DOI: 10.15849/IJASCA.230720.16

Int. J. Advance Soft Compu. Appl, Vol. 15, No. 2, July 2023
Print ISSN: 2710-1274, Online ISSN: 2074-8523

A Hybrid Sentiment-Discourse Analysis Model for Ukraine Crisis Facebook Posts with a Jordanian Dialect

Riham Jaradat[a], Ghaith M. Jaradat[b,*], Mutasem Alsmadi[c], Malek Alzaqebah[d], Ibrahim Almarashdeh[c], Ahmad Althunibat[e], and Hassan Al-Tarawneh[f]

[a]Department of Modern Languages, Faculty of Arts, Yarmouk University, 655-21110, Irbid, Jordan
e-mail: rihamja@yu.edu.jo

[b,*]Department of Computer Science and Information Systems, Faculty of Computer Sciences and Informatics, Amman Arab University, 2234-11953, Amman, Jordan
e-mail: g.jaradat@aau.edu.jo

[c]Department of Department of Management Information Systems, College of Applied Studies and Community Service, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam, Saudi Arabia
e-mail: mkalsmadi@uod.edu.sa; iaalmarashdeh@iau.edu.sa

[d]Department of Mathematics, College of Science, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam, Saudi Arabia
e-mail: maafehaid@iau.edu.sa

[e]Department of Software Engineering, Faculty of Science and Information Technology, Al-Zaytoonah University of Jordan, P.O. Box 130-11733, Amman, Jordan

[f]Department of Data sciences and Artificial Intelligence, Al-Ahliyya Amman University, Amman, Jordan.
e-mail: h.altarawneh@ammanu.edu.jo

Abstract

Sentiment analysis has gained attention in computational linguistics, where algorithms are dedicated to recognizing sentiments to take better decisions in each context. However, this is difficult to achieve in social networks due to linguistic features. A text classification model of combined lexical-based and machine-learning techniques is presented in this paper; it uses discourse structure including its relations to improve sentiment analysis. The model aims at extracting information about opinion polarity posted by users; then modelling users' usual polarity to detect significant polarity. The model extracts posted texts on Facebook, then it classifies them according to their polarity, which shows the detection of polarity. It provides interesting results using Ukraine crisis argued by Jordanian news followers. Obtained results show a high accuracy (94.68%). It is useful to have information about users' sentiments. It may serve feedback for politicians and jurists, where direct contact with war victims is extremely difficult.

Keywords: *Discourse analysis; Sentiment analysis; Computational linguistics; Social Networks; Jordanian Facebook users; Ukrainian crisis.*

Received 5 April 2023; Accepted 25 June 2023

1 Introduction

Over the last decade, sentiment analysis (SA) has gained a great attention in computational linguistics, natural language processing (NLP), human computer interaction, and decision support systems, where algorithms are dedicated to recognizing sentiments to take better decisions in each context (e.g., [1], [2], [3]). Online social media has become a significant communication on the web. Hence, SA includes mining user's opinion which has a great social and commercial importance. SA is particularly useful for texts that need to carefully analyze the information spread through the Internet, or rather posted on social networks. For instance, industries that want to get feedback from their customers about their newly launched products for financial or service assessment and evaluation. However, it is known that the information reproduced through social networks is difficult to analyze accurately, since people use stereotypes, metaphors, writing styles, multi-word, and ironies, expressed in an informal language that is difficult to interpret.

According to [3] and [4] categorized SA into opinion mining (OM) of documents, sentiment classification of sentences, and polarity prediction of words. Recently, the effectiveness of discourse relations in SA has increasingly been recognized and approved. Where, in traditional lexicon-based methods, all words and sentences are treated equally, ignoring the structural aspects of a text. However, discourse structure knowledge is crucial to texts dedicated to polarity prediction. For example, Jordanian observers or news followers on social networks expressing their opinion about the Ukrainian crisis (#الحرب_الروسية_الأوكرانية) in a variety of news agencies' Facebook accounts: "*defiance and defending homeland is glorious, Russian precautious measures is also understandable, but warfare and offensive acts are destructive. In general, I think both sides should reconsider negotiation peacefully*."

By analyzing the discourse structure of a text is split into spans with different semantic relations. With this discourse knowledge, it is possible to assign text spans with different weights by their contribution to the overall sentiment of a document. From the given example, there are two positive words "glorious" and "understandable," and two negative words "offensive" and "warfare" which would give a neutral prediction using the lexicon-based method. However, spans introduced by connective "but" has higher degree of importance, denoting a contrast relation; and spans introduced by connective "in general" has the highest degree of importance, denoting a generalization relation. This leads to overall negative sentiment.

This paper exploits discourse relations using connectives and conditionals for sentiment classification of Jordanian-Arabic texts. Discourse analysis (DA) is a qualitative and interpretive methodology that has been employed to analyze different socio-cultural phenomena in an effective manner via a variety of communication media. In which, a combination of discourse relations with sentiment classification for Jordanian news followers of the Ukrainian crisis posted on social networks (e.g., Facebook). Hence, the contributions of this study are: (1) proposing a relatively complete discourse relation hierarchy and listing their corresponding connectives and conditionals in Jordanian-Arabic dialect and validate their effectiveness in SA; (2) conducting weighting schemes at various granularities of discourse relations; (3) concluding influential discourse relations that contribute most to the overall meaning of texts.

In this study, a collected corpus from a set of randomly chosen Facebook accounts for the case study. Starting by preprocessing the corpus and tag sentences using various levels of negativity and positivity to improve the accuracy of text classifications. These classifications are assessed by DA. Interesting results are discussed about the use of this combination to get helpful feedback for the analyzed case study and in general for any communication context.

This paper is organized as follows: Next sections are related work, proposed method, results, analysis, and discussion followed by conclusions and directions of future work.

2 Related Work

2.1 Discourse Analysis

DA describes the methodology of studying language and text that draws variously upon linguistics, literary theory and cultural studies, philosophy of language, sociology, and psychology [5]. According to [6], the aim of DA is to show hidden ideological and power relations embedded in text. In a similar fashion, the perspective of [7] texts can be written or spoken and must be described in linguistic terms and in terms of their intended meaning. Discourse, as text in context, is defined by its effect. Because DA perceives language as a social practice, it predetermines that it cannot function in isolation, but only within a cultural or social setting [8].

DA including critical analysis is an amalgamation of a variety of micro-sociological theories and theories on society and power based on Michael Foucault’s definition of power [9]. According to [10], the analysis is top-down, where analysts begin with their understanding of the content; or bottom-up, where the starting point is the linguistic detail. Practically, a combination of both is considered, where the analyst looks for signification and significance in sentences.

The study of [11] confirmed that the DA could be done by the critical DA method. One of the paradigms was from the well-known Teun A. van Dijk’s model. Van Dijk’s model can analyze the text, where the text can be separable from the context that covered it. They also revealed another fact that the text can affect the reader. Hence, the reader thinks and acts appropriately with the writer. By Van Dijk’s model, this text consisted of macrostructure, superstructure, and microstructure.

2.2 Sentiment analysis based on social media

SA is essential for different domains where it is crucial to know users’ public opinion about events, products, brands, politicians. Many studies have concentrated on English or Arabic texts including Twitter feeds and user reviews on hotels, movies, and products. On the other hand, Facebook, as an online social network, has attracted quite limited attention from the research community. However, a handful of works conducted studies and experiments on Arabic Facebook posts and more sparsely on the Arabic dialects, specifically the Jordanian dialect.

For example, Facebook ranks first amongst the present social media platforms [12]. A blog provides users the ability to exchange a variety of contents (e.g., posts or tags) such as short sentences, paragraphs, individual pages, images, or videos over the internet. One of the first studies of applying SA on social media platforms (e.g., micro-blogging websites), was provided by [13], where they described a distant supervision-based approach for sentiment classification.

Other social media platforms for SA Twitter and blogs have been intensively adopted due to their simplicity which has remained a primary focus of researchers (e.g., [14] [15] [16]) using the concepts and techniques of OM. While other research papers (e.g., [17] [18] [19]) considered sentiment lexicons for classification and twitter-specific features into positive and negative. Others also considered an ontology-based analysis (e.g., [20]) and managing imbalanced online scrapped data (e.g., [21] [22] [23]). On the other hand, Facebook has been less addressed [24]. In [24] a SA and OM framework for Facebook posts and comments analysis is presented. They used computational linguistics to measure sentiments of user's opinion about different entities. They evaluated their framework's performance on a randomly selected user case study of posts and comments with 85% accuracy compared to human judgment.

Similar works on analyzing sentiments via machine learning (ML) techniques in Facebook and its applications such as in the context of e-learning, were conducted by [25] with high accuracy (83.27%); a case of Turkish users conducted by [26] with a very high accuracy (91.6%) using deep learning techniques (e.g., neural networks); and [27] using ML techniques (e.g., decision tree). Hence, in this study, the framework of [24] is used to analyze Facebook posts and comments for opinions and sentiments of the Jordanian news followers.

Approaches used in literature for SA and OM are primarily ML based on lexicon with a supervised learning (e.g., [28] [29] [30]) where the text is compared to a human developed list of sentiment bearing words. Claimed by [24] that the literature on twitter hashtags used an overall score (positive, negative, or neutral) is assigned to the text based on the human designed list. This technique works better for short informal text (as in twitter tags) where people are less formal in using grammar, which is the case in the people comments on Facebook [24]. Another technique is based on proper grammatical checking on the text using various methods of NLP (e.g., [31]), which is mandatory for text where proper grammar has been used. Technically, it is a combination of principal component analysis for features reduction and supervised learning for prediction of opinions [32]. Other approaches are based on hybrid techniques that use combination both ML and NLP for SA and OM.

Therefore, our study focuses on the SA including OM with DA from the text of Facebook posts and comments carries formal as well as informal text expressions.

2.3 Discourse analysis and sentiment analysis

A comprehensive survey was conducted by [3] tackling both DA and SA. They argued that polarity calculation is critically affected by discourse structure. The survey also categorized the application of discourse relations to SA into two groups: constraint-based approaches and weight-based schemes.

[33] and [34] represented text spans targeted at the same entity with reinforce relations are constrained to have same polarities, while text spans targeted at opposing entities with reinforce relations are constrained to have opposite polarities. [35] Improved SA by applying conditional relations. [36] Described constraints to eliminate the intra-sentence polarity ambiguities. Where a sentence holding contrast relation has two text spans with opposite polarities. [37] Focused on reweighting each discourse unit depending on the relations in which it takes part.

Our work adopts the discourse-related works of [38] [39] and [40] and carries the idea further in the SA of social media blogs. Our work exploits a variety of features discussed in the literature of Twitter and Facebook-based works to develop a bag-of-words model, in which the discourse features are incorporated to give better sentiment classification accuracy. Our model is evaluated on a dataset using a lexicon-based classification and a supervised classifier, namely, support vector machine (SVM). The dataset consists of 7198 posts and comments using hashtags (#), to measure the accuracy of the classification method.

An early example of a classification enhancement was conducted by [4], where they exploited discourse relations for improving SA in Chinese contexts. Their study focused on using explicit connectives for sentiment classification of texts, which yielded better results than the state-of-the-art at the time. A work conducted by [41] and [42] applied machine-learning techniques on a subjective lexicon dataset fetched from products' reviews on a blog. A similar work conducted by [27] but on Facebook posts with English translation. Others such as [43] have conducted SA by classifying short Arabic political, sports, and social tweets without considering a supervised DA, they employed an unsupervised lexicon-based DA. In our case of Facebook posts, unsupervised lexicons are impractical for classifying long and diversified posts (e.g., status, pictures, and videos). In contrast to their work, our datasets are pre-labeled automatically into three classes (positive, neutral, and negative) of polarity.

In this paper, we aim at performing SA on public Facebook data collected from Jordanian user accounts. Our study differs from existing studies in terms of the dataset scale, source, and the natural language of the texts in the dataset and the extent of experimental analysis that include both ML and discourse relations techniques.

Hence, our method is based on a hybrid technique, the framework of [24] with a DA, specifically a discourse relation. Where, the proposed method uses both ML-based technique (for sentiment classification) and NLP-based technique (for discourse analysis). To the best of our knowledge, this is the first time a hybrid approach of SA and DA is employed for text classification in the context of a Jordanian Arabic dialect extracted from Facebook posts and comments related to the Ukrainian crisis.

3 The Proposed Method

SA is the process of detecting whether a piece of text indicates positive, negative, or neutral feelings. Humans have their natural ability to find out sentiments, however, human-based SA and OM comprises of limitations such as not scalable, time-consuming, unsuitable for real-time decision making, inconsistent if reviewed by different human. Hence, rule-based automation is needed for a clear judgment. To deal with these limitations, a computational model with a hybrid approach for SA and DA is proposed. The flow and functionality of the proposed model is described as follows. The proposed method consists of four main phases shown in Fig. 1.

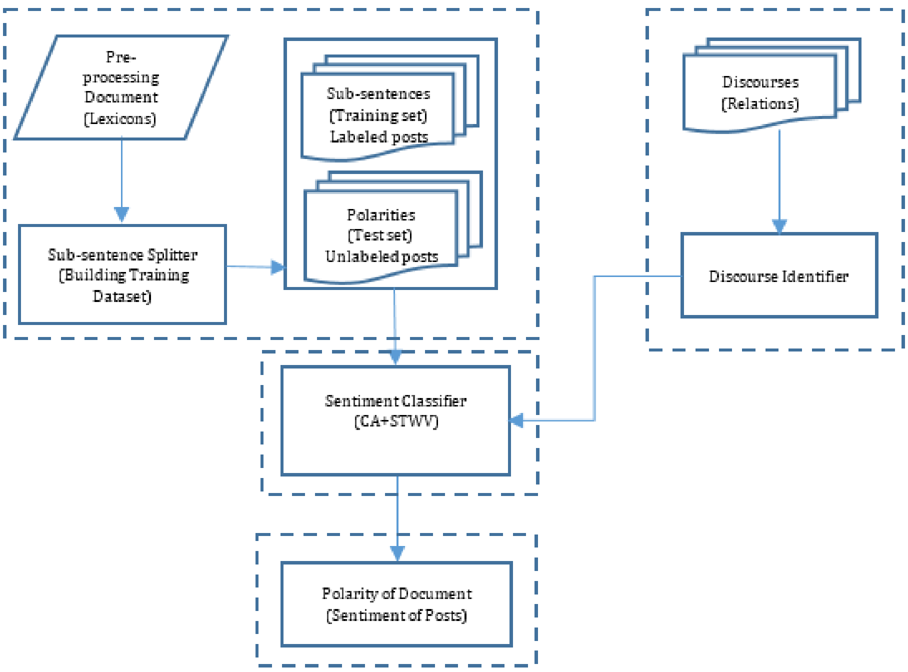

Fig. 1. Model for SA based on DA.

3.1 Phase 1

A document's (post, tag) lexicons is divided into sub-sentences by sub-sentence splitter, in ML concepts, the collected dataset of Facebook posts is first preprocessed by cleaning and removing links and noises. Preprocessing conducts removing unwanted punctuation, removing stop words (when necessary), stemming (when necessary), part of speech tagging and calculating the score of tagged words, which will be passed to the classifier. Then, the document is divided into two sets namely, training set with three labeled classes (positive, neutral, and negative), and a test set without labels. In this phase, the training dataset is built from rows of single sentimental words, instances, and their label. A labelled word tagged positive, negative, or neutral according to the word sentiment. The training dataset has 1037 sentimental words, 317 words were positive and the remaining 720 words are negative. Jordanian dialect sentimental words were added manually with the help of 7198 posts and comments. In this study, we use Weka ML software; therefore, the dataset is represented in an attribute-relation file format. This approach is supervised as each line has one instance, a post or a comment. This instance holds words that do not affect the sentiment but causes confusion in the classifier, hence decreasing the accuracy. If these words are removed from the instance, only sentimental words will remain which is like lexicon in an unsupervised approach.

The collected data for our study was focused on Facebook posts about the Ukrainian crisis. These posts were collected from three different hashtags, namely: (#Russian_Ukrainian_War); (#الحرب_الروسية_الاوكرانية); and (#اوكرانيا). With those hashtags, Facebook posts and comments are automatically collected using a developed Python crawler software that scrapes web contents for extracting specific data. Using more plug-ins for advanced search options, only posts and their comments in Arabic were retrieved to generate the corpora for this study. We were able to extract 7198 posts and comments from each hashtag, which are then compiled into an Excel document. Each Excel sheet is organized according to the hashtags where they were posted. Columns on the Excel sheet (a comma separated value format) included: (a) a column for the post (b) a column for translation; (c) a column for the username and ID; (d) a column indicated whether the posts were for or against the crisis; and (e) a column indicated whether a photo or link was attached to the text. Considering that the posts are written in a Jordanian dialect instead of the standard Arabic, the posts were translated as accurately as possible so that they still conveyed the same meaning despite the cultural and structural perspectives.

3.2 Phase 2

A polarity (training set) is assigned to each sub-sentence-by-sub-sentence sentiment classifiers with a string-to-word-vector (STWV) to detect polarity. The Classification Algorithms (CA) from different learning families used are baseline (ZeroR), decision tree C4.5 (J48), Naïve Bayes (NB), k-nearest neighbor (KNN), and random forest (RF). In which, the preprocessed dataset is transferred as input to the classification algorithms. Here we have tweaked the values of parameters of the classifier to cope with the content of the dataset such as gamma and margin constant parameters.

To minimize the confusion of the classifier mentioned in phase 1, a sentimental lexicon is used as a training set. Therefore, the hybrid SA model incorporates the advantages of supervised SVM classifier and unsupervised lexicon-based technique to build the training set properly that has a positive effect in the classification accuracy. It also reduces preprocessing time including collection of instances, normalization with linear time complexity.

3.3 Phase 3

A discourse identifier contributes to phase 2 at the same time which identifies the discourse type holding by a sub-sentence.

In finding the polarity of a sentence, negation handling is used in sentence-level SA to examine the problem of identifying the scope of negation. Stop word removal, stemming, part of speech tagging and calculated sentiment score in the preprocessing phase, to enable the classification algorithm to classify opinions as either positive or negative. The dataset is analyzed using the unigram feature-extraction technique and the content's polarity provided. Sentiment classification requires working with syntax features for the word-bag method in ML techniques to reveal the grammatical and logistical relationships between the words in sentences. Later, using classification algorithms, the sentiment polarity classification with low-level discourse-based features is automatically extracting connectives and their senses as low-level features to reveal polarity classification of reviews. Hence, a context-aware method is used to analyze sentiment at the level of individual sentences and develop sentence-level sentiment classification using the classification algorithms. A lightweight method for using discourse relations for polarity detection of posts and comments is used with the connectives and conditionals to improve sentiment classification accuracy.

The model first extracts sentiments from a document (e.g., text dataset) and generates entities around sentiments. Eventually, the model leads to decide sentiment orientation of a text around entities in the document. In which, all sentences of the document are broken into its parts of speech, which detects the elements of a document depending upon its grammatical structure (e.g., nouns, adjectives, verbs, and adverbs etc.), connectives, clause boundaries, and discourse arguments. Then the rule based on the grammatical structure is used to find sentiment orientation in the text. It is found by showing whether bigram words are mutually independent or not. For example, in phrase "horrible war" translated from Arabic "حرب مروعة," first word in a bigram is adjective while second is a noun, in Arabic is vice-versa. These two words are mutually dependent followed by any word or by a noun word.

Pre-tagged sentiment lexicons are used after finding sentiment orientation in the text for documents comparison to figure out sentiment-bearing phrases. In social media including Facebook some phrases also embed emoticons. Pre-coded emoticon sentiments are used to decide sentiment orientation of emoticons phrases, e.g., a smiley face emoticon is coded as positive sentiment. Emoticon phrases are highly prioritized in the case of sentiment-bearing phrases. Finally, each phrase polarity is combined to find the eventual polarity of a sentence and entities in those sentences.

Training the classifier via Weka ML software is not compatible with string data type. For this reason, a filter *StringToWordVector* is used to convert string attributes to numeric attributes. We also used *NGramTokenizer* tokenizer, which splits a string into n-gram with minimum and maximum gram. In this case, only unigrams are possible to be applied since the training set has one word in each row.

3.4 Phase 4

A classifier generates the polarity of the document by calculating the weighted sum of sub-sentences by their discourse types.

In the ML perspective, a confusion matrix is generated to show classified positive, negative and neutral opinions. Classification accuracy is calculated based on confusion matrix along with other calculated measures such as F-measure, precision, recall, area under curve, and ROC.

The classifier is evaluated on the training set and then on the test set to produce the expected label of each post or comment. A 10-fold cross-validation splitting strategy is used to evaluate the classifier, which is suitable for small datasets. The result of the evaluation is measured by computing precision and recall as follows [44]:

$$Precision = TP / (TP + FP) \quad (1)$$

$$Recall = TP / (TP + FN) \quad (2)$$

where, TP, FP, TN, and FN are true positive, false positive, true negative and false negative, respectively. The performance matrix is used to calculate classification accuracy which is calculated as the ratio of number of correctly predicted reviews to the number of total number of reviews present in the dataset. The accuracy is calculated as follows:

$$Accuracy = (all\ correct / all) = (TP+TN) / (TP+TN+FN+FP) \quad (3)$$

The classification of a new dataset (namely, test set folded from the training) using the classification algorithms with its tweaked parameters to expect overall sentiment polarity of each post or comment in the dataset. It also needs to be normalized and filtered using *StringToWordVector* filter and *NGramTokenizer* tokenizer, while training the classifier to match words correctly. Classification accuracy is measured by computing the number of correct classified sentiments from posts and comments.

4 Results, Analysis and Discussions

4.1 Experimental results

This research has applied an experiment that is based on supervised ML upon the collected dataset. The datasets were collected using a web crawler that we developed in Python programming for scraping Arabic texts on Facebook for the Ukrainian crisis related posts and comments. It is implemented using the Weka ML software [45] on a PC machine with core-i7 CPU and 12 GB of RAM. The classification algorithms have trained on news domain using 7198 posts and comments. Weka eases ML tasks dedicated for data preprocessing and classification with a variety of filters, functions, and optimization algorithms.

To evaluate how well the suggested hybrid ML model performs, Table 1 shows the results of different classifiers from different learning families, namely baseline (ZeroR), decision tree C4.5 (J48), Naïve Bayes (NB), k-nearest neighbor (KNN), random forest (RF), and Support Vector Machine (SVM).

Table 1 presents the results using the tested classifiers alone, which shows unsatisfactory results using all performance measurements. Fig. 2 also clearly shows the inability of the classifiers alone to discover the sentiments in the dataset. Therefore, a hybrid SA model is proposed in this work for SA.

Table 1: Performance and classification measure of different classifiers

Measure	ZeroR	NB	KNN	J48	RF	SVM
Accuracy	50.54%	50.54%	45.05%	54.94%	54.94%	52.74%
ROC Area	50.00%	57.80%	58.40%	65.12%	66.20%	66.68%
F-measure	34.00%	56.00%	33.00%	53.00%	53.00%	42.00%
Precision	26.00%	64.00%	52.00%	57.00%	55.00%	50.00%
Recall	51.00%	51.00%	45.00%	55.00%	55.00%	53.00%
TP Rate	0.00%	27.20%	41.10%	33.70%	31.40%	4.39%
FP Rate	0.00%	9.09%	52.20%	34.80%	31.46%	3.29%

The following shows the extent of improvement that occurred in the SA results. To verify the performance of the proposed hybrid SA model, Table 2 shows the results of the classification algorithms used in our proposed hybrid SA model, where it is related to the minimum term frequency, which is a key factor in the overall accuracy of the classifiers. For example, a word that is repeated the same number of the minimum term frequency; classifiers consider it a sentimental word. Using Weka, training the classifier on a set of labelled sentimental words is conducted. Then, testing the classifier on a new set of non-labeled sentimental words is also conducted to evaluate and validate the accuracy of the classifier.

Fig. 2. Performance of classification algorithms.

In some cases, misclassification is caused by: comments have negation words which invert the polarity; comments have two sentiments; and comments have ambiguous sentiments. These factors are managed manually to hinder increasing classification confusion. In addition, negation has been considered to inverse the meaning of word.

Table 2: Performance of the hybrid SA model measured for different classifiers.

Measure	ZeroR	NB	KNN	J48	RF	SVM
Accuracy	57.04%	94.39%	79.16%	85.99%	69.96%	94.68%
ROC Area	49.60%	98.70%	79.10%	85.80%	69.50%	94.60%
F-measure	-	94.40%	79.30%	85.90%	69.70%	94.70%
Precision	-	94.50%	80.10%	85.90%	69.70%	94.70%
Recall	57.00%	94.40%	79.20%	85.90%	69.70%	94.70%
TP Rate	57.00%	94.40%	79.20%	85.90%	69.70%	94.70%
FP Rate	57.00%	5.40%	19.70%	14.10%	30.30%	5.40%

Fig. 3 clearly shows the performance of each algorithm using different indicators, it can be observed that the SVM has obtained the best accuracy and outperformed other classifiers, the SVM is known to work efficiently in text classification that showed superior performance in the literature such as [46] [47]. SVM has obtained the highest classification accuracy (94.68%) followed closely by NB with (94.39%) accuracy. This might be due to their native capabilities and structures suited for text classification. However, NB is not truly reliable in our case; the main drawback of NB is that it assumes conditional independence among the linguistic features. If the key features are the tokens extracted from texts, they cannot be considered as independent, since several types of syntactic and semantic dependencies somehow link words co-occurring in a text. Although even NB could produce an oversimplified model, its classification is accurate in some cases.

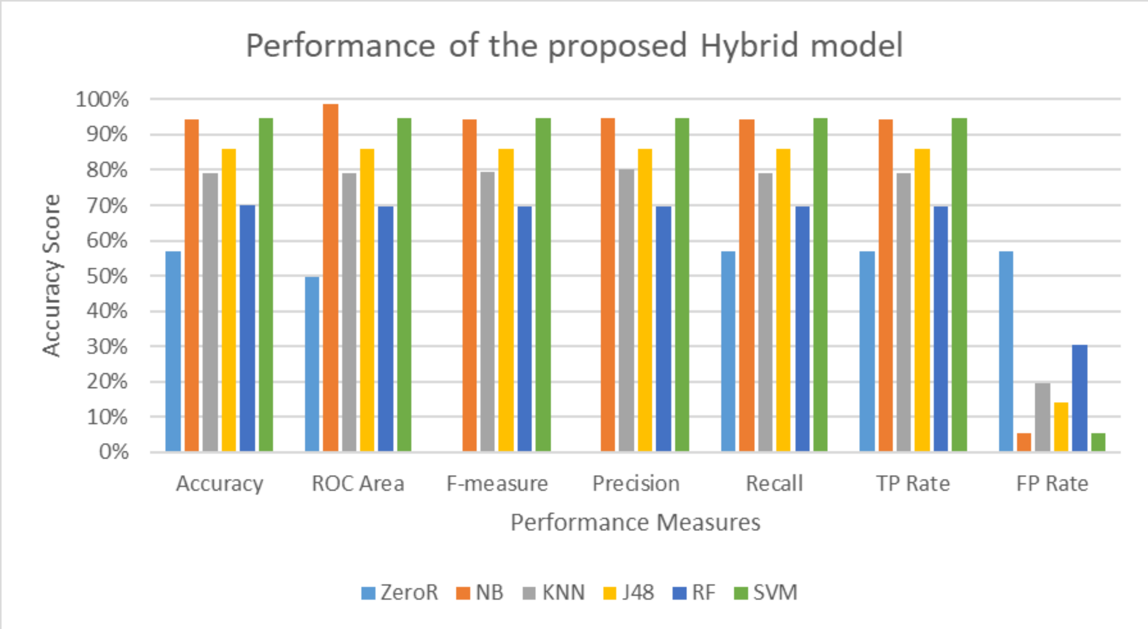

Fig. 3. Performance of the hybrid SA model using different classification algorithms

To verify this, all classifiers have been evaluated and validated upon the same training and testing strategy, which is the 10-folds cross-validation evaluation. Evaluation measurements in the table show the performance of each classifier to support the arguments on their classification capability (see ROC, F-measure, precision, recall) and their ability to avoid overfitting (see TP and FP rates). In all aspects, the SVM has obtained the best evaluation measurements among the classifiers.

TP and FP rates show how accurate the classification is, and how suitable the training model is for the collected dataset. Suitability is the ability of the classifier to avoid model overfitting or underfitting. Overfitting is learning a model with uncleaned training data, which may negatively affect the performance of the model on test data, which eventually limits the ability of the model to generalize across different datasets. On the other hand, underfitting is the inability of a model of neither training data nor generalize to test data. Both issues lead to an inferior performance of a model in the classification task. Hence, we can experimentally avoid or limit overfitting via the implementation of a cross-validation for estimating the model's classification accuracy.

4.2 Discussion

In the context of political and humanitarian news, it is useful to have information about users' sentiments. It may serve as feedback for politicians and jurists, especially in the case of judging the situation by both news followers and humanitarian organizations, where direct contact with war victims is extremely difficult. The proposed hybrid method integrates discourse relations with SA and performed better than baseline methods, confirming the effectiveness of using discourse relations in Jordanian-Arabic dialect SA. We also concluded the most influential discourse relations in SA.

It is important to quickly know the sentiment of the user's posts and comments, so that it improves a social media strategy, and prevents any negative waves from escalating into something critical. Many tools and platforms automate the hard tasks of harvesting, collecting, processing, and classifying opinions to which is handy and time saver for a fast response or drawing solid conclusions. In this aspect, SA considers automating the entire process. For a robust SA an artificial intelligence approach is utilized based on ML methods that analyzes a text and returns the sentiment of one of the following classes: positive (e.g., Ukrainian are keen to negotiate more than ever before), negative (e.g., Russians have initiated the unforgivable), neutral (e.g., my sympathy goes to all victims in Mariupol), and in some cases no sentiment (e.g., none sense). SA along with DA can be applied to Facebook's user post, comment on any post, reply on any post, and a direct conversation. On the other hand, we are aware that it cannot be applied to Facebook page posts or admin comments created by the profile that created the post.

The Ukrainian crisis is an interesting case study from which we have gathered a corpus of 7198 posts and comments that used the (#الحرب_الروسية_الاوكرانية) hashtag from randomly selected accounts to avoid an unfair evaluation. The evaluation is performed by an automatic SA on texts, while employing DA. Our purpose in applying SA was to figure out the opinion of people on the Internet about the Russian-Ukrainian war, trying to identify users who are against or in favor of the crisis.

5 Conclusion

Large data generated by individuals through social networks could be of excellent value when analyzed properly. In this research work we have studied the combination of SA and DA. SA provides a systematic study of affective states using NLP techniques, while DA is a qualitative and interpretive methodology to analyze different socio-cultural phenomena.

This paper proposed a hybrid approach (SA with DA) for Jordanian-Arabic dialect on Facebook examining the classification of randomly collected posts and comments in Ukrainian crisis. We have collected a set of Facebook posts and comments, which have been analyzed and classified using the SA-DA (SVM + discourse relations) hybrid model. We have concluded that this combination is of superior performance, in which the combination has simplified the processing of the language complexity and ambiguity.

The results confirm that the hybrid model has better accuracy score and scalability than standalone ML techniques, meaning using SA without considering DA. The collected posts and comments were analyzed using supervised ML technique based on SVM classifier, which has trained the collected dataset on a single words dictionary. The obtained results by the cross-validation on the dataset clearly confirm the high accuracy of the SVM classifier. The results of this study have shown that the hybrid approach can find opinions of the users and sentiments around those opinions with 94.68% accuracy.

From the existing works, it is seen that there has been no work developed in Arabic (e.g., Jordanian dialect) for the identification of connectives and arguments. Results showed the efficiency of the proposed hybrid model in discourse relation and their argument identification task. This application helps develop a sentence-level SA system using clause and discourse connectives in Jordanian-Arabic dialect. For future work, a more comprehensive sentiment lexicon could be used or designed to improve the classification accuracy or generalize across a variety of Arabic dialects with different domains such as universities, drugs and any other types of entities which would be of significant use.

References

[1] Buimer, H.P., Bittner, M., Kostelijk, T., van der Geest, T.M., van Wezel, R.J.A., & Zhao, Y. (2016). Enhancing emotion recognition in VIPs with haptic feedback. In: *Stephanidis, C. (ed.) HCI 2016. Communications in Computer and Information Science*, 618, 157–163. Springer International Publishing, Cham.

[2] Khan, A., & Lawo, M. (2016). Recognizing emotional states using physiological devices. In: *Stephanidis, C. (ed.) HCII 2016*, 164–171. Springer International Publishing, Cham.

[3] Ali, O. Gegov, A., Haig, E., & Khusainov, R. (2020). Conventional and Structure Based Sentiment Analysis: A Survey, *International Joint Conference on Neural Networks (IJCNN)*, 1-8. doi: 10.1109/IJCNN48605.2020.9207027.

[4] Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle. In *Proceedings of the ACL 2012 System Demonstrations*, 115–120. Jeju Island, Korea. Association for Computational Linguistics.

[5] Luke, A. (1995). Text and discourse analysis in education: An introduction to critical discourse analysis. *Rev. Res. Ed.*, 21, 1-48.

[6] Elahe, R., & Sharififar, M. (2015). Critical Discourse Analysis and Its Implication in English Language Teaching: A case study of Political Text. *Theo. and Prac. in Lang. Stud.*, 5(3), 504-511.

[7] Widdowson, H.G. (2004). Text, Context, Pretext: Critical Issues in Discourse Analysis. *Oxford: Blackwell*.

[8] Tarone, E. (2014). The issue: Research on materials and their role in classroom discourse and SLA. *The Modern Language Journal*, 98(2), 652-653.

[9] Tenorio, E.H. (2011). Critical Discourse Analysis: An Overview. *Nordic J. of Engl. Stud.*, 10(1).

[10] Rahimi, F., & Riasati, M. J. (2011). Critical Discourse Analysis: Scrutinizing Ideologically Driven Discourses, *Int. J. of Human. and Soc. Sci.*, 1(6).

[11] Aini, N., & Widodo, P. (2018). Critical Discourse Analysis of the Bombing Attack News: An Analysis of Teun A. van Dijk's Model, *Advances in Social Science, Education and Humanities Research, Proceedings of the International Conference of Communication Science Research (ICCSR 2018)*, https://doi.org/10.2991/iccsr-18.2018.49.

[12] Prelovac, V. (2010). *Top social media sites*. Web.

[13] Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *Proceedings of Association for Computer Linguistics (ACL)*.

[14] Pak, A., & Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In *Proceedings of LREC*.

[15] Barbosa, L., & Feng, J. (2010). Robust sentiment detection on twitter from biased and noisy data. In *23rd International Conference on Computational Linguistics: Posters*, 36–44.

[16] Bermingham, A., & Smeaton, A. (2010). Classifying sentiment in microblogs: is brevity an advantage, *ACM*, 1833–1836.

[17] Gonzalez-Ibanez, R., Muresan, S., & Wacholder, N. (2011). "Identifying sarcasm in Twitter: a closer look," In *Proceedings of ACL: short paper*.

[18] Joshi, A., Balamurali, A.R., Bhattacharyya, P., & Mohanty, R. (2011). C-feel-it: a sentiment analyser for microblogs. In *Proceedings of ACL: Systems Demonstrations, HLT '11*, 127–132.

[19] Breen, J.O. (2012). Mining twitter for airline consumer sentiment. *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*, 133.

[20] Kontopoulos, E., Berberidis, C., Dergiades, T., Bassiliades, N. (2013). Ontology-based Sentiment Analysis of Twitter Posts. *Expert Systems with Applications,* 40, 4065-4074. 10.1016/j.eswa.2013.01.001.

[21] Montejo-Ráez, A., Díaz-Galiano, M., Martínez-Santiago, F., & López, L. (2014). Crowd Explicit Sentiment Analysis. *Knowledge-Based Systems*. 10.1016/j.knosys.2014.05.007.

[22] Vidya, N.A., Fanany, M.I., & Budi, I. (2015). Twitter Sentiment to Analyze Net Brand Reputation of Mobile Phone Providers, *ProcediaComput. Sci.*, 72, 519-526.

[23] Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets, *Int. J. Hosp. Manag.*, 55, 16-24.

[24] Junaid, S.M., Jaffry, S.W., Yousaf, M.M., Aslam, L., & Sarwar, S. (2017). Sentiment Analysis and Opinion Mining - A Facebook Posts and Comments Analyzer, *Technical Journal, University of Engineering and Technology (UET) Taxila*, 22, 98-104.

[25] Ortigosa, A., Martín, J.M., & Carro, R.M. (2014). Sentiment analysis in Facebook and its application to e-learning, *Computers in Human Behavior*, 31, 527-541, ISSN 0747-5632, https://doi.org/10.1016/j.chb.2013.05.024.

[26] Çoban, Ö., Özel, S.A., & İnan, A. (2017). Deep Learning-based Sentiment Analysis of Facebook Data: The Case of Turkish Users, *The Computer Journal*, 64(3), 473–499. https://doi.org/10.1093/comjnl/bxaa172

[27] Shudhanshu, T., & Anurag, S. (2020). Sentiment Analysis of Facebook Data using Machine Learning. *International Journal of Innovative Research in Applied Sciences and Engineering*, 4, 2456-8910. 10.29027/IJIRASE.v4. i4.2020.735-742.

[28] Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal texts. *J. Am. Soc. Inf. Sci. Technol.*, 61, 2544–2558.

[29] Basiri, M.E., & Kabiri, A. (2018). Words Are Important: Improving Sentiment Analysis in the Persian Language by Lexicon Refining. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, 17(4), 18 pages. https://doi.org/10.1145/3195633

[30] Ravi, K., & Ravi, V. (2015). A Survey on Opinion Mining and Sentiment Analysis: Tasks, Approaches and Applications. *Knowledge-Based Systems,* 89, 14-46. https://doi.org/10.1016/j.knosys.2015.06.015

[31] Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis, In: *Mining Text Data, Aggarwal, Ch.C., Zhai, Ch.X. (eds.)*, 415-463.

[32] Vinodhini, G., & Chandrasekaran, M.R. (2014). Opinion mining using principal component analysis-based ensemble model for e-commerce application, *CSI Trans. ICT*, 2(3), 169–179.

[33] Somasundaran, S., Wiebe, J., & Ruppenhofer, J. (2008). Discourse Level Opinion Interpretation. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, 801–808. Manchester, UK. Coling Organizing Committee.

[34] Somasundaran, S., Namata, G., Wiebe, J., & Getoor, L. (2009). Supervised and Unsupervised Methods in Employing Discourse Relations for Improving Opinion Polarity Classification. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 170–179. Singapore. Association for Computational Linguistics.

[35] Narayanan, R., Liu, B., & Choudhary, A. (2009). Sentiment Analysis of Conditional Sentences. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 180–189, Singapore. Association for Computational Linguistics.

[36] Zhou, L., Li, B., Gao, W., Wei, Z., & Wong, K.-F. (2011). Unsupervised Discovery of Discourse Relations for Eliminating Intra-sentence Polarity Ambiguities. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 162–171). Edinburgh, Scotland, UK. Association for Computational Linguistics.

[37] Heerschop, B., Goossen, F., Hogenboom, A., Frasincar, F., Kaymak, U., & de Jong, F. (2011). Polarity analysis of texts using discourse structure. *In Proceedings of CIMK*, 1061-1070.

[38] Polanyi, L., & Zaenen, A. (2004). Contextual valence shifters. In *Exploring Attitude and Affect in Text: Theories and Applications*. AAAI Spring Symposium Series.

[39] Wolf, F., & Gibson, E. (2005). Representing discourse coherence: A corpus-based study, *Computational Linguistics*, 31(2), 249–287.

[40] Taboada, M., Brooke, J., & Voll, K. (2008). Extracting Sentiment as a Function of Discourse Structure and Topicality. *Simon Fraser Univeristy School of Computing ScienceTechnical Report*.

[41] Medagoda, N., Shanmuganathan, S., & Whalley, J. (2015). Sentiment Lexicon Construction Using SentiWordNet 3.0. 10.1109/ICNC.2015.7378094.

[42] Jadav, B.M., & Vaghela, V.B. (2016). Sentiment Analysis using Support Vector Machine based on Feature Selection and Semantic Analysis. *International Journal of Computer Applications* (0975 – 8887), 146(13), 26-30.

[43] Alhumoud, S., Altuwaijri, M., AlBuhairi, T., & Alohaideb, W. (2015). Survey on Arabic Sentiment Analysis in Twitter. *International Journal of Social, Behavioral, Educational, Economic and Management Engineering*, 9, 364-368.

[44] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Inf. Process.Manag.*, 45(4), 427–437.

[45] The official website of WEKA machine learning software https://www.cs.waikato.ac.nz/ml/Weka/ [Online] Accessed on: 2/11/2022.

[46] AlBadani, B., Shi, R., & Dong, J.A. (2022). Novel Machine Learning Approach for Sentiment Analysis on Twitter Incorporating the Universal Language Model Fine-Tuning and SVM. *Appl. Syst. Innov.*, 5(13). https://doi.org/10.3390/asi5010013

[47] Basiri, M.E., & Kabiri, A. (2020). HOMPer: A new hybrid system for opinion mining in the Persian language. *Journal of Information Science*, 46(1), 101–117. https://doi.org/10.1177/0165551519827886