volume 2517 of *CEUR Workshop Proceedings*, pages 92–99. CEUR-WS.org.

Hamada A. Nayel and Shashirekha H. L. 2019. DEEP at HASOC2019: A machine learning framework for hate speech and offensive language detection. In *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019*, volume 2517 of *CEUR Workshop Proceedings*, pages 336–343. CEUR-WS.org.

Hamada A. Nayel, Walaa Medhat, and Metwally Rashad. 2019. BENHA@IDAT: Improving Irony Detection in Arabic Tweets using Ensemble Approach. In *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019*, volume 2517 of *CEUR Workshop Proceedings*, pages 401–408. CEUR-WS.org.

Sergios Theodoridis and Konstantinos Koutroumbas. 2009. Chapter 3 - Linear Classifiers. In *Pattern Recognition (Fourth Edition)*, fourth edition edition, pages 91 – 150. Academic Press, Boston.

Deirdre Wilson. 2006. The pragmatics of verbal irony: Echo or pretence? *Lingua*, 116(10):1722–1743. Language in Mind: A Tribute to Neil Smith on the Occasion of his Retirement.