Sudan University of Science and Technology
College of Graduate Studies

Development of a Deep Neural Networks Based on Framework for the Discovery of User's Interests and Sentiments Analysis

تطويرإطارعمل بإستخدام الشبكات العصبية العميقة لإكتشاف إهتمامات المستخدمين وتحليل مشاعرهم

A thesis submitted to the College of Graduate Studies,
Faculty of Computer Science and Information Technology
Sudan University of Science and Technology
In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Computer Science

By

Abubakr Hassan Dafallah Ombabi

Supervisor

Professor: Adel M. Alimi

January 2021

ACKNOWLEDGMENT

Immeasurable appreciation and deep gratitude for the help and support to all who in one way or another has contributed to making this Ph.D. study. I would like to express particular appreciation to my Ph.D. project director Professor Adel M. Alimi for his guidance and persistent help that allows the achievement of this dissertation. Many thanks to Doctor Dr. Wael Ouarda the quality manager of Search Research Groups in Intelligent Machines (REGIM-Lab) at Sfax University for his integration assistance and for his advice. Also i would like to thank Sudan University of Science and Technology which have given me this great opportunity to conduct this thesis.

Abstract

Recently, the world has witnessed an exponential growth in the number of social networks subscribers. Social networks have opened a venue for online users to interactively share their opinions in different life aspects. Despite that Users interests classification and sentiment analysis have become more critical for services providers, online advertising, and E-commerce, minimal research efforts have been conducted on these areas. However, the majority of existing approaches are unsupervised or partially supervised they primarily ignores the semantic information within the sentences, the complicated structure of the natural languages, in addition to lack of tools and available sufficient corpora are the main challenges in this area. In this study, we propose a deep learning model for users interests discovery and sentiment classification. The proposed model takes advantage of the FastText model and WordNet lexical database for words representation. Convolutional Neural Network architecture is used to capture local features. To overcome the limitation of the Convolutional Neural Network in capturing the temporal information over long-distances, in this study a recurrent neural network is designed to capture the contextual information and long-term dependencies, thus, increasing the classification accuracy. The proposed model is implemented on Tensorflow under Python environment. The model is evaluated using our constructed multi-domains Arabic sentiment corpus which contains 32,950 instances, Amazon corpus which contains 60,000 instances, in addition to many other corpora. Experimental results have demonstrated the outstanding performance of the proposed architecture. Also, intensive experiments are conducted to validate the impact of each deep learning component in the classification performance. The proposed model has achieved a classification accuracy of up to 88.58% for Arabic text and up to 90.20% for English text. Furthermore, the proposed model has remarkably outperformed many existing approaches for both tasks.

المستخلص

شهد العالم مؤخراً نمواً هائلاً في عدد مشتركى شبكات التواصل الاجتماعي، التى اتاحت المجال لمستخدمى الانترنت لمشاركة آرائهم فى مختلف جوانب الحياة بطريقة تفاعلية. على الرغم من ألاهمية الكبيرة لمجالات إكتشاف اهتمامات المستخدمين وتحليل المشاعر بالنسبة لمزودى الخدمات، شركات الإعلانات، والتجارة الإلكترونية، إلا ان هذة المجالات وجدت القليل من المجهودات البحثية. ومع ذلك، فإن غالبية التقنيات الحالية تعتمد على طريقة التعلم الغير خاضع للإشراف أو التعلم شبة الخاضع للإشراف، فهي تتجاهل بشكل أساسي المحتوى المعنوى للجمل، كما أن التركيب المعقد للغات الطبيعية، بالإضافة إلى نقص الأدوات ومجموعات البيانات التجريبية المتاحة تمثل تحديات اساسية في هذا المجال. هذه الدراسة، تقترح نموذج مبنى على تقنية التعلم العميق للاله لإكتشاف إهتمامات المستخدمين وتصنيف مشاعرهم. يستفيد هذا النموذج من نموذج (النص السريع) وقاعدة بيانات معجمية فى تمثيل الكلمات. ايضا تم استخدام الشبكات العصبية الالتفافية لإستخراج السمات المحلية للنص. للتغلب على ضعف الشبكة العصبية الالتفافية في التقاط تسلسل المعلومات فى السلاسل الطويلة، في هذه الدراسة، تم تصميم شبكة عصبية ارتجاعية لالتقاط المعلومات السياقية وتسلسل المعلومات فى السلاسل الطويلة للتاكيد على السمات السياقية المؤثرة على التصنيف، وبالتالي زيادة دقة التصنيف النهائى. تم تنفيذ النموذج المقترح باستخدام بيئة (تنسور فلو) على بيئة لغة (بايثون). تم تقييم اداء النموذج باستخدام مجموعة بيانات تحليل المشاعر متعددة المجالات للغة العربية والتي تحتوي على ٣٢٫٩٥٠ عينة بيانات نصية، ومجموعة بيانات (امازون) والتي تحتوي على ٦٠٫٠٠٠ عينة بيانات، بالإضافة إلى العديد من مجموعات البيانات الأخرى. أظهرت النتائج التجريبية الأداء المتميز للمعمارية المقترحة. كما تم إجراء تجارب مكثفة للتحقق من تأثير كل خوارزميات التعلم العميق وبدائلها في ألاداء العام للنموذج. اثبتت التجارب ان النموذج المقترح حقق دقة تصنيف تصل إلى ٨٨٫٥٨ % للنصوص العربية و ٩٠٫٢٠ % للنصوص الإنجليزية. علاوة على ذلك، فقد تفوق النموذج المقترح بشكل ملحوظ على العديد من التقنيات الحالية فى كل من اكتشاف اهتمامات المستخدمين وتحليل المشاعر.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One

Introduction

Social networks (SNs) such as Instagram, Twitter, and Facebook, etc have become attractive communication channels that enable online users to express their interests, experiences, tendencies, and opinions towards different life aspects such as events, services or products, etc. In the SNs users can widely share their opinions using different forms of social data (SD) such as textual data (e.g. comments, tweets, reviews, etc.), visual data (e.g. shared and liked images), or voice data, therefore, analyzing this huge volume of data to capture the sentimental tendencies of the audience has direct impacts on products/services quality improvements, predicting of upcoming marketing trends, changing sales strategies, etc (Sohangir et al., 2018; Liu et al., 2018; Cinar, Zoghbi, and Moens, 2015). With the exponential growth of SD, conventional analysis is a very time and resources consuming task. Therefore, many researchers have proposed an automated or semi-automated techniques that can significantly help in handling and analyzing this massive amounts of social data. Social Network Analysis (SNA) is an approach based on networks and graph theory. It uses social networks structures to detect people's hidden information such as demographic information, interests, sentiment analysis (Chang, 2018). In fact, social networks analysis can be applied in several area of application since it allows for the analysis of the user's generated contents to discover the overarching feelings on a product, service, or brand, thus make enhancements. Opinion Mining is a considerable research area that is concerned with the use of Natural Language Processing (NLP), computational linguistics, and text analysis techniques (Y. Chen and Z.

Zhang, 2018) to analyze and understand the opinions, emotions, and attitudes toward a particular topic, event, entities, etc (Mostafa, 2017; Alowaidi, Saleh, and Abulnaja, 2017; Al-Smadi et al., 2018). User interest discovery is the process of automatically identifying the areas or topics of interest for the individual user based on his/her information collected from SN or websites. In SNs users tend to follow a particular community or other users according to their own area of interest Interests. Therefore, discovering users interests have significant roles in user-centric applications and personalized web services development (Darabi and Tabrizi, 2017). Also, recommendations systems depend mainly on user interests discovery to provide users with huge options of new or unknown releases of products/services as they need (McAuley, Pandey, and Leskovec, 2015), Such as recommending relevant book from Amazon to users whose interested in reading. In fact, the revenue of social media websites comes mainly from online advertisements, so users interest identification allows to obtain more personalized advertising (Lewenberg, Bachrach, and Volkova, 2015; H. C. Yang, C. H. Lee, and C. Y. Wu, 2018). In fact, there are two types of user interests: explicit interests in which users explicitly express their interests orientations using likes, posts, etc or provide feedback as comments or ratings (Darabi and Tabrizi, 2017). Implicit interests which are obtained from analyzing of user's profiles contents (inferred from explicit ones) such as browsing behaviors (e.g. time duration on a page) (e.g. Yunfei Ma et al., 2011). Meanwhile, sentiment analysis (SA) is concerned with computationally determine whether opinions toward a specific event, product, topic, or service are positive, negative, or neutral. According to Yunfei Ma et al., 2011 textual data sentiment analysis focuses on obtaining subjective information using text mining, linguistics and statistical knowledge. Recently many studies have proposed several Machine Learning approaches for textual data SA tasks either on a document, paragraph, sentence, or word levels (Y. Lu et al., 2018). In fact, sentiment analysis can be used to evaluate for example movie reviews, to determine which movie has more positive or negative opinions, and to answer questions such as why those opinions are negative or positive?.

According to Attia et al., 2019, it's challenging to capture the users interests and sentiment orientations from textual data regardless of deep considerations of syntactical and semantic rules of the language. Also, meanings of words are strongly tied to the context and the long-short distances dependencies between words in sentences. Deep Learning (DL) is a new area of neural networks. It emulates the human learning behaviors using multi-layer neural networks architectures to handle a cognitive task, in contrast to conventional ML-based-LNP approaches (e.g SVM and logistic regression) which depend mainly on hand-crafted features, DL models can automatically perform multi-level abstract features representation of the original data (Young et al., 2018a; Y. Chen and Z. Zhang, 2018). Which make a simple DL model able to obtain superior performances in AI tasks (Sohangir et al., 2018). Inspired by the successful application of DL in speech recognition (e.g. Graves, A.-r. Mohamed, and G. Hinton, 2013) and computer vision, DL models have presented remarkable advances in several NLP challenges such as sentence and document representation (Mikolov, K. Chen, et al., 2013), sentiment analysis (e.g. Y. Kim, 2014a; H. C. Yang, C. H. Lee, and C. Y. Wu, 2018) machine translation (e.g. Merri and Fellow, 2014) and named-entity recognition (e.g. Gridach, 2016).

In this thesis, we make an attempt to come up with a novel deep learning architecture to predicate the topics of interest such as (hotels, music, sport, etc) for social networks users in addition to detect the sentiment orientation over the predicated topics whether its positive or negative to get the correct user interest topics.

1.1 Motivations and problem statement

Despite that, the majority of existing text classification approaches are unsupervised or semi-supervised, they can weakly categorize the user's generated textual data, but can not pay attention to the overall meaning of the sentences, therefore, these approaches are not effective in detecting the correct orientation of the sentences as confirmed in (Hai et al., 2017). Also it is challenging to perform text categorization regardless of deep consideration of syntactical and semantic rules of the language, therefore, traditional approaches ignore the sentiment orientation over the predicted topics. Furthermore, due to the complicated structure of the Arabic language, lack of tools and available sufficient corpora, text categorization has comparatively attracted minimal research efforts than in the English language.

1.2 Important of the research

This research is concerned with discovering hidden knowledge about the users i.e (customers - consumers), this knowledge have a significant roles in user-enteric applications and personalized web services development. Consumers are interested in receiving recommendation based on their areas of interest, there an ubiquitous application that depend mainly on user interests discovery such as recommendations systems which is incorporated with a wide range of e-commerce, online news, SNs, and governments platforms to provide users with immense options of new releases of products/services as they need, this could remarkably improve users experience and satisfaction. In fact the revenue of social media websites comes mainly from online advertisements, therefore, users interest identification enable to target specific audiences with personalized advertising. It allows to target users whose interest orientation is political with a correct promotional materials. Also enterprises are allowed to analyze opinions about their products for further improvements.

1.3 Summary of related works

Text categorization is one of the fundamental tasks in Natural Language Processing (NLP) with wide applications including interests or preferences prediction, and sentiment analysis. Recently, users interests discovery have gained much research attention, different approaches have been proposed for different purposes such as recommendation systems, user's hidden information discovery (e.g preferences), predicating political trends in public opinions, users modeling and categorization, communities' detection etc. Almost all of the proposed approaches are unsupervised which focused on using Natural Languages Processing, linguistic analysis, lexicon based to detect the topics of individual user interest. A very few partially supervised approaches are proposed to predict user interests, those approaches are not efficient as confirmed in Hai et al., 2017. Tables 1.1, 3.1, and table 1.3 summarize some state of the art in user interest discovery and sentiment analysis for several languages, more details about the approaches is presented in chapter tow.

1.4 Research questions

- What are the limitations of the existing Deep Learning algorithms ? How to overcome these limitation ?
- Which deep learning algorithms can be used to effectively predict users interests and sentiments using textual contents ?
- How to improve the state of the art Deep learning to perform text classification **semantically** ?
- How to make deep learning algorithms focus on the important parts of the text which can influence the classification correctness ?
- Which is the best performing words representation model for Arabic text contents ?

1.5 Research hypotheses

- Hypothesis 1: Integrating different Deep learning algorithms can provide better classification performance in a complex morphological language such as Arabic.
- Hypothesis 2 :Using WordNet lexical database enhances the quality of the representation, thus improving the overall performance.

- Hypothesis 3: Stacking LSTM over CNN provide more substantial contextual information extraction.
- Hypothesis 3: Stacking a neural network classifier (e.g SVM) improves the classification performance of Deep Learning architecture.

1.6 Research objectives

The main objective of this research is to design a joint topic-sentiment based deep learning architecture to automatically detect the user's topics of interest and the sentiment inclination over the detected topics using users' generated textual data, also this research aims to achieve the following objectives:

- optimize the processes of features extraction and classification by integrating Convolutional Neural Network with Recurrent Neural Network.
- investigate the impact of different representation models, contextual information extractors, and classifiers on the text classification of multilingual textual contents.
- contribute to fulfilling the limitation of Arabic opinion mining resources by constructing Arabic sentiment analysis corpus which can be also used in topics classification.

1.7 Research methodology

This section provides an overview of the research methodology followed in this thesis. This thesis have adopted the constrictive and action research methodologies to achieve the goals of understanding and analyzing users interests discovery and sentiment analysis. The constructive research methodology aims at constructing a solution to the research questions in form of theoretical model of users interests discovery and sentiment analysis. This model is constructed based upon the existing theories form which the theoretical body of knowledge assists as a tool in the creation of the model. In order to construct the framework and to perform practical evaluation, a methodology based on action research have been used in which each actions are decided iteratively and the phases provide feedback to each other.

The proposed users interests discovery and sentiment analysis framework has been developed in five phases each with questions that address the research aims and objectives. Phase 1. investigate the current state of opinion mining based techniques. Phase 2. Creation of the theoretical model which seeks to be the solution to the main research objective. Phase 3. introduces an enhanced features representation and extraction techniques based on contextual information extraction and attention mechanisms. Phase 4. Validation of the Framework's practical relevance by including it in user's interests discovery and sentiment analysis's contexts. Finally, Phase 5. compares the result achieved by the proposed framework with the state of the arts techniques. Chapter three presets deeper details of the research methodology used in this thesis.

1.8 Research scope

This research is concerned with the discovery of the subjective information expressed in textual data and determine the mind-set of the audience towards an issue. The main focus of the research is to discover the topics of interests for individual user and to detect the sentiment orientation of the user in whether its positive or negative based to the user's generated textual comments over online social networks particularly, Twitter and Facebook.

1.9 Thesis outline

This thesis has the following structure: Chapter 2 provides the necessary theoretical background in addition to a detailed discussion of related works and methodologies to our research. Chapter 3 introduces the research methodology and the proposed solution for Users Interests Discovery and Sentiment Analysis Based on Text Social Data. Chapter 4 presents further interpretation of the results. Finally, conclusions are drawn, and future work set out in chapter 5.

1.10 List of Publications

A. H. Ombabi, O. Lazzez, W. Ouarda and A. M. Alimi, "Deep learning framework based on Word2Vec and CNNfor users interests classification," 2017 Sudan Conference on Computer Science and Information Technology (SCCSIT), Elnihood, 2017, pp. 1-7, doi: 10.1109/SCCSIT.2017.8293054.

A. H. Ombabi, Ouarda, W. Alimi, A.M. Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks. Springer Nature. 10, 53 (2020). https://doi.org/10.1007/s13278-020-00668-1

A. H. Ombabi, W. Ouarda and A.M. Alimi, " A Deep Neural Network Model for Users Interests Discovery and Sentiment Analysis Based on Text Social Data," 2020, International Arab Journal of Information Technology. (Accepted)

Table 1.1 Users interests discovery based English text

Study	Approach	Features	Dataset	N.classes	Acc%	Limitation
(Liang and Lai, 2002)	Networks structure and user's profiles analysis with WordNet	contents viewing time	SRI and HLA	10	49.8	inaccurate predictions low accuracy.
(Bhargava, Brdiczka, and Roberts, 2015)	semantic relatedness, NER, social tagging	user's profiles and activities	488 Facebook users accounts	4	88.1	UI is confined to the viewed webpages.
(Pennacchiotti and Popescu, 2011)	Latent Dirichlet Allocation, lexical analyzer	behaviors, network structure and the linguistic contents	10.000 Twitter accounts	2	88.9	Low accuracy Small data size
(K. Xu et al., 2018)	unified probabilistic topic model	N. most likely followers	Sina Weibo and Epinions	N\A	83.9	Accuracy is reduced as N increased
Lee2011	TF and weighting	TFs and likes	715 Posts	10	70.5	Words vagueness reduce accuracy
(Darabi and Tabrizi, 2017)	domain ontology, bag of terms, Porter Stemming and WordNet	TFs of nouns	20.000 documents	10 concepts	-	small dataset
(Mangal, Niyogi, and Milani, 2013)	Stanford coreNLP, taggers, and semantic similarity from WordNet	TFs noun, adjective	2,02,578 tweets	9	73.2	ambiguity of terms and keywords
(Liu et al., 2018)	stacked autoencoder and stacked LSTMs	RNN	Yelp and Epinions	30	-	low evaluation values
(H. Xu et al., 2016)	CNN-channel, CNN-concatenation, and LDA	Skip-gram CNN	143,853 citations	2	82.7	N/A
(Fornaia et al., 2015)	concept graph, and community detection	bag of concepts graph	(77K news articles)	18 concepts	-	low quality, lack of coherence, among others
(Fornaia et al., 2015)	profiling agent RBPNN	RBPNN	250 users profiles	N\A	-	users profiles contents as data
(Hong, C. Choi, and Shin, 2018)	CNN, hierarchical topics	CNN for text and images features	DMOZ datset	16	79.8	Some aspect not presented

Table 1.2 Users interests discovery based Arabic text

Study	Approach	Features extraction	Dataset	N.classes	Acc%	Limitations
(Mourad Abbas and Daoud Berkani, 2006)	TF-IDF and SVM classifier	TF-IDF of interest terms	Alkhabar, includes 6000 news articles	4	84.37	TFIDF considered as classifier and feature extractor
(Mourad Abbas, Smaïli, et al., 2017)	triggers using Average Mutual Information, TR-Classifier	KNN, TR-distance	Alwatan,9000 articles.	6	-	words representation not presented
(Zrigui et al., 2012)	stacking Latent Dirichlet Allocation -SVM	LDA topics terms	1500 documents	9	88.2	Small dataset
(Kelaiaia and Merouani, 2013)	Latent Dirichlet Allocation and K-means	StemmerLight10	Reuters and TDT2	11	75.6	weak performance
(Koulali, El-Haj, and Meziane, 2013)	TFIDF and a distance-based approach similarity measure	TF-IDF	Al-Wattan	6	84.6	Words embedding not presented
(Koulali and Meziane, 2014)	Named Entity couples, NE boundaries	mutual information measure	Wattan	6	78.4	Results are not well presented
(Hmeidi, Hawashin, and El-Qawasmeh, 2008)	Knn and SVM	KNN	2250 Arabic articles.	2	76.5	N\A
(Cheng and Soon, 2006)	propagation neural network (BPNN) , Porters stemming	N\A	Reuter-21578	10	-	BPNN cannot cope with Arabic text well
(Soucy and Mineau, 2005)	KNN and SVM	TFIDF , ConfWeight	Reuters-21578, Reuters Vol 1 and Ohsumed	N\A	74.9	Results not clear

Table 1.3 Sentiment analysis state of the art

Study	Approach	Dataset	Features	Classifier	Acc%	Limitations
(Ouyang et al., 2015)	word2vec, CNN , and two-layer LSTMs	7226 comments	CNN	SVM	87.2	Could try more classifiers.
(Medrouk and Pappa, 2017)	Multi-lingual CNN architecture	N\A	CNN N-gram features	SVM	93.2	Using individual words polarities technique
(Al-Smadi et al., 2018)	Companied CNN and BiLSTM.	LARB dataset of 63,257 reviews	CNN	CRF	86.4	CNN-BiLSTM must achieve higher accuracy.
(Preethi and Krishna, 2017)	Recursive Neural Networks	Amazon dataset	RNN	NB, SVM	90.47	N\A
(R. Ghosh, K. Ravi, and V. Ravi, 2016)	two layers Restricted Boltzmann with Probabilistic Neural Network	MOV, BOO, DVD, ELE, and KIT	POS tagger, TF-IDF	PNN	94.9	Could try more classifiers.
(Taj, Shaikh, and Fatemah Meghji, 2019)	WordNet and SentiWordNet lexical databases	BBC news articles	TF-IDF	TF-IDF	-	Results not clear Small data size
(X. Li et al., 2017)	combines Word2vec, RSM, and BTM model	Wikipedia Talk Pages, and Twitter	LSM	BTM	-	Could try more classifiers.
Abd-Elhamid, Elzanfaly, and Eldin, 2016	tree structure, terms weighting approach	N\A	Part Of Speech, rules	N\A	92.1	Not dataset details. Only 5 rules are experimented
(Alowaidi, Saleh, and Abulnaja, 2017)	lexicon-based Arabic WordNet	consisted of 1000 tweets	N\A	NB and SVM	85.9	N\A

Chapter Two

Theoretical Background and Related Works

2.1 Overview on text analytic

Recently, a huge volume of textual contents is being generated on a daily basis. For instance, Twitter users generate 500 million tweets in one day (Hmeidi, Al-Ayyoub, N. A. Abdulla, et al., 2015). These huge volume of user generated data contain valuable information which need to be analyzed effectively. These data could be used in a variety of applications to improve human life. Analyzing this massive amount of textual data using conventional approaches is very resources consuming task, therefore, more advanced algorithms are involved to discover the hidden patterns in these data. Text analytic provides approaches to process and analyze this huge and unstructured textual data to obtain high quality of information about the audience. According to El-halees, 2014 text Analytic provide a collection of machine learning, linguistics and statistical approaches which can be use in analyzing and structuring the textual contents to support the process of decision making in organization, services providers, governments, etc. Text analytic techniques including text categorization (TC) and sentiment analysis (SA) can be used in to analyze the audience opinions, users attitudes and emotions toward different life aspects. Actually, text categorization is a challenging task that involves a great effort to process features (Vo et al., 2017). Basically, text analytic framework consists of three main processes which are prepossessing, representation, and knowledge discovery as illustrated by (Aggarwal and Zhai, 2013) in Fig. 2.1.

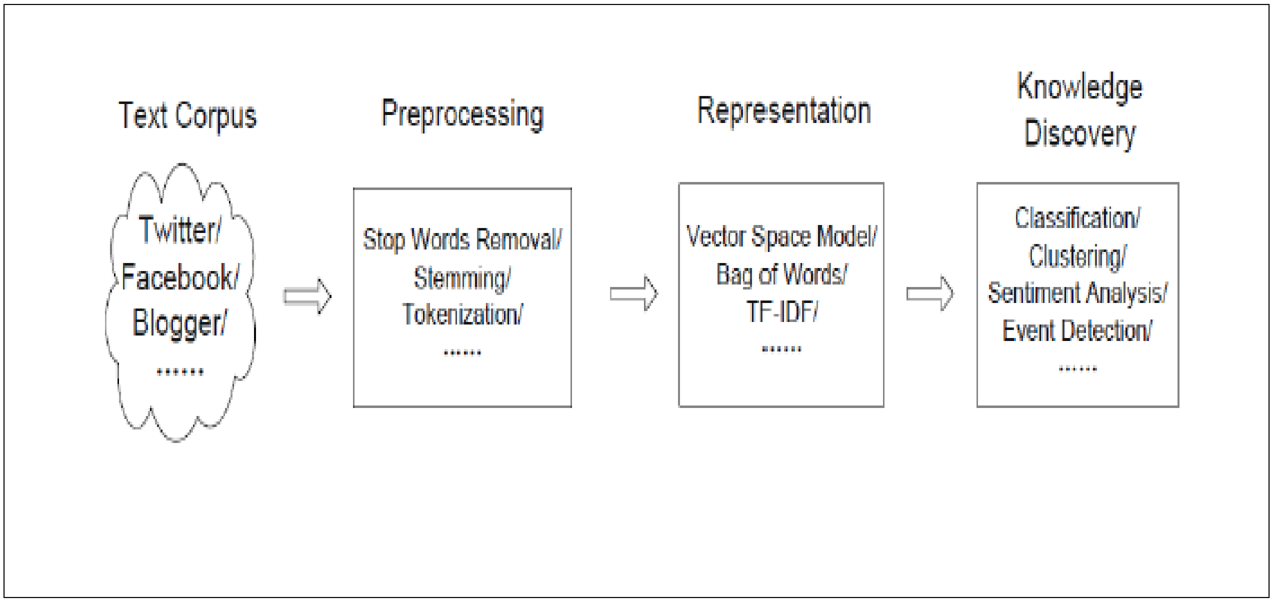


Figure 2.1 Overall processes of text analytic

2.2 Text classification (TC)

Text Categorization (TC) is recently described as a rapidly developed branch of the applied research area (Mahalakshmi and Duraiswamy, 2012). Due to the increased volume of textual data generated daily by online users and thire activities. Text categorization is the field that utilizes natural language processing, information retrieval, machine learning and data mining techniques to provide solutions for real life problems (Mesleh, 2007). The main aim of TC is to assign each piece of text or document to one or more predefined label. According to (Lam and D. K. Lee, 1999). Also, text categorization aims at successfully finding the most proper scope or topic of among categories for the word or sentence at hand. Text categorization plays significant role in different areas including topic detection and tracking, word sense disambiguation, information retrieval, web pages classification, as well as any application requiring document organization. The text categorization applications includes Automatic Indexing, Document Organization, Document Filtering, Word Sense Disambiguation (Odeh et al., 2014). Text Categorization is divided into three major tasks: development of linguistic resources, sentiment classification, and opinion memorization (Elawady, Barakat, and Elrashidy, 2014).

2.3 Text classification general approach

Text classification or opinion analysis is based on analysis level and process definition. Indeed, textual data is commonly used to express opinions on several levels (i.e. sentence, paragraph, document, phrase or aspect level) (Al-ayyoub and Nuseir, 2016), accordingly, opinions can be obtained either at several levels including word, sentence, paragraph, document or aspect levels. Sentence-level opinion mining focuses on categorizing whether an opinion expressed on a sentence is subjective or objective, sentence level deals with a grammatical form of the sentence that expresses an independent statement, comment, request, or exclamation, etc (e.g. we find here such a good food). Paragraph level examines the a unit of more than two words associated with a grammatical construction (e.g. such a good food). Document-level OM serves on determining whether an opinion document is containing positive or negative sentiment, finally, aspect-level OM serves on categorizing the opinion with respect to a particular aspect. Some studies considered the opinion orientations to be simply binary (i.e positive or negative), however others added more labels (Alsmearat et al., 2015). According to Medhat, Ahmed Hassan, and Korashy, 2014, text categorization or opinion analysis task is mainly passes through three stages: opinion identification, features selection and opinion classification as illustrated in Fig. 2.2. Every stage represents a research area.

2.3.1 Opinion identification

It is also called subjectivity identification. Opinion identification can be considered as a process of selecting the emotions containing parts of the sentences (S. Verma and Pushpak Bhattacharyya, 2009). This process aims to classify the language contents at hand (i.e. sentence, phrase or word) into two groups: subjective and objective. This process is very

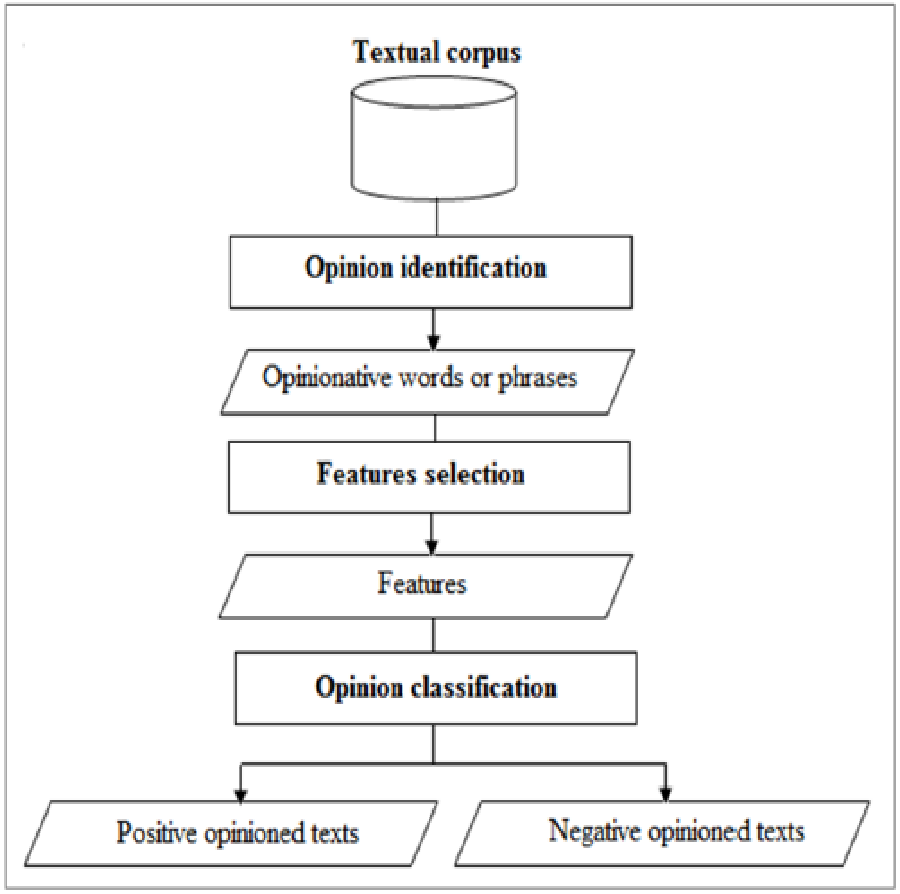


Figure 2.2 Opinion analysis general steps

challenging because it is context dependent.

2.3.2 Features selection

The second step in opinion analysis consists on text features selection. Text features can be term-based, term presence and frequency, part of speech, opinion word or phrase, negation, aspect-based, etc. Term-based is concerned with the identification of term position in the text. In the case of term presence and frequency, only some terms (i.e. graphical words or word n-grams) as being more effective for opinion classification are supported. Parts of speech deal with identifying adjectives since they are significant opinion identifiers. Opinion words or phrases are some words or phrases that express usually opinion. Sometimes, they are or include opinion words. While in other times they doesn't. With regard to negation, the use of negative word or particles changes opinion orientation. For features selection many techniques and methods are used. They represent texts as a bag of words BOW (i.e. a simplifying model for text representation as mustiest of its words ignoring grammar and order but retaining multiplicity (Harris, 1954). Indeed, each dimension encodes how often the feature occurs in the text) or a string which retains the sequence of words in the text. Many methods were used for feature selection. They can be grouped into two main classed: lexicon-based and statistical. Lexicon-based methods need human intervention for annotation. However, statistical methods are automatically realized e.g. Point-wise mutual information (Cover et al., 1991), Chi-square (Yang Y., 1997), Latent semantic indexing (Deerwester et al., 2015).

Feature selection for classification

Overall feature selection process for classification framework is presented by Aggarwal, Kong, et al., 2014 in Fig. 2.3. The quality of the training process of classification model essentially influenced by the performance of feature selection. After features generation, instead of feeding the classification algorithm with the completed data and features, it is better to perform feature selection to identify a subset of important features, then passes these features directly to the learning algorithm. The feature selection process could be separated of the entire learning algorithm. Feature selection can be seen as a filter model, or it can be used to assess the performance of the learning algorithms like wrapper models. A classifier is then uses the finally selected features to predict the label according the classification task. Actually, feature selection for classification aims to focus on a subset of the minimum size of features dose not reduces the classification accuracy.

Feature selection techniques try to focus on the important and the most context related subset of features this process is conducted based on some evaluation functions. However,

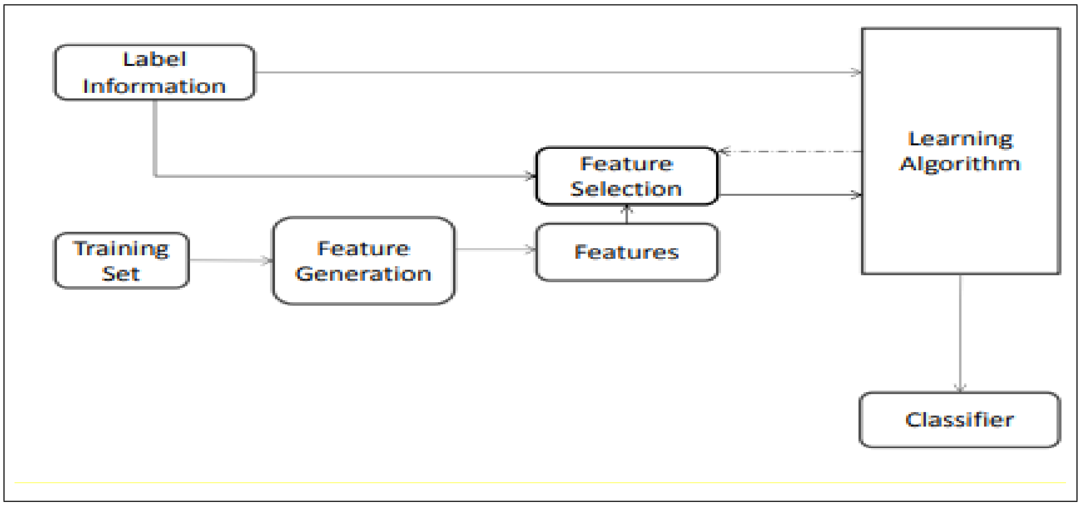


Figure 2.3 Features selection for classification framework

this process can impact the training time because it tries to identify the best set of features among all extracted features. However, random or heuristic based search methods proved to be effective in reducing the computational complexity through compromising performance. These methods applies a criterion to avoid the extensive search of subsets. Based on the feature structure, feature selection can be classified into three approaches: flat features selection methods, structured features selection methods and streaming features selection methods.

2.3.3 Opinion classification

This process aims at classifying text into certain class according to the opinion expressed in the text. There are two main Opinion classification techniques: lexicon based and machine learning.

2.4 Machine learning

Machine learning approach takes advantages of the famous machine learning algorithms to address the opinion classification challenges such as text classification task which utilizes linguistic features. Many of the existing opinion classification approaches are based on machine learning algorithms that uses a Bag-Of-Words (BOW) sentences representation as their basis. These approaches are fall into either supervised or unsupervised learning approaches.

2.4.1 Supervised learning

It aims to build classification models that assign class labels C (i.e. negative or positive) to a particular problem unit D (i.e. document, paragraph, sentence, phrase or aspect D= d1, d2, ..., dn). Every unit is represented in a form of features vector:$F = (n1(d), n2(d), \dots, nm(d))$. Where $X = x1, x2, \dots, xm$ is predefined set of the m features that can appear in d and $ni(d)$ the number of time xi occurs in d. The supervised learning methods take advantage of annotated datasets that contains high quality domain representative training data. In this approach we distinguish between probabilistic algoeithms such as Naïve Bayes, Maximum Entropy, and Nearest Neighbor , linear like Support Vector Machines and tree based as Decision Tree (DT) classifiers.

2.4.2 Probabilistic classifiers

NB is a very simple technique, based on Bayes theorem which assumes that a each particular feature value is unique and separated from the other feature values (Rich, 2001). Indeed NB assigns to d the class $c = *argmaxp(c|d)$ based on Bayes theorem described in the Eq. 3.1.

$$PNB(c|d) = \frac{p(d|c)p(c)}{p(d)} \tag{2.1}$$

However, the probability of a document assigned to specific category for a particular context maximizes the entropy of the overall classification system (Ratnaparki, 1997), thus, it's guaranteed that the biases are not introduced in the system. ME is used to estimate $P(c|d)$ as exponential as presented by Eq. 3.2.

$$PME(c|d) = \frac{1}{Z(d)} EXP(\sum_i \gamma_i c f_i, c(d|c) \tag{2.2}$$

note that: Z(d) denotes a normalization function, γ_i, c denote feature-weight parameters and f_i, c denote a function for feature, see Eq 3.3.

$$f_i, c(d|c) = \begin{cases} 1, & ni(d) > 0 \text{ and } \tilde{c} = \text{c} \\ 0, & \text{otherwise} \end{cases} \tag{2.3}$$

ME is work according to the principle of Maximum Entropy, despite that it identify the model with the maximum entropy value, it takes longer training time compared with NB.

Regarding Neural Networks methods, they are considered as the simplest and the best performing classification algorithms. The working principle of theses approach is based on the theory which assumes that the category of the new unseen in the training set instance is identified as the nearest neighbor instances from the training set (Abd-Elhamid, Elzanfaly, and Eldin, 2016). Therefore, the k-NN algorithm examines the k-closest instances from the training set with the new unclassified instance at hand to predict the correct class based on the class of the k neighbors. The nearness (distance) between two points in the dimensional space is measured using a distance function such as the standard euclidean distance function employed between two points in x-dimensional space, x denotes the number of attributes of the dataset.

2.4.3 Linear classifiers

They are margin classifiers such as SVM (Shawe-Taylor et al., 2000). This latest is considered as a binary classifier that tries to identify a hyperplan H detailed in the equation 4 that separates training instances vectors in one class from those of the other class, accordingly, this separation is as wide as possible as illustrated in fig. 2.4.

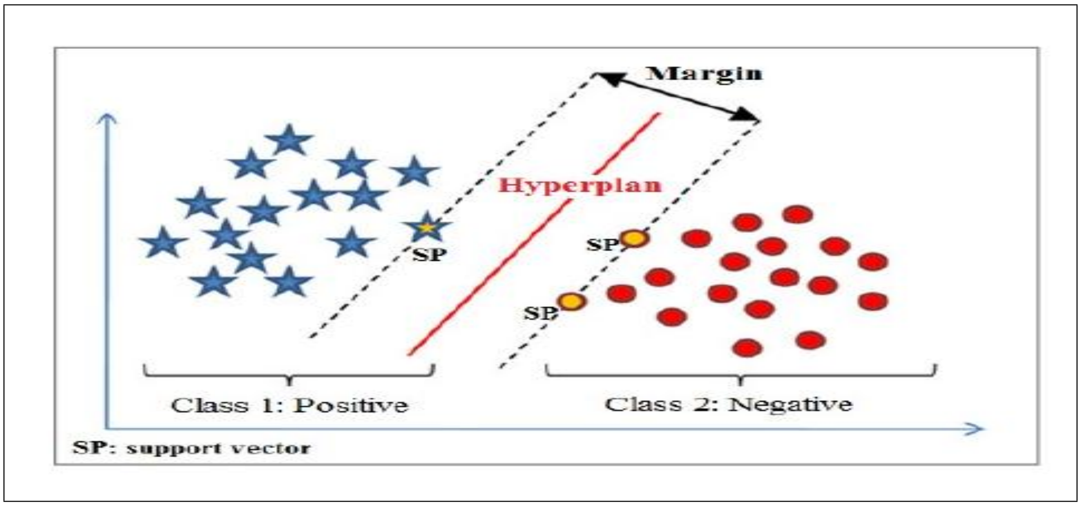


Figure 2.4 Support vector machine classification

where: H = $\sum_j \alpha_j c_j d_j$, $\alpha_j \geq 0$, Test classification is based on identifying that which side of the hyperplan H they are existed.

2.5 Artificial Neural Networks (ANNs)

Also known as connectionist systems can be simply defined as a computing system. The working principal of the Artificial Neural Networks have been inspired by the biological neural networks system in the human's brains. In fact, this system have an impressive ability to take advantage of training examples in enhancing their ability to automatically perform tasks, furthermore, these Neural Networks can generalize a task-specific training to wide range of other tasks (Cabreira, Tripode, and Madirolas, 2009) (Cabreira, Tripode, and Madirolas, 2009). The basic building block of these Neural Networks is sets of connected artificial neurons similar to the biological neurons in a humans brain (Aggarwal, Kong, et al., 2014), Connection between neurons is established using synapse which are used as a transmission medium over which the signals are transmitted to another neuron (postsynaptic) which processes the received signals and then the signals is transmitted to the nearest connected neurons. It is worth mention that every neuron in the system is assigned a state which is typically represented using real numbers fall between 0-1. Both synapses and neurons have also set of weight that can be learned and modified during the training process, these weights can influence the strength of the signal out of the current neuron.

Neural Networks architecture is organized as layers which can make diverse types of transformations on their inputs. Typically, signals propagate from the input layer (first) to the output (last) layer, signal may traverse the current layer several times (Harrag and El-Qawasmah, 2009). The main idea behind neural network algorithms is to effectively provide a solution to the problems using exactly the mechanism that a human brain uses. Later on, researches have paid more attention on simulating a particular humans mental abilities this efforts come up with several concepts including back propagation or sending the information in the recurrent direction and fine-tuning the network to match that information. Typically a neural network contains number of units ranging from a thousand to a few millions, in addition to millions of connections. However, the numbers of neurons and connection in neural networks are minimal compared a human brain, neural networks can outperform humans performance in the same task such as faces recognition, and games playing (Marinai, Gori, and Soda, 2005).

2.5.1 Deep neural network (DNN)

Is well known type of artificial neural network which consists of multiple layers I.e. hidden layers between the input layer and the outputs layer (Kacser, 1991), that is why called "deep" networks. Actually deep neural network algorithm tries to identify the optimal mathematical manipulation to calculate the probability of the output in each layer from the current input. For instance, the deep neural network which have been trained to detect the type of animal will scan over the image and compute the probability that the animal presented in the image is fall into a certain animals type. Deep Learning (DL) is recent advances in neural networks that emulates human learning behaviors using multi-layers neural architectures to handle a cognitive task. In contrast of conventional Machine Learning-based-LNP approaches (e.g. SVM) which depend mainly on hand-crafted features, DL models can automatically perform multi-level abstract features representation, which make a simple DL model achieve superior performance in any artificial intelligence tasks. Convolution Neural Network (CNN) has achieved promising results in computer vision and different NLP tasks including text classification, sentence modelling, etc (Yunfei Ma et al., 2011). CNN takes advantages of at least one convolution layer, fully connected, and pooling layers to capture local features and words relationships (Medrouk and Pappa, 2017), however, it fails in capturing long-distance dependencies. Recurrent Neural Network (RNN) is used mainly in text data classification due to its ability to capture long-term dependence, and to maintain the sequences of variable length data Alowaidi, Saleh, and Abulnaja, 2017. In RNN each neuron is connected to its previous one i.e the output of the current hidden layer is calculated using the current and previous neurons outputs. Despite that, RNN suffers from vanishing gradient problem.

DNNs is proved to be effective in modeling (provide a compositional models) of a complex non-linear relationships problems. DNN algorithm can effectively provide features composition and abstraction from lower layers using the extra layers which enable a simpler (fewer units) DNN architecture effective in modeling of a complex data problem than other similar shallow network. As illustrated in Fig. 2.5, typically, Deep Neural Networks architectures can be visualized as feedforward networks where the information and data streams only in one direction from the input layer to the direction of the output layer (Geoffrey E. Hinton, Osindero, and Teh, 2006). The learning mechanism of these networks depends mainly on creating virtual neurons map with associated random weights (numerical values). At the current neuron, the input (output of the other neuron) values and the weights of the current neuron are multiplied to generate the output of the current neuron in the range (0-1). Note that the network may incorrectly identify a certain pattern at the earlier stages, If that the case, the networks modifies its weights during the training (Kacser, 1991). The network influentially continue adjusting the weights till reaching a correct mathematical manipulation.

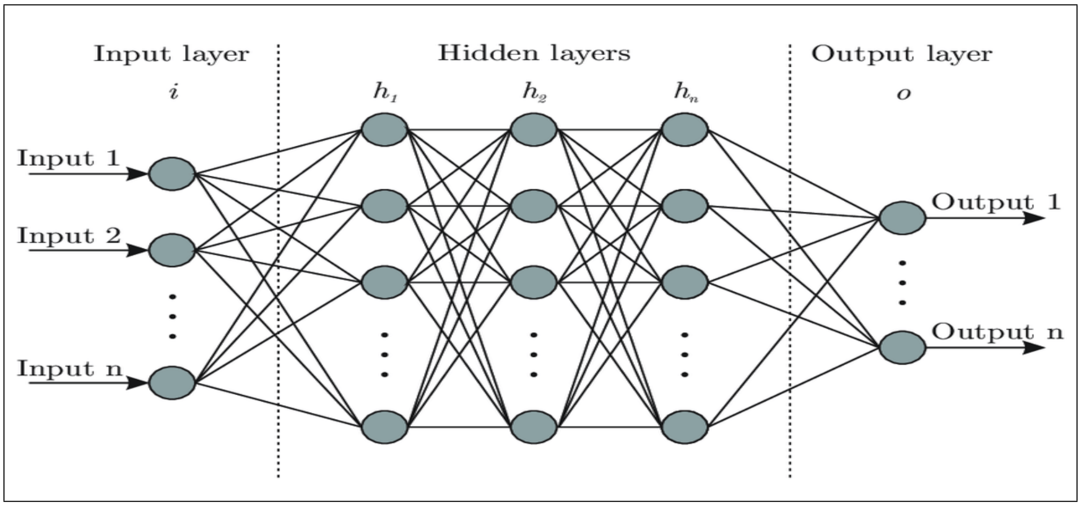


Figure 2.5 Artificial neural network architecture

Commonly used DNNs examples includes: Recurrent neural networks (RNNs) which are commonly used in sequential modelling due to its unique ability in capturing terms from any direction, in contrast to CNNs, RNNs are not able to extract features in a parallel way. However RNNs are considered as deep neural networks, for long data sequences RNN has problems of gradient vanishing or explosion. Another example is Convolutional deep neural networks (CNNs) which have presented a remarkable performance in computer vision, sentences modeling and automatic speech recognition (ASR) (Y. Kim, 2014a).

2.6 Words representation (Word2Vect)

In order to transform any NLP task into machine learning algorithms, text must firstly be transferred into corresponding vector representation. For this there are two vectorization algorithms. One-hot representation, in which very long vector is used to represent the words, the vector length is the same as the size of the dictionary used in the corpus. It uses only 1 and 0 weights. With One-hot representations it is not easy to depend only on words vectors to define the relationship between words. Another approach is distributed representation which has recorded the best performance in deep learning field. This method is based on mapping each word into fixed length vector, distributing these vectors to form the vector space Mikolov, Sutskever, et al., 2013. A word vector is described as a low dimensional vector representations that encode semantic features of words learned in unsupervised neural nets models on a very big text corpus. Word2vec is a neural network used to process the text before this text is received by deep-learning algorithms. It takes text corpus as an input and generates the word vectors as output. The vector representation of words is obtained after word2vec builds vocabulary from the training corpus. The resulting word vectors file could be used as features to deep learning algorithms. In this algorithm, the sentence words are initially represented in form of words matrix, then it transferred into vectors in an n-dimensional vector space. In this method, similar words are represented near each other in the vector space Gupta, Tiwari, and Robert, 2016. Moreover, with Word2vec features can be obtained without human intervention. Word2Vec can also perform effectively even when its input is an individual word. With this tool, very accurate predictions about a word's meaning can be obtained and the semantic relationship between words can be easily evaluated.

2.6.1 Continuous bag of words

Continuous Bag Of Words (CBOW) generates the word representation using the surroundings words (similar meaning words) in the vector space (Bojanowski et al., 2017). As illustrated in Figure 2.7, the CBOW architecture is proved to be effective in learning distributed vector representations that takes into consideration the syntactic and semantic relationships between words inn the sentences (Mikolov, K. Chen, et al., 2013). As presented in Figure 2.7, the continuous bag of words architecture receives the input words represented as a one hot vector. x denotes the surrounding context words, V denotes vocabulary size, C denotes the context, N denotes the number of hidden units (N.dimensional). y denotes the words representation vector which have been encoded during the training. By adapting the weights

W weights W0, the hidden layers learns to generate the words representations of the target words, this process in achieved during a backpropagation passes which aims at reducing the errors between the prediction and the target then the weights are updated accordingly.

2.6.2 Skip-gram model

In skip gram model, the input words and their associated context are used to in neural network training process. In this approach a unique identifier is assigned to each pairs appear of the same context which appear frequently. As shown in Figure 2.6 the Skip-gram model generates the representation of the sequence by predicting the context of another word in the sentence. Let $w_1, w_2, \ldots, w_T$ denotes a sequence of input words, here the skip-gram algorithms aims at maximising the average log probability as in Eq (2.4):

$$\frac{1}{T}\sum_{t=1}\sum_{-c\geq j\geq C, j\neq 0}\log(P(w_{t+j}|w_t) \tag{2.4}$$

where c is the size of the training context and T the amount of words in the sentence. The Bayesian component: $P(w_{t+j}|w_t)$ in Eq. 3.3 is defined as:

$$P(w_i|w_j) = \frac{e^{ut}w_t v_{wj}}{\sum_{i=1} e^{ut}w_t v_{wj}} \tag{2.5}$$

u_w denotes the representations vector and v_w denotes the word context. V denotes the vocabulary size. The probability of predicting word w_i using w_j is computed using the softmax function 2.5.

2.7 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is one of the important deep learing argorhtims developed for computer vision tasks (Gu, M. Wu, and C. Zhang, 2017), CNN is considered as the first supervised learning architecture based multi-layer network structure. Convolution

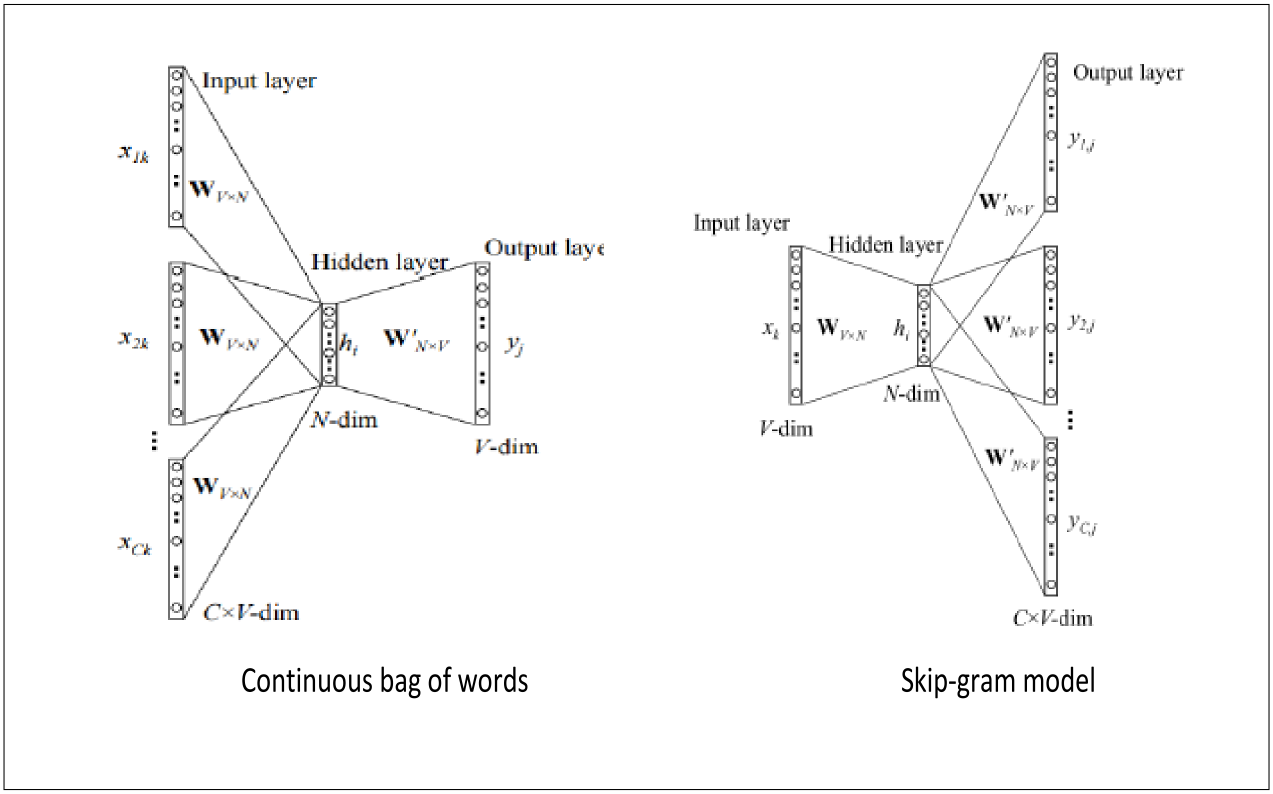


Figure 2.6 Continuous bag of words and Skip-gram models

Neural Network (CNN) has achieved state-of-the-art performances in computer vision and different NLP tasks including text classification Wint, Manabe, and Aritsugi, 2018, sentence modeling Kalchbrenner, Grefenstette, and Blunsom, 2014, semantic search query retrieval Y. Shen et al., 2014, etc. With much fewer parameters and connections, CNN takes advantages of convolutional layers to capture local features and words relationships Ombabi, Lazzez, et al., 2017. CNN is composed of at least one convolution layer, fully connected and pooling layers. However, CNN is not able to capture long-distance dependencies (L. Lei, J. Lu, and Ruan, 2019).

2.7.1 Convolutional Neural Network architecture

As illustrated in Figure 2.7 Convolutional operation contains three main operations. Initially, a set of linear activation is produced out of several the convolutional layer performs. Then, these linear activation is passed through a nonlinear activation function e.g. "rectified linear RLU" function. An finally type of a pooling function is used to adjust the output of the layer.

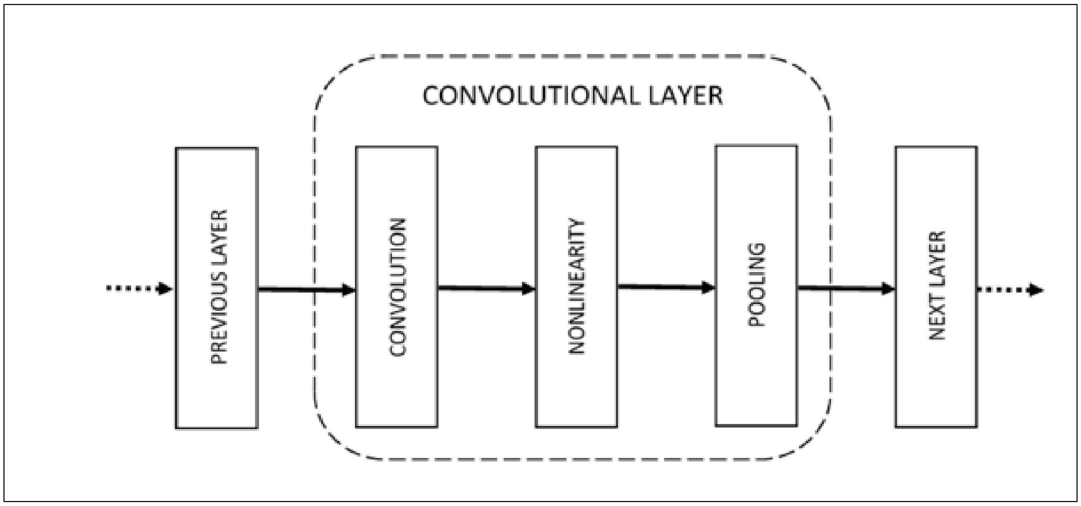


Figure 2.7 CNN layer processes

Convolution operation

The representation vectors generated by the words representation layer is received on by a 1-dimensional convolution layer with a specific convolution filter size c. With e as embedding size. The convolution kernel generates (c x e) features thought one-dimensional convolution operation. Multiple outputs is produced out of this layer due the use of multiple filters, CNN aims at maintain the words sequences of the filter size c, thus understanding the semantic relationship between words, for instance if c = four, then the layer takes four words at a time, so, forming a sense of four word combinations. Figure 2.8 presents a schematic illustration of a 2Dimension convolution operation. Letters in boxes denote the values combination of the input and of the kernel.

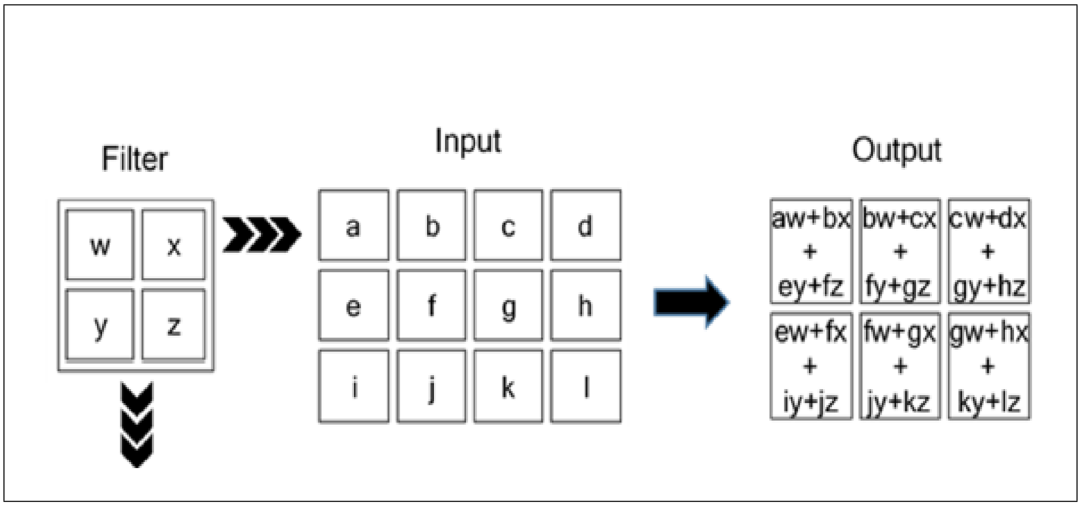


Figure 2.8 2-Dimensional convolution operation

Activation

After the convolutional operation a rectified linear unit (ReLU) is applied to network at this phase, after introducing the non-linearity into the architecture using a RELU activation function, the outputs takes the same shape as the input and the negative outputs are transformed into zero.

Max-Pooling

The main reason of adding max pooling just after ReLU activation function is to prevent overfitting while signals (features) propagates, also it aims to map the output of each filter into single stream with maximum number, then performing down sampling on the output. Typically, max pooling is performed based on town parameters filter size p (height dimension) and data width, max pooling implies reducing the height of the data according to filter size (p).

Dropout

After the max pooling, the output of each channels received by a dropout layer, this layer aims at randomly transforms a potion (specified d set) of the inputs into 0. At this layer, overfitting is prevented and generalization is presented to the network to eliminate focusing on particular part of the input. The corresponding output shape of dropout layer is the same shape as the input.

2.8 Long short-term memory (LSTM)

As special type of artificial recurrent neural network (RNN). Contrasting the standard feed-forward neural networks architecture, LSTM process the input data in two directions i.e forward and backward directions (connections), this architecture allows LSTM to be an effective general purpose system able to handle any task handled by Turing machine (Y. Lu et al., 2018). Unlike many other neural networks, LSTM can process the data that have a sequential nature such as speech or video, however, many other neural networks can only focus on a single data points (images). LSTM architecture has recently been applied to a wide range of application areas including handwriting identification and speech recognition (Chiu and Nichols, 2015). As illustrated in Figure 2.9 typical LSTM unit consists of an input gate, output gate and a forget gate to control the information flow through cells, in addition to a new added memory cell to selectively maintain the cells states values over long and short. LSTM architecture have been proved to be effective in processing and classifying the time sensitive data. LSTMs were introduced to prevent the limitation of data exploding and vanishing gradient encountered with the traditional RNN (Abdalraouf Hassan and Mahmood, 2017). Despite that, RNNs can maintain long-term dependencies in sequences of arbitrary length, during the training of vanilla RNNs and particularly over back-propagation, the back-propagated gradients vanishes. In order to handle this limitation, RNNs partially stacks LSTM units to prevent the vanishing gradient, because LSTM permits the gradients to stream unaffected. However, LSTM architecture still needs to deal with exploding gradient problem. The following block functions explain the working principle of LSTM :

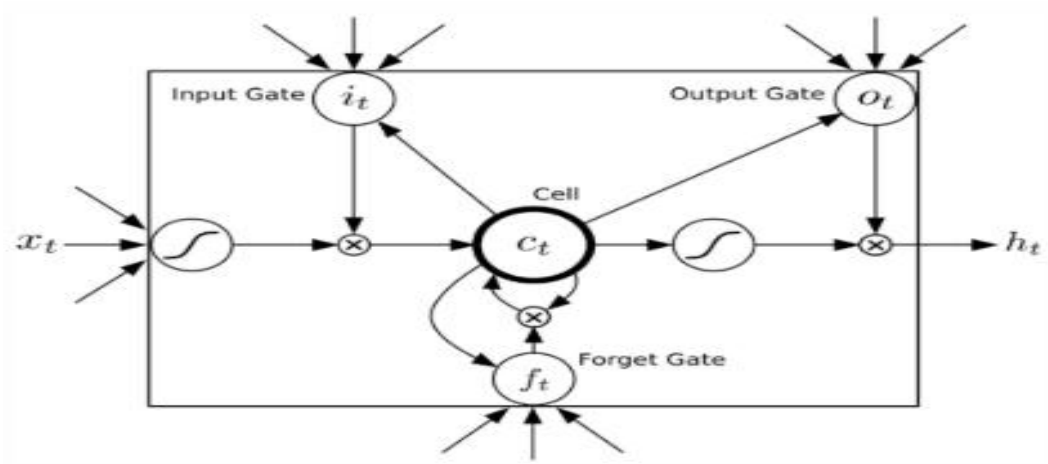


Figure 2.9 Long Sort Term Memory with a memory cell for text prediction

$$g(t) = \phi(W_{gx}x_{(t)} + W_{gh}h_{(t-1)} + b) \tag{2.6}$$

$$i(t) = \sigma(W_{ix}x_{(t)} + W_{ih}h_{(t-1)} + b_i) \tag{2.7}$$

$$f(t) = \sigma(W_{fx}x_{(t)} + W_{fh}h_{(t-1)} + b_f) \tag{2.8}$$

$$s(t) = (g(t) * i(t) + s(t-1) * f(t) \tag{2.9}$$

$$h(t) = s(t) * o(t) \tag{2.10}$$

2.9 Related Works on Users Interests Discovery

Recently, Text categorization have become an essential tasks of Natural Language Processing (NLP) with wide applications including interests or preferences prediction, and sentiment analysis both tasks are considered as a classification task. In this section we present state of the art in user interest discovery on different languages including Arabic.

Recently, users interests discovery have gained much research attention, different approaches have been proposed for different purposes such as recommendation systems, user's hidden information discovery (e.g preferences), predicating political trends in public opinions, users modeling and categorization, communities detection etc. Almost all of the proposed approaches are unsupervised which focused on using NLP, linguistic analysis, lexicon based to detect the topics of individual user interest. A very few partially supervised approaches are proposed to predict user interests, those approaches are not efficient as confirmed by Hai et al., 2017. Several approaches are competing to effectively discover and classify user's interest, text mining and NLP approaches (e.g. Seo and B. T. Zhang, 2000) are commonly used in users interests discovery, they depend on using weighted keywords (explicit indicators) to classify the text into certain topic label, however keywords vagueness and terms ambiguity (one word expresses several emotions) can reduce the classification accuracy (Darabi and Tabrizi, 2017). Besides, semantic relationships of the words are not considered (Yunfei Ma et al., 2011). At the other side, some existing studies (e.g. Liang and Lai, 2002) focus on tracing user's browsing behaviors (implicit indicators) (e.g duration) and user's browsed contents (e.g viewed web pages) to detect users interests. However, these methods give inaccurate predictions because users may view web pages but he/she is not interested in the contents of the web pages. Also, these methods confined the users interest to the viewed web pages. Liang and Lai (2002) have proposed users interest classification approach based in structure analysis, reader profile analysis, and learning structure analysis. WordNet is exploited to calculate words frequencies. Report reading time is considered as the interest level of that report. Bhargava, Brdiczka, and Roberts (2015) have introduced a novel approach which focused on user's profiles and activities to predict use interests. Several NLP techniques are utilized: semantic relatedness for implicit interests detection, and NER, social tagging, document categorization features for fine-grained interests modeling. Pennacchiotti and Popescu (2011) have proposed an approach to detect tweeter users tendency and attributes based on users behaviors, network structure and the linguistic contents. They used Latent Dirichlet

Allocation, lexical analyzer to obtain set of features. K. Xu et al. (2018) have proposed a novel framework to discover users preferences based on interest and social connections, they exploited a unified probabilistic topic model to capture interest and social topics. Then a community-based approach is employed to perform recommendations, N most likely followers are generated through Matrix Factorization.

Another types of studies have investigated both methods: network structure and text data as network contents, K. Lee et al. (2011) have proposed tweeter topics classification model based on Bag-of-Words for text-based and network-based classifications, TF-IDF weights were employed with NB Multinomial classifier for text-based classification. Secondly, top similar topics are identified for the topic at hand based on the number of influential users, then a decision tree learner is exploited for final classification. J. Kim et al. (2014) have presented a model for user's interests prediction based on Term Frequency (TF) of nouns, and Likes. TFs of nouns, and "likes" are used to calculate the interest weights which are later been ranked descending. Final user interest is determined when interests weightings exceed a specified threshold.

Another unsupervised Traditional methods (lexicon-based) focused on using keywords with associated weights for users interests classification and users modeling. Darabi and Tabrizi (2017) have presented a novel methodology for users interests detection and preferences management. In this study, users model is constructed based on domain ontology. With the proposed Interest Extractor, the frequently used words were determined using bag of terms, then Porter Stemming is employed, WordNet is employed to detect the highest frequency terms. Mangal, Niyogi, and Milani (2013) have investigated topics classification and sentiment analysis in tweeter. For SA Stanford coreNLP is exploited. The sentiment score is computed with sentiment property with 0-4 range. Using 36 taggers and semantic similarity and relatedness measures from WordNet, tweets are categorized into certain topic label. However, the vagueness and ambiguity of terms and keywords are the major shortcoming of these methods lead to poor performance particularly when dealing with users interest classification or sentiment analysis. Liu et al. (2018) have presented a novel hybrid deep learning model aimed to predict the type of PoIs/items of which a user is interested with (will buy/visit). This model is featured with stacked autoencoder-based deep model for social influence learning, and stacked LSTMs for user interests sequential learning. H. Xu et al. (2016) have proposed a novel topic-based semantic word embeddings with two multimodal CNN. LDA was used to compute topic-based words relationships.In this study, two CNN architectures are described i.e CNN-channel and CNN-concatenation for addressing: biomedical articles indexing, clinical text fragments, newsgroups classification. Fornaia et al. (2015) have proposed multi-agent driven system for user interests and behavior analysis. To find the categories of the users, profiling agent with ANN technologies was employed, it retains relevant data of given user profile. Radial Basis Probabilistic Neural Networks (RBPNN) is employed for users categorization. There are another type of works which aim to discover user's tendencies using combination of different type of social data such text and images together: Hong, C. Choi, and Shin (2018) have proposed a method for user interests categorization based on text and images, then providing follow suggestions based on their interests similarity. Convolutional neural network and hierarchical topics of interests categorization were used to learn and classify text and images features. Table 1.1 summarizes and compare the state of the art in English text user interest discovery approaches.

At the other hand, there is lack of research in the field of users interests discovery for the Arabic language, however, in this section we presets recent state of the art of different Arabic NLP and ML techniques: Mourad Abbas and Daoud Berkani (2006) have performed Arabic topic identification based on TF-IDF and SVM in the field of Arabic topic identification. Each document is represented as a vector, by combining TF(w,d) and IDF(w) the document's topic is determined as the topic with the highest similarity with the document. Also Mourad Abbas, Smaïli, et al. (2017) have proposed a novel technique for topic identification based on triggers which are defined using the Average Mutual Information. Topics and documents are presented using triggers which are a set of words that have the highest degree of correlation. Then, based on the TR-distance, the similarity is calculated between triggers to identify the document's topic. Zrigui et al. (2012) have presented a novel hybrid algorithm for topics identification. The proposed algorithm is based on the stacking LDA before SVM, the main aim of this combination is to minimize the feature dimension. The LDA technique is exploited to classify documents, and the SVM is employed to determine class label. Kelaiaia and Merouani (2013) have presented another approach of using LDA in topic identification. They depend on topic models to determine the documents distribution over topics. Koulali, El-Haj, and Meziane (2013) have used automatic text summarization for Arabic topic identification. Documents summaries is generated using Gen-Summ. Then cosine measure is applied to determine the similarity between the documents summaries which represented as TF-IDF and the corresponding vectors of topics and documents summaries which are represented by TF-IDF. Furthermore, Koulali and Meziane (2014) have proposed to perform Arabic topic identification using named entities. The main aim is to reduce the vectors dimensionality by using the segments bounded by named entities pairs. Then, the similarity between topics and documents is determined using mutual information. Hmeidi, Hawashin, and El-Qawasmeh (2008) have performed Arabic topics categorization based on two machine learning algorithms i.e K nearest neighbor (KNN) and support vector machines (SVM). The obtained results proved that these are efficient, SVM has generate more accurate predictions. Cheng and Soon (2006) have proposed to use back propagation neural network (BPNN) for topics classification. BPNN is enhanced to be more effective for this task. It is proved that the improved BPNN model achieved high performance on standard Reuter-21578 dataset. Soucy and Mineau (2005) have proposed a novel approach for text categorization method (ConfWeight) based on statistical estimation of importance word. The proposed approach is evaluated in three commonly used data sets: Reuters-21578, Reuters Corpus Vol 1 and Ohsumed. Table 3.1 summarizes and compare the state of the art in Arabic text user interest discovery approaches.

2.10 Related works on sentiment analysis

Typically, Sentiment Analysis can be conducted using three approaches i.e. supervised, unsupervised, hybrid approaches. The supervised approach is carried out using machine learning algorithms such as Support Vector Machine (SVM), Naïve Bayes (NB), Maximum Entropy, Artificial Neural Networks (ANN) and K-Nearest Neighbor (KNN). Ouyang et al. (2015) have proposed a framework based on pre-trained Word2vec and 7-layers CNN architecture for SA, Generalizability was improved using Parametric Rectified Linear, Normalization, and dropout unites. this model achieved better performance than RNN and Matrix-Vector Recursive Neural Network. Also (Ouyang et al., 2015) have proposed a combination of CNN and LSTM models for sentiment representation. For words embedding pre-trained word2vec was exploited, then CNN was incorporated for textual local features extraction, the outputs of the CNN are passed into two-layer LSTMs for context-dependent features management. Medrouk and Pappa (2017) have Proposed multi-languages deep learning model for textual data sentiment analysis, based on CNN with N-gram features level, Single input layer was designed to receive multi languages texts. Multi-lingual dataset which contained three languages was built to evaluate this mode. RNN also was used widely for SA. Al-Smadi et al. (2018) have proposed to perform aspect-based sentiment analysis using two implementations of LSTM i.e Bi-LSTM with CRF classifier to extract aspect Opinion Target Expressions (OTEs), aspect-based LSTM to perform aspect sentiment polarity classification, aspect-OTEs were considered as attention expressions for the sentiment polarity classification. Also RNN was utilized by Preethi and Krishna (2017) to enhance places recommendations based on SA. Experimental results shown the improved performance of the proposed approach in sentiment classification which have led to better recommendations. R. Ghosh, K. Ravi, and V. Ravi (2016) have proposed to incorporate two layers Restricted Boltzmann with Probabilistic Neural Network for sentiment classification, RBM was used to perform dimensionality reduction, and TF-IDF to represent the input data. The proposed model outperformed the other models. Taj, Shaikh, and Fatemah Meghji (2019) have utilized lexicon-based approach for textual data sentiment classification. TF-IDF was used to determine and assign weights to frequently used words in the documents, then WordNet and SentiWordNet lexical databases were employed to assign sentiment scores to the keywords. also X. Li et al. (2017) have presented a novel HNN architecture which combines Word2vec, RSM, and BTM models for social emotion classification, this HNNs were trained to obtain semantic features using Latent semantic machines and regularized transfer learning models. Sasmita et al. (2017) have introduced an unsupervised approach for aspect-based sentiment analysis. Tow subtasks were performed, firstly, aspect extraction in which sets of indicator words were constructed using seed words, each set corresponds to a different aspect, then the extracted pronoun or noun is compared against the indicators words. Secondly, an opinion lexicon was used to determine the sentiment orientation of a particular opinion term. A semantic Arabic Twitter Sentiment Analysis approach was introduced by Alowaidi, Saleh, and Abulnaja (2017). Arabic WordNet was utilized to enhance tweets representation. NB and SVM classifiers were evaluated with several words representation approaches on Arabic tweeter corpus. Classification performance can be improved by using the proposed representations concepts. The following tables summarize and compares the reviewed state of the art researches in these areas.

2.11 Limitation of the existing approaches

In SNs users tend to follow particular community or other users according to their own topics of interest. Therefore, discovering users interests have significant roles in user-centric applications and personalized web services development to gain more personalized information interest with wordNet. Moreover, consumers are interested to receive recommendations based on their areas of interest, an ubiquitous application which depend mainly on user interests discovery is recommendations systems which is incorporated with a wide range of e-commerce, online news, SNs, and governments platforms to provide users with huge options of new or unknown releases of products/services as they need, this could remarkably improve users experience and satisfaction amazon, Such as recommending relevant book from Amazon to users whose interested in reading, or suggesting a furniture to home decoration interested users. In fact the revenue of social media websites comes mainly from online advertisements, thus, users interest identification enable to target specific audiences with personalized advertising based on their discovered interests Lewenberg, Bachrach, and Volkova, 2015. Furthermore interests and sentiment orientations detection have a great influence in user's hidden information detection (e.g preferences, age), political trends predication, and community detection, etc. Despite that, all social networks analysis, sentiment analysis, NLP and unsupervised machine learning approaches are competing in user's interest detection, Text mining and NLP approaches are commonly used to infer user interests, they depend on using keywords (explicit indicators) each assigned weights, however keywords vagueness and terms ambiguity (one word expresses several emotions) can reduce the predictions accuracy. Besides, semantic relationships of the words are not considered Yunfei Ma et al., 2011. On the other side, some existing studies such as Seo and B. T. Zhang, 2000 focused on tracing user's browsing behaviors (implicit indicators) (e.g duration) and user's browsed contents (e.g type of web pages viewed) to detect the users interests and orientations. However, only the previous users' interests can be obtained, predictions are not actual and not precise because users may view web pages but heis not interested in the contents. Meanwhile, these methods confined the users interest to the viewed web pages. Unsupervised approaches are commonly used in SA. However, keywords vagueness and ambiguity can decrease the accuracy of predictions. These approaches cannot consider the semantic relationships between words in the sentences. For Arabic sentiment analysis, unsupervised approaches cannot be effective due to the numerous words from several dialects to be included in the lexicons. Also, it is observed that using only CNN or using only LSTM is inadequate to achieve the desired results on Arabic sentiment analysis (Q. Huang et al., 2017), this is because CNN fails to maintain long-term dependencies, and LSTM is weak to capture local features.

Falling to provide the correct users interests and sentiment orientations have a direct impact in providing right recommendation to the customers or consumers, thus falling to provide precise and accurate answers to such questions:

- How do the company policies impact customers perception of the brand?
- What do customers think of the products and brand?
- Are customers happy with the service?
- What do customers like about the brand's competition?

Therefore, it is challenging to capture users sentiments and interests orientations from textual data regardless deep considerations of syntactical and semantic rules of the language. Also meanings of words are strongly tied to the context, long and short distances dependencies between words in sentences multilingual. User's interests are not always expressed explicitly.

Chapter Three

Research Methodology

As this research is based upon constructive research and action research methodology, this chapter presents and discusses the adopted research methodology for the proposed users interests discovery and sentiment analysis framework. It begins with providing an overview of the theoretical model, then providing an overview of the research framework, then presents the constructed New Constructed multi-domains Arabic Database. Finally, the experimental environment is described along with the performance measurements and evaluations. Figure 3.1 illustrates the constructed research methodology processes.

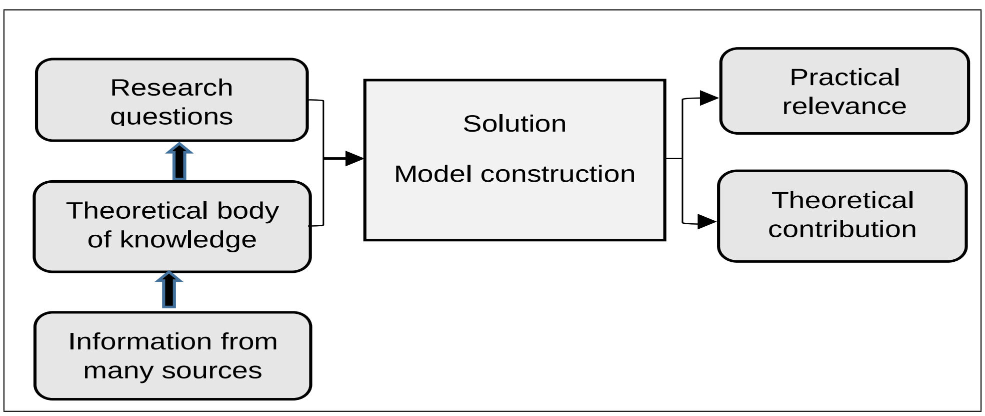


Figure 3.1 Constructive research methodology process.

The main research question is: How to best perform topics classification and sentiment analysis for individual user in social networks based on textual data ?. The thesis project is divided into five phases with their own sub research questions. Answers to these questions will lead to a conclusion to the main research question.

3.1 Overview of the theoretical model

3.1.1 Phase 1: investigate the current state of opinion mining based techniques

The goal of this phase is to clarify the current state of opinion mining based techniques and practice and get an overview of their processes. Therefore, the questions of this phase are:

- Question 1: What is the current state of the opinion mining based techniques ?.
- Question 2: How opinion mining based techniques processes look like ?.

A theoretical body of knowledge is created using a literature survey from peer-reviewed research and published literature. The literature survey is presented in Chapter Two. The developed theoretical model seeks to be the solution to the main research question.

3.1.2 Phase 2: Constructing of the theoretical model

using constructive research, this solution is built in form of a model from the existing theories which created through a comprehensive survey of a diverse sources including peer-reviewed research and published literature. Then the model should be tested for its practical relevance and its theoretical contribution. The theoretical body of knowledge is presented in Chapter Two. It is in this phase the theoretical model is constructed, which is considered to be the solution to the main research question. The goal is to identify best performing techniques for opinion mining based textual data.

- Question 1: What is the best techniques for textual features representation?
- Question 2: What is the best deep learning algorithm for features extraction and selection?
- Question 3: How to maintain the contextual information and terms sequences?

Figure 3.2 shows the constructed theoretical model for users interests discovery and sentiment analysis

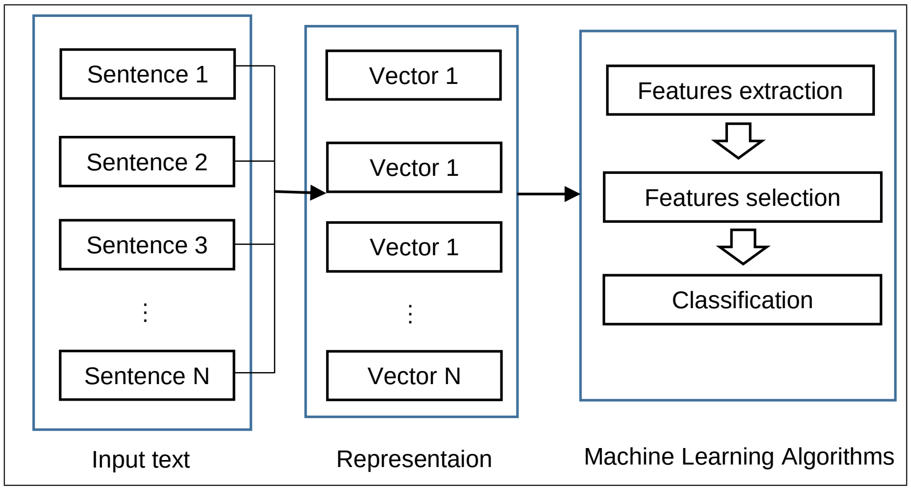


Figure 3.2 Constructed theoretical model

3.2 Overview of the research framework

The action research methodology which combines theory and practice used to construct the proposed framework and practical evaluation. Figure 3.3 illustrates the action research process.

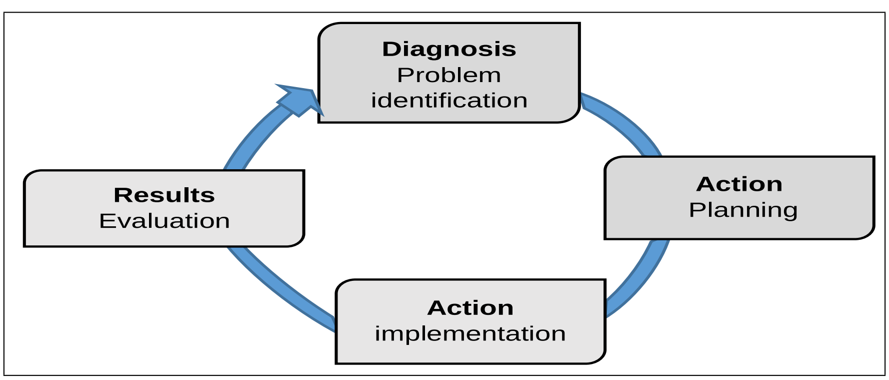


Figure 3.3 Overview of the action research methodology process

Actions then are taken to construct the framework. Finally, the results are evaluated.

This process is done iteratively and the phases provide feedback to each other. As the main focus of this research is to propose a novel deep neural network based system for users interests discovery and sentiment analysis. The proposed system intends to perform two tasks: firstly, to identify the topics of interests such as "sports", "restaurants", etc for the individual user, secondly to perform sentiment analysis which determines users inclination over the detected topics. Both tasks are considered as a classification task. The proposed takes full advantage of FastText model to construct vectors representation of the input sentences, also it incorporates WordNet knowledge base FastText to increase the coverage and the initialization of the embedding space and to embed the semantic information of words, If FastText fail to generate the corresponding word vector for a particular word, then WordNet can identify a similar word (synonyms), the generated words vectors are captured by a one-layer Convolutional Neural Network (CNN) to perform an n-gram local-region features and information extraction, then feature maps generated by the CNN are captured by a Recurrent Neural Network (RNN) for contextual information extraction, and to maintain the long-term dependency of the original sentences. Finally, a linear Support Vector Machine is used to classify the sentences into a certain label according to which tasks the architecture was trained.

3.3 Deep Neural Network Model for Users Interests Discovery and Sentiment Analysis

In this section we propose a multilingual deep learning architecture namely Deep-InterSent for automatic users interests discovery (UID) which identifies the user's topics of interest such as “sports”, “restaurants”, etc, furthermore, Deep-InterSent is intended to perform sentiment analysis (SA) which determine the users inclination over the detected topics whether it is “positive” or “negative”. Both tasks are considered as classification task, therefore we decided to train the same architecture to perform both tasks separately. We decided to train this architecture on English and Arabic languages.

3.3.1 Proposed deep InterSent framework architecture

The proposed DL architecture takes advantage of a new semantic layer which combines FastText model to obtain words vectors representation of the input sentences and WordNet knowledge bases to incorporate the semantic information of words at the embedding layer, then traditional convolutional neural network is used for n-gram local-region features and information extraction, the extracted features are fed into two layers LSTM to handle long-term dependency. Finally SVM classifier is exploited to determine the final classification labels according to which tasks the architecture is trained on. Fig. 3.4 illustrates the overall processes of the proposed architecture to perform user interest classification and sentiment analysis tasks, Fig. 3.5 visualizes in details the fundamental architecture and processes of Deep-InterSent.

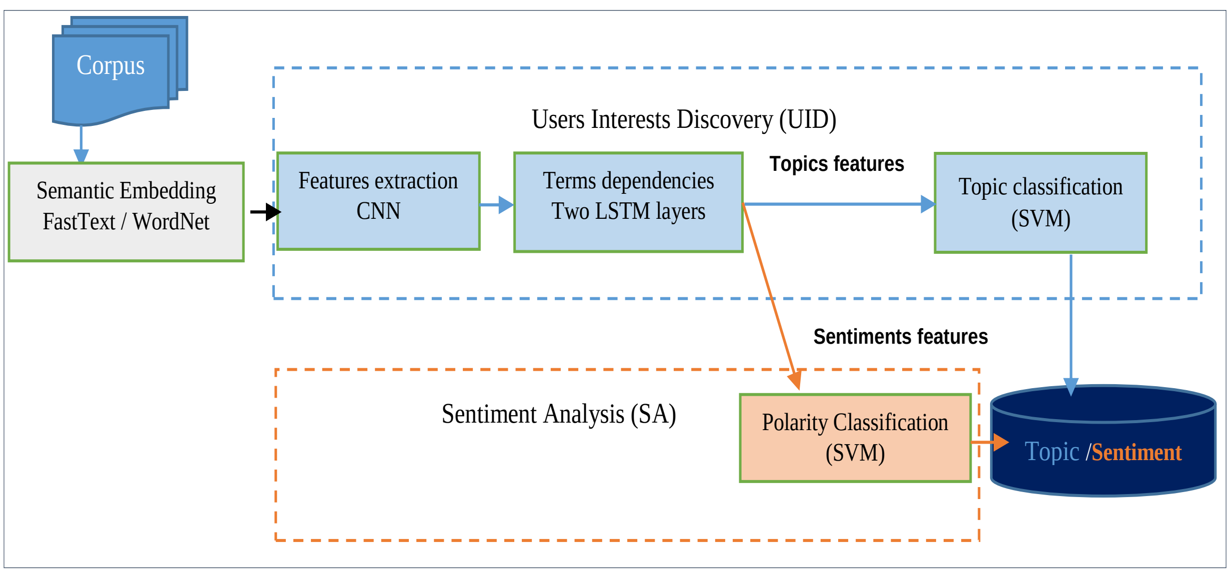


Figure 3.4 Overall processes of the proposed InterSent framework

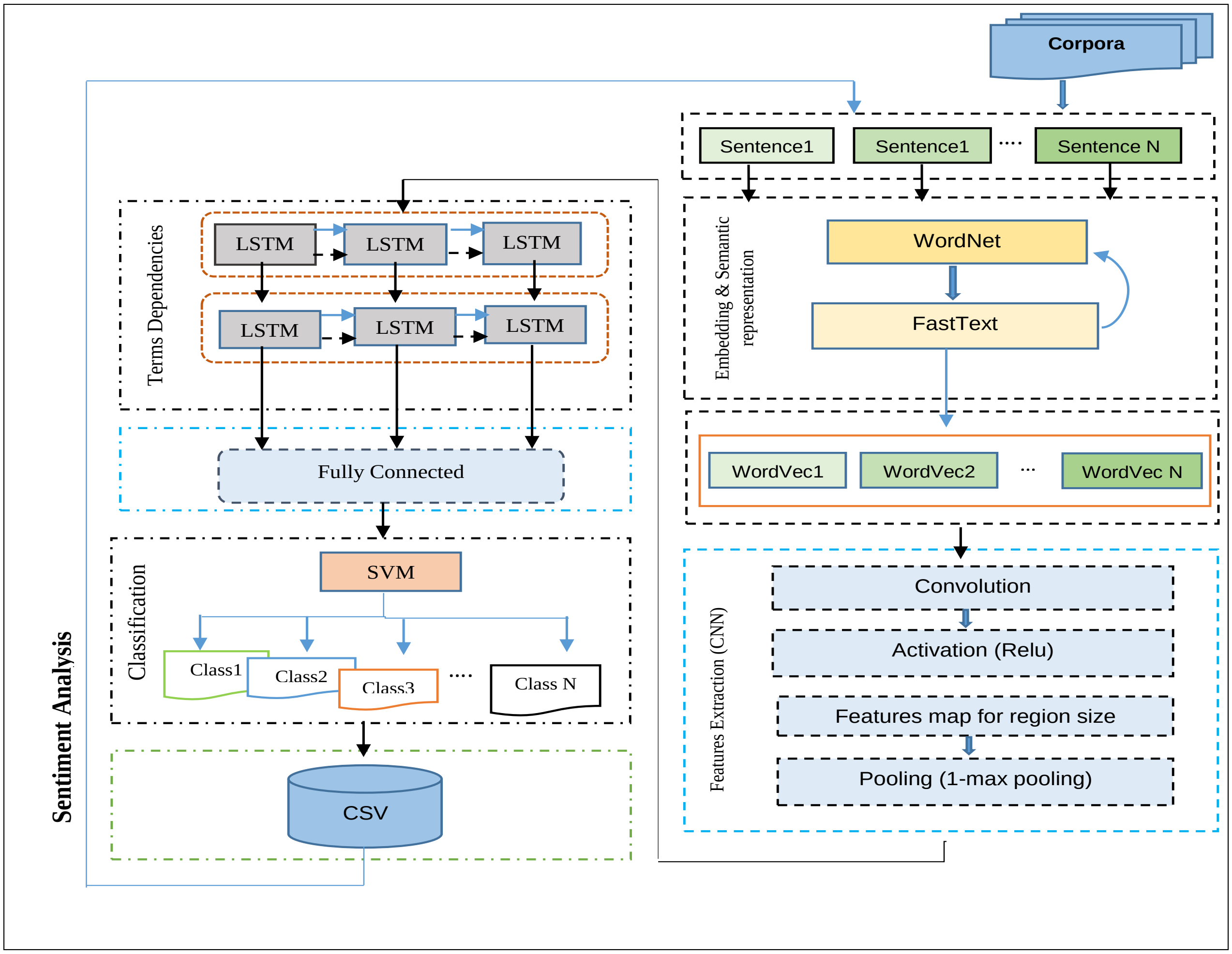


Figure 3.5 Flow Diagram of the proposed InterSent framework

3.3.2 Words embedding and semantic representation

This study take advantage of using the recent FastText model which introduced by Facebook researchers team (Mikolov, Grave, et al., 2017) for word vectors representation which is trained on wide range of languages including English and Arabic. As in word2vev FastText provides two models: Skip-gram model which is used to predict a target word using the closed neighboring words, while CBOW uses the surrounding words in the context to predict the target word, both methods generates text file which contains numerical representation (vectors) of the learned words. FastText has great capabilities to generate the corresponding vectors for the Out-Of-Vocabulary words (OOV) and rare words (Bojanowski et al., 2017) because these words still share their n-gram characters with other common words, however both word2vec and Glove cannot generate vector representations for OOV words (Wint, Manabe, and Aritsugi, 2018). In our experiments we used FastText skip-gram model in which each word is represented as a bag of character n-grams, sentences are then represented as a summation of these words vector. FastText skip-gram was run in default configurations: 100 dimension vector space, sub-word size is 3-6 characters.

WordNet

WordNet is a large lexical database which is constructed by George A Miller, 1995, then it have been deveoped for many other languages sucg as French, German, Arabic (Black et al., 2006), and even Arabic dialects (Cavalli-Sforza et al., 2013), WordNet groups the related words into one set of cognitive synonyms called Synsets, in WordNet the lexical categories such as verbs, nouns, adjectives are connected using conceptual-semantic relations, thus words meaning can be given via these associations. In each synset, senses are classified as synonym sets, however the synonym sets contain several keywords, they give the same meaning. WordNet is provided for several languages separately,

In this research, to increase the coverage and the initialization of the embedding space, we proposed to incorporate WordNet and ConceptNet just before FastText to perform semantic information injection. If FastText fail to generate the corresponding word vector for particular word, then WordNet and ConceptNet can identify a similar word (synonyms) for that word as shown in Algorithm.1.

WordNet is provided for several languages separately, therefor, we used WordNet Lexical Database (G.A. et al., 1993) for English language text, and we used Arabic WordNet (AWN) lexical and semantic database (Black et al., 2006) for Arabic language text. Arabic WordNet (AWN) was also used by (Alowaidi, Saleh, and Abulnaja, 2017) to enrich the text representation with concepts.

Algorithm 1 Semantic words vectors representation

Input: Text sentence. **Output**: Vectors representation
for each input sentence s $\in$ {1, ...,N}
do generate words vectors
if FastText (sn) $\neq$ NULL **then**
xn = FastText (sn)
else if WordNet (sn) $\neq$ NULL **then**
sn = WordNet (sn)
xn = FastText (sn)
else xn = Random(sn)
end if
end for

3.3.3 Features extraction

In this research, we used the CNN architecture proposed by kim Y. Kim, 2014a to perform features extraction for both UID and SA tasks separately. This CNN architecture uses 2-dimensional matrix input layer to receive sequences of words represented as vectors of the same length. Figure 3.6 presents the proposed neural networks architectures including the CNN and LSTM, the basic building blocks of the CNN architecture contains: Convolutional Layer with three different region size filters, and fully-connected layer. Convolution operation is applied when the filters slide over the words vectors from the sentence matrix, according to the filter size few words are taken at a time to obtain local features, note that the filters are sharing the same weights and parameters.

To explain the working principle of the CNN, let $X_i \in R^k$ denotes to the k-dimensional vector corresponding to the i^{th} word in the input sentence, the input sentence with length (n) is represented as a concatenation of its words vectors using Eq.3.1. Zero padding is applied to the sentences with length less than (n).

$$X_{1:n} = X_1 \ \oplus \ X_2 \ldots . \ \oplus \ X_n \tag{3.1}$$

Where $\oplus$ denotes a concatenation operator. Let $X_{i:i+j}$ represents a concatenation of the

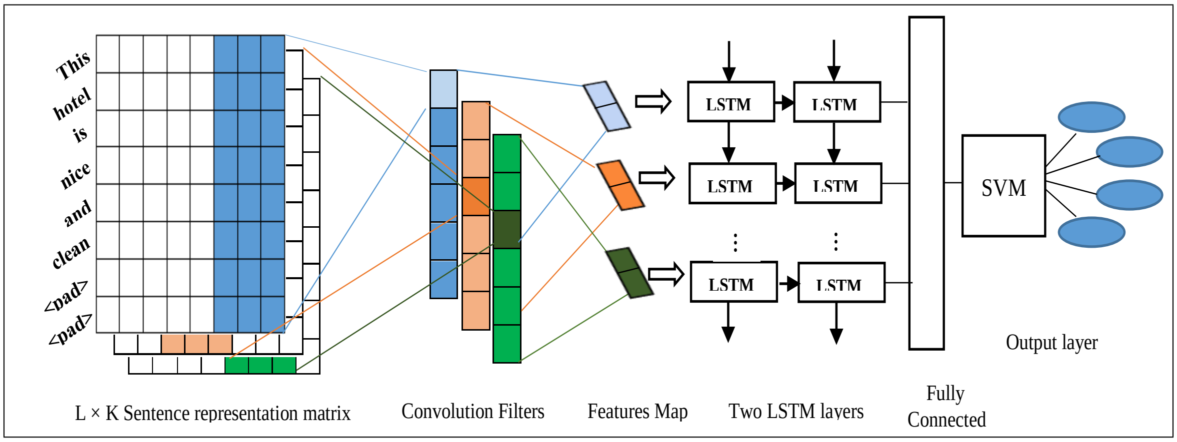


Figure 3.6 Deep CNN-LSTM Arabic-SA NNs Architecture

words $X_i, X_{i+1,} \ldots X_{i+j}$, the convolutional layer applies its convolution filter $W \in R^{hk}$ with h window size on each window of k width word to generate new $h \times k$ features matrix, $X_{i:i+j}$ denotes the basic element from the i^{th} to the $(i+j)^{th}$ corresponding to the feature matrix from the i^{th} line to the $(i+j)^{th}$ line of the current sentence vector. for exmple a feature C_i (i^{th} feature value) can be generated using convolution process over a window of words $X_{i:i+h-1}$ using Eq. 3.2:

$$C_i = f(W.X_{i:i+h-1} + b). \tag{3.2}$$

Where $b \in$ R and it refers to bias term, f is a nonlinear activation function (i.e. sigmoid and hyperbolic tangent). b and W are learned during the training. The convolution filter convolves over each window of words in the input sentence $X_{1:h}, X_{2:h+1}, \ldots X_{n-h+1-n}$ to generate local features map using Eq. 3.3.

$$C = [C_1, C_2, \ldots C_{n-h+1)}] \tag{3.3}$$

with $C \in R^{n-h+1}$.

In this architecture, max pooling function which performs features sampling is not applied because we need to maintain the original sequences of the feature maps to be directly fed into two layers LSTM.

3.3.4 Capturing high level dependencies using LSTM

LSTM can efficiently control the information through preventing vanishing gradient and capturing long-term correlations in arbitrary length sequences (Abdalraouf Hassan and Mahmood, 2017). LSTM architecture contains a new added memory cell to selectively maintain the information for longer time without getting degenerated, in addition to input, output, forget gates as illustrated in Figure 3.7. Hochreiter and J?urgen Schmidhuber, 1997 proposed LSTM architecture to handle term dependencies over long period of time, later, several versions of LSTM have been introduced. In this study, we proposed to stac two-layers LSTM after the CNN.

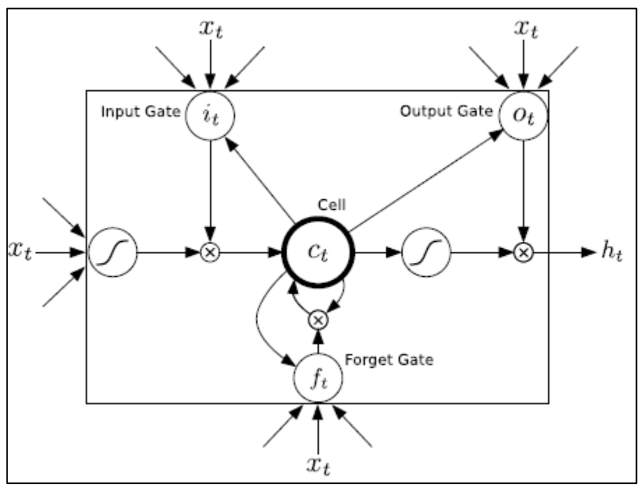


Figure 3.7 LSTM unit structure used to maintain the sequentiality of sentences

LSTM applies recursive operation to control and process the coming sentence vectors, the recursive execution of the current cell block is performed using the old hidden state (h_{t-1}) and the current input x_t, where (t) and ($t-1$) refer to the current time and the former time respectively. Now i_t denotes to the input gate, f_t denotes to the forget gate, o_t denotes to the output gate, and $\tilde{C}_t$ denotes to the state value of the current memory cell at a time t in cell block. The working principle of LSTM is explained using the following equations: Using Eqs. 3.4 and Eq. 3.5 the values of i_t, and $\tilde{C}_t$ are computed for the memory cells states at a time t.

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{3.4}$$

$$\tilde{C} = tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{3.5}$$

Eq. 3.6 calculates the activation value f_t of the forget gate at time t:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{3.6}$$

Eq. 3.7 calculates the new state values C_t for the current memory cells at a time t:

$$C_t = i_t * \tilde{C} + f_t * C_{t-1} \tag{3.7}$$

Using C_t, the values of the memory cells output gates are computed, then their outputs using Eqs. 3.8 and 3.9 respectively.

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o c_t + b_o) \tag{3.8}$$

$$h_t = o_t * tanh(C_t) \tag{3.9}$$

Where x_t refers to the input of the memory cell at t. W_i, W_C, W_f, U_i, W_o, U_C, U_f, U_O, and V_o are the weights matrices. b_i, b_f, b_c, b_o are the bias vectors. σ denotes to logistic sigmoid function, o refers to element-wise multiplication. During training the model learns the values of W_i and U_i. The values of f_t, i_t and $o_t \in [0, 1]$ the output of the first LSTM layer is passed to the second LSTM layer where the final deep representation of the sentence is constructed. In this architecture, Max Pooling function which performs features sampling is not applied because we need to maintain the original sequences of the feature maps to be directly fed into two layers LSTM.

3.4 Validation of the Framework's practical relevance

3.4.1 New constructed multi-domains Arabic database

Due to the complex morphology of Arabic language and the lack of high quality and appropriate free accessible Arabic textual corpora, minimal researches efforts have been conducted in Arabic linguistics and NLP research fields (Saad and Ashour, 2010), therefore, in this research we have built Arabic corpus for user interests discovery and sentiment classification tasks, Figure 3.8 shows the main steps to construct this dataset.

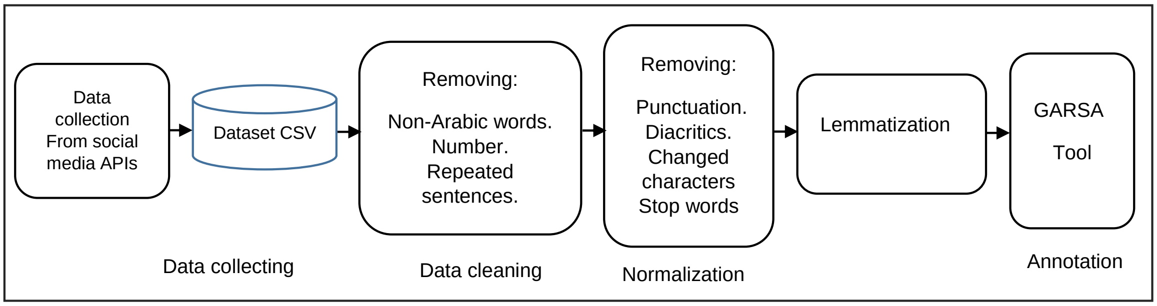


Figure 3.8 Overall processes of dataset preparation

Data collection

Arabic reviews acquisition from the internet is a very hard task, due to the minimal activity in the Arabic based e-commerce and reviewing websites resulting reduced amount of pure Arabic reviews on the internet. In this research we focused on the available reviewing Arabic contents to construct multi-domain Arabic corpus that is reliable for both tasks of users interest discovery and sentiment analysis. For the automatic generation of this corpus, an open-source Scrapy framework is utilized, which is a framework for constructing custom web crawlers. The constructed dataset covers eight domains. The text reviews are collected form several websites including: BBC Arabic, the topics collected form this site include (Economy, Sports, politics, and Culture). CNN Arabic, the topics collected from this sites include (books, movie, hotels, and restaurant), in addition to other websites such as aljazeera.net and khaleej.com etc. As detailed in Table 3.1, the constructed dataset contains 32,950 instances (reviews) about eight topics, for each topic the reviews are divided into two sentiment classes (positive and negative), the positive class contains 16,425 while the negative class contains 16,525 number of instances.

Data cleaning

This step focuses on removing unwanted contents such as URLs, HTML tags, numbers, Arabize. Also Figure 3.9 presents an example of removing non-letter contents from the sequences. Any non-Arabic words such as glossaries are always written in English. Therefore, in this step, these words are removed as shown in Figure 3.10. Strange words are also considered as non-Arabic contents, those words are removed using predefined list as Figure 3.11 shows sample of Arabize words removing list. Figure 3.12 presents an example of sentence before and after removing the Arabize contents.

Before	فندق رائع فندق جميل يستحق فوق 100% الاستقبال والخدمة والنظافة ممتازه ومستوى الطعام فوق الممتاز والانيمشن amazing
After	فندق رائع فندق جميل يستحق فوق الاستقبال والخدمة والنظافة ممتازه ومستوى الطعام فوق الممتاز والانيمشن amazing

Figure 3.9 Non-letter removing example

Before	فندق رائع فندق جميل يستحق فوق الاستقبال والخدمة والنظافة ممتازه ومستوى الطعام فوق الممتاز والانيمشن amazing
After	فندق رائع فندق جميل يستحق فوق الاستقبال والخدمة والنظافة ممتازه ومستوى الطعام فوق الممتاز والانيمشن

Figure 3.10 Removing non Arabic words

قراند	ستاف	اوتيل
هارتس	لاتيه	سيتى

Figure 3.11 Sample of Arabize removing list

Before	أنظف الفنادق و أفخمها فى مسقط سيتى الواجهات التي تطل على البحر جميله
After	أنظف الفنادق و أفخمها فى مسقط الواجهات التي تطل على البحر جميله

Figure 3.12 Arabize words removing example

Normalization

Normalization is the task of making a text more reliable by changing some characters, eliminating extra white spaces, removing diacritics etc. Generally, this phase consists of two tasks: removing and replacing, as show in figure 3.8, normalization is done as follows: 1. Removing: In this step the punctuation and diacritics are removed, thus making the text processing more easily. It's worth mentioning that the punctuation are used in Arabic to organize sentences to be readable in proper format. Figure 3.13 presents a list of punctuation to be removed, Figure 3.14 shows an example of punctuation removing.

“	&	#	^	(	)
%		=	-	+	?
@	:	;	[	]	,
!	*	\	/	{	}

Figure 3.13 Punctuation list

Before	الفندق نظيف ,الخدمة جميلة! الاخوة في الاستقبال رائعين ,والاهم هو القرب من الحرم المدني..
After	الفندق نظيف الخدمة جميلة الاخوة في الاستقبال رائعين والاهم هو القرب من الحرم المدني

Figure 3.14 Punctuation removing example

Another thing to be removed is the diacritics. Furthermore, the extra white spaces should be removed because it makes the tokenization return an empty word. Therefore all of the white spaces are replaced with single space. Figure 3.15 presents an example of removing diacritics and white spaces.

Before	عَزّف رائع على البيانو لأشهر الأغاني الرومانسية الجَميلة جداً إهداء خاص لَجَميع المحبين...
After	عزف رائع على البيانو لاشهر الأغاني الرومانسية الجميلة جدا إهداء خاص لَجَميع المحبين...

Figure 3.15 Punctuation and extra white spaces removing

Stop words can be categorized in into two classes extend in suffixing or prefixing. All forms of stop-words should be removed from the sentences.

Lemmatization

This task is concerned with obtaining the root of words using morphological analysis. This process involved predefined dictionaries to association every word with its root: thus, for this task we used light10 stemmer because it is faster than lemmatization.

Also, the main aim of this process is to enhance the text classification by removing insignificant information for the text. In our constructed corpus we applied the following processes: removal of numbers, punctuation, and stop words such as أبو أبٌ أجل أخو, in addition to removing any un-Arabic text. Text normalization including Hamza ء and Taa ة to ا and ه آ respectively, and removing the duplicated letters such as "جميييييل جددداا" beautiful to be آ جميل جدا آ. removing of diacritics such as shedah and tanween, and finally removing the duplicated reviews for the corpus. For this task Weka 3.7.10 which is open source software which combines a large set of Machine Learning Algorithms is used for data preparation and preprocessing.

Data annotation

The process of annotation started for earlier steps (collecting), when collecting the data, we have been the initial text label of the sentences have been taken into consideration, we already know the topics and the sentiment orientation of the collected text, then sentences of the same sentiment orientation are grouped together in the same CSV file. This process as accomplished by a team of linguistics experts, the process consists of three primary sub tasks: annotation, review, and approval. A tool is used to achieve this process, this tool allows the annotators to retrieve the sentences in sequentially from the CSV then sentences is read, analyzed and assigned appropriate label according the measuring results. After that, the annotators will double-check if the associated label is matching the overall sentiment of the review based on their judgment after understanding the reviewer point of view. Once the annotators complete the annotation, another member will verify the annotation, and report the final approval. Each sentence is given a score, this score serves as a measure positivity and negativity of the sentence using a set of specific features. With the help of adjectives weighting table the annotator calculates the total number of the adjective presented in the sentence (N), determine the (ni) the number of positive and negative adjectives for each label.

Data division

After preprocessing, the corpus is ready for the experiments, for UID and SA tasks, the entire corpus is divided into two parts, training and testing sets. According to the training set, the proposed Deep InterSent model is trained to produce a classification model. The testing set is then used to assess the classification performance of the proposed model. As there is no optimal ratio for the training data size to testing data size, we divided the constructed corpus into 70% for the training and 30% for the testing for both tasks i.e UID and SA. Table 3.1 presents the labels distribution details of this dataset.

Table 3.1 Arabic dataset for UID and SA tasks

N	Class	N.instances	Positive	Negative
1	Sport	4,092	2,110	1,982
2	Culture	4,000	1,865	2,135
3	Politic	4,274	2,237	2,037
4	Books	4,198	2,000	2,198
5	Economy	4,000	2,000	2,000
6	Restaurant	4,150	2,095	2,055
7	Hotel	4,110	2,155	1,955
8	Movie	4,126	1,963	2,163
Total		32,950	16,425	16,525

Table 3.2 presents sample of review from hotels domain with positive sentiment orientation with its English translation, meanwhile, Table 3.3 shows sample of negative sentiment reviews from the same domain.

Table 3.2 Sample of positive polarity review.

Arabic Comment	فندق رائع يتميز بالهدوء ومستوى النظافه ممتاز تعامل لطيف من العاملين الغرف نضيفه ورائعة كل ماتحب اوتبحث عنه تجده في هذا الفندق..
English Translation	wonderful hotel,quiet and clean nice treat from the staff rooms are clean and wonderful, shortly all you like and looking for you find in this hotel..

Table 3.3 Sample of negative polarity review.

Arabic Comment	اسواء مكان ممكن تزوره اوتقيم فيه غرف سيئة وخدمة متدنية - بصورة عامة مخيب للآمال
English Translation	It's the worst possible place to visit or to stay in, where bad rooms and low service -in general disappointing

3.4.2 English database: Amazon products reviews

Due to the lack available corpora for the purposes of users interests discovery and sentiment analysis we used subsets from amazon products reviews corpus McAuley and Leskovec, 2013, this corpus contains about 34,686 reviews about 24 topics and metadata, it has built in 2013.

For users interests discovery task, we sampled subsets from 10 top-level categories including as books, movies and TV, clothing, etc. For the sentiment analysis, we annotated each review text into positive or negative sentiment labels. Finally, all the reviews are stored together into one text file which contains reviews text, topic labels and the sentiment labels for a given reviews. Table 3.4 presents labels distribution of this dataset which is divided into 70% for training and 30% for testing.

Validation of the Framework'practical relevance by including it in uszzer's interests discovery and sentiment analysis's contexts. This phase aims to evaluate the performance of the proposed framework interests discovery and sentiment analysis's contexts in English and Arabic textual data, also the details to the constructed dataset and the slandered perfor-

Table 3.4 English dataset for UID and SA tasks

N	Class	N.instances	Positive	Negative
1	Books	6,000	3,320	2,680
2	Movies and TV	6,000	3,445	2,555
3	Video Games	6,000	2,858	3,142
4	Sports and Outdoors	6,000	3,269	2.731
5	Digital Music	6,000	3,342	2,658
6	Clothing	6,000	3,452	2,548
7	Electronics	6,000	2,880	3,120
8	Office Products	6,000	3,157	2,843
9	Automotive	6,000	3,211	2,789
10	Health and Personal Care	6,000	3,024	2,976
Total		60,000	31,958	28,042

mance measures are presented. This phase seeks answers the following question

- Question 1 : What are the slandered measures used to evaluate the classification performance of deep learning frameworks?
- Question 2: What are the optimal parameters of the algorithms that lead to achieve the highest performance?
- Question 3: How to validate the contribution of each components in the final predictions?

3.5 Experimental settings and parameters

3.5.1 Model hyperparameters and settings

Empirically, different hyberparameters and settings are examined for each task and also for each language, this subsection presents the basic hyberparameters of the CNN and LSTM which are used on the experiments of Deep InterSent as presented in Table 3.5. Note that the number of features is varies according to the training task.

Table 3.5 Model hyperparameters and settings

Parameter	Value
Filter window size	3, 4, 5
Feature maps	256
Activation functions	ReLu
Padding	Zero
Dropout	0.5
LSTM hidden state dimension	128
Number of epochs	10
Learning rate	0.001

3.5.2 Implementation environment

The proposed system architecture is implemented in Tensorflow which is an open source framework for NLP and machine learning problems developed by Google to meet their needs for systems capable of building and training neural networks. Tensorflow runs in Python under Linux platform. Table 3.6 shows the development environment used for our implementation.

Table 3.6 Implementation environment

Development Platform	Ubuntu 14.04.1
Scripting Language	Python 2.7.12
Software library	TensorFlow 2.0 – CPU

3.6 Experimental Results

Although we used the same deep learning architecture, the experiments of UIC and SA are conducted in separately for each language, where different features are extracted according to the training data, for topic classification, the model is trained to extract and classify the features (nouns) into certain topic label such as sport. For the SA the mode is trained to extract and classify the features (adjectives) into cetin sentiment labels i.e positive/negative which represent the sentimental orientation of the users towards the detected topic.

3.6.1 Evaluation metrics

To assess the classification performance of the proposed Deep system in both UID and SA, we used accuracy, precision, recall and F1-measure (Alexander Clark, Fox, and Lappin, 2013) which are commonly used evaluation metrics to evaluate the performance of deep learning models in text classification (N. et al., 2014). The Precision is used to determine how precise the interest or sentiment categories are classified, its determined using the number of correctly classified (TP) text divided by the total number of predictions as in formula 3.10

$$Precision = \frac{TP}{TP + FP} \tag{3.10}$$

The Recall is the ratio of the correctly classified interest or sentiment categories and the total number of relevant classes of instances as defined in formula 3.11

$$Recall = \frac{TP}{TP + FN} \tag{3.11}$$

F1-score Measure is used evaluates the balance between the precision and the recall as in formula 3.12

$$F1Measure = \frac{2 * Precision + Recall}{Precision + Recall} \tag{3.12}$$

Finally, the classification accuracy is calculated as:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{3.13}$$

Where: TP denotes true positive, FP false positive, TN denotes true negative and FN denotes false negative.

3.6.2 Results of users interests discovery in English database

Initially, Deep InterSent model is trained on topics classification using 10 topic categories as shown in Table 3.4, then the classification performance is tested with the testing set. to assess the classification performance of Deep InterSent we focus on Image result for colors confusion matrix Confusion Matrix as illustrated in Figure 3.16 where each number in the cells represents the percentage of true positive and false positive of each topics category.

For deeper assessment of Deep InterSent model classification performance, precision, recall, F1 score, and accuracy measures are used, Table 3.7 and Figure 3.17 presets the obtained performance of Deep InterSent for each topic individually and the overall model performance.

	Books 1	Movies, TV 2	Digital Music 3	Electronics 4	Clothing 5	Sports and Outdoors 6	Automotive 7	Office Products 8	Health and Personal Care 9	Video Games 10	Accuracy %
Books	82.17	2.83	2.94	0.11	1.39	1.78	1.11	1.33	1.72	0.67	83.19
Movies, TV	2.89	87.22	2.28	1.00	1.00	2.78	1.94	1.72	1.78	2.17	88.22
Digital Music	2.83	1.94	85.17	2.50	0.72	1.17	1.33	1.67	1.28	1.56	86.85
Electronics	1.44	1.61	2.22	88.78	0.83	0.56	2.33	3.61	1.39	2.44	88.93
Clothing	1.83	1.22	1.44	1.39	83.28	2.89	1.22	1.22	1.39	1.33	83.59
Sports and Outdoors	1.39	1.44	0.56	0.33	3.78	85.33	0.78	0.78	2.11	1.22	85.81
Automotive	2.00	1.50	1.00	0.61	2.67	1.22	86.11	2.67	0.94	1.11	86.82
Office Products	2.61	0.50	1.33	1.72	0.56	0.78	2.50	84.67	0.83	1.00	84.94
Health and Personal Care	1.50	0.56	0.89	0.56	4.33	2.39	0.89	1.22	87.50	1.33	87.96
Video Games	1.33	1.17	2.17	1.94	1.44	1.11	1.78	1.11	19.00	87.17	87.84

Figure 3.16 Topics classification Confusion Matrix

Table 3.7 InterSent model classification Performance

Category	Precision	Recall	F1	Accuracy
Books	82.17	84.61	83.37	83.19
Movies, TV	87.22	83.24	85.19	88.22
Digital Music	82.15	85.02	83.57	86.85
Electronics	88.78	84.37	86.52	88.93
Clothing	83.28	85.66	84.45	83.59
Sports and Outdoors	85.33	87.32	86.32	85.81
Automotive	86.11	86.25	86.18	86.82
Office Products	87.14	85.52	86.32	84.94
Health and Personal Care	87.50	86.49	86.99	87.96
Video Games	87.17	86.93	87.05	87.84
Overall	90.68	89.18	89.92	87.41

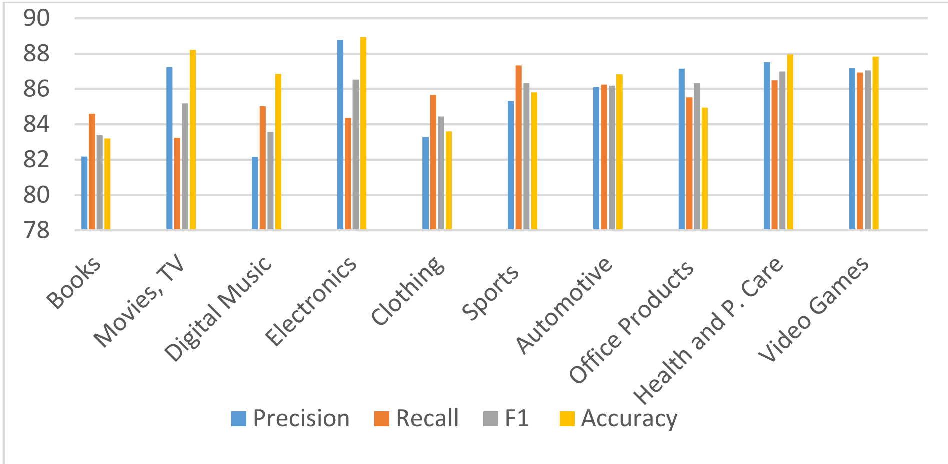


Figure 3.17 Evaluation metrics per topic

The superior classification performance is recorded on "Electronics" topic i.e 88.78%, 84.37% and 88.93% of precision recall and accuracy respectively, whihle the lowest performance is recorded on "Books" topic with 82.17%,84.61, and 83.37% of precision recall and accuracy respectively. the overall performance the model reached 90.68%, 89.18%, and 87.41% of of precision recall and accuracy respectively.

Best performing classifier

The classifier performance is eventually determine the quality of the word embedding and features extraction approaches, therefore, we trained and tested our proposed deep learning architecture with Naive Bayes (NB), and K-Nearest Neighbors (KNN) classifiers on the same takes using same training parameters and dataset splits. Figure 3.18 shows the overall classification performance of Deep InterSent using SVM classifier compared against the NB, and KNN where K=10. SVM achieved the superior performance over NB and KNN with a significant difference in terms of precision, recall and accuracy. KNN is the worst preforming classifier.

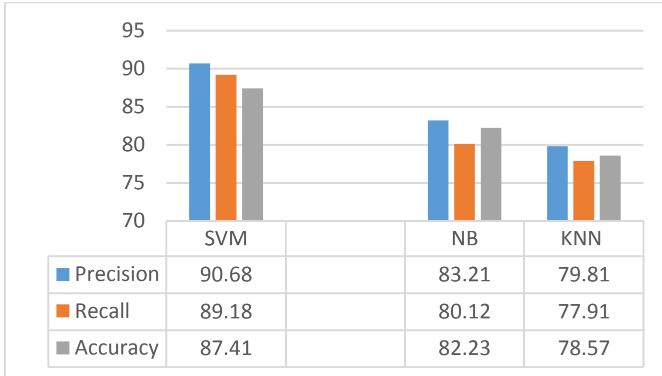


Figure 3.18 Effect of classifiers on the classification performance

Best performing representations model

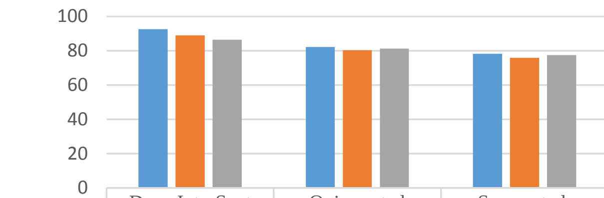
Extensive experiments are conducted to assess the classification performance of Deep InterSent using word2vec model which introduced by Mikolov et al (2013), based on two-layers neural network word2vec performs a distributed representation of the words, it contains 3 million words represented in 300-dimensional vector space. Table 3.8 shows the classification performance of Word2vec-Deep InterSent compared against FastText-Deep InterSent models, according to the results, FastText- Deep Deep InterSent outperformed Word2vec-Deep InterSent with enhancements of +5.77%, +6.57%, and +2.1% of precision, recall and accuracy respectively, therefore FastText model is the best choice for this task, which is consistent with Bojanowski et al., 2017 which stated that FastText model produces high quality words vectors representations using the semantic and syntactical information, also it can generate vectors representation for out-of-vocabulary words.

Table 3.8 Performance with different embedding models

Embedding model	Precision	Recall	Accuracy
FastText-InterSent	**90.68**	**89.18**	**86.41**
Word2vec-InterSent	84.91	82.61	84.31

Performance comparison with related works

The classification performance of InterSent is evaluated against Quispe et al. Quispe, Ocsa, and Coronado, 2018 which used Latent Semantic Indexing, CNN, and Multi-Layer Perceptron, in addition to Senge et al. Senge, Coz, and Hüllermeier, 2019 which used classifier chains approach for multi-label classification. As illustrated in Figure 3.19, InterSent achieved the superior performance.



	Deep InterSent	Quispe et al	Senge et al.
Precision	92.54	82.14	78.31
Recall	88.94	80.25	75.91
Accuracy	86.45	81.15	77.43

Figure 3.19 Comparison with the state of the arts

3.6.3 Results of Sentiment Analysis in English

After performing topics classification as reported in the previous section, in this section we report the classification performance of Deep InterSent in sentiment analyses which is a standard binary classification task to classify the opinions towards the identified topics. We calculated the overall confusion matrix for all positive and negative examples in the dataset as presented in Figure. 3.20 where 90.81% of the total positive sentences are correctly classified into positive class while 9.19% are mis-classified, also 89.58% of the total negative examples are correctly classified into negative class while 10.42% are mis-classified by the Deep InterSent.

For deeper evaluation on SA task, Table 3.9 reports the classification performance in terms of Precision, Recall, F1 score and accuracy measures.

Figure 3.20 Sentiment Analysis confusion matrix

Table 3.9 Sentiment analysis evaluation metrics

Measure	Value
Precision	89.14
Recall	90.35
F1 score	89.19
Accuracy	89.85

Comparison with English SA related works

The classification accuracy of InterSent is evaluated on Sanders and movie reviews datasets, and compared against different baseline studies including: Lu et al Y. Lu et al., 2018 which joint Bidirectional Long Short-Term Memory with Lexicons, Ouyang et al. Ouyang et al., 2015 which used Word2vec over a CNN, Sahu et at. Sahu and Ahuja, 2016 which used structured N-grams with RF classifier, Chen et al. S. Chen, G. Chen, and W. Wang, 2016 which used semantic and syntactic embeddings with extended static and non-static CNN architectures. In addition to Pawar et al. Pawar and R. R. Deshmukh, 2015 which used ML classifiers as well as lexicon based approach, Fouad, et al. Fouad, Gharib, and Mashat, 2018 which used BoW as features extraction, Information Gain (IG) as feature selection technique in addition to SVM, NB and LR classifiers. And Attia et al. Attia et al., 2019 which used CNN architecture. As shown in Table 3.10 Deep InterSent outperformed all of the baseline studies on both datasets.

Table 3.10 Comparison with the existing SA approaches

Dataset	Study	Accuracy(%)
Sanders	(Attia et al., 2019)	78.60
	(Fouad, Gharib, and Mashat, 2018)	90.11
	(Pawar and R. R. Deshmukh, 2015)	88.65
	(Y. Lu et al., 2018)	81.30
	Deep InterSent	**91.78**
Movie reviews	(Sahu and Ahuja, 2016)	88.95
	(Ouyang et al., 2015)	45.40
	(S. Chen, G. Chen, and W. Wang, 2016)	81.30
	Deep InterSent	**90.35**

3.7 Results of Users Interests Discovery in Arabic

Separate experiments are conducted to train and test the proposed model on the Arabic textual data. Similar to what we done for English language, in this subsection we report the classification performance of Deep IntrSent on user interest classification. To achieve our experiments, we trained and tested the model on the dataset in Table 3.1 which contains eight topics categories. 70% of the dataset for training and 30% for testing. Table 3.11 shows in details the corresponding results for the testing set and the model performance in terms of Precision, Recall, F1-measure and accuracy for each topic individually, in addition to the overall mode performance in "overall" column. Also Figure 3.21 visualizes the obtained metrics values for each topic category.

Table 3.11 Deep InterSent evaluation metrics per topic

Category	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)
Culture	80.48	82.22	81.34	80.48
Politic	78.63	81.14	79.86	78.63
Movie	80.52	82.10	81.30	84.75
Books	83.90	80.19	82.00	83.90
Economy	82.48	81.06	81.76	82.48
Restaurant	83.75	82.65	83.20	83.75
Hotel	82.55	83.74	83.14	82.55
Sport	81.13	83.25	82.17	81.13
Overall	82.23	80.20	81.25	82.21

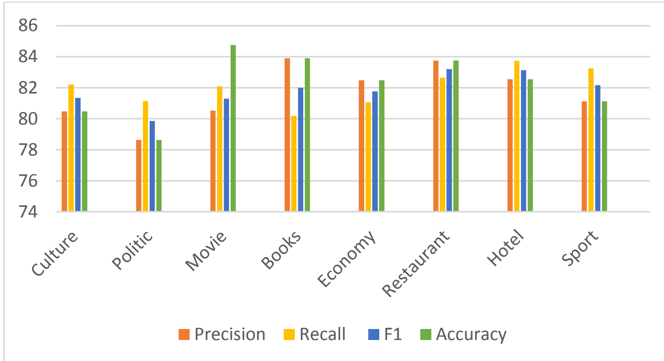


Figure 3.21 Evaluation metrics per category

For this task, the superior classification performance is recorded on "Restaurant" topic i.e 83.75%, 82.65% and 88.93% in terms of precision and recall respectively, meanwhile the highest accuracy percentage is 83.90 % which is recorded on "Books" topic. The lowest performance is recorded on "Politic" topic with 78.63%, 81.14%, and 78.63% of precision, recall and accuracy respectively. The overall performance of the Deep InterSent model on this task is 82.23%, 80.20%, and 82.21% of precision recall and accuracy respectively.

Its observed that the performance on Arabic topics classification is lower than in English due to the smaller size of the dataset in addition to quality of the words vectors which are generated by the embedding model that can effect the features extraction and classification processes, in addition to the labels distribution among topics.

Best performing classifier

For further evaluation we compared the overall classification performance of the proposed Deep InterSent for Arabic text using Naive Bayes (NB) and K-Nearest Neighbors (KNN) classifiers against SVM classifier on the same takes using same training parameters and dataset splits. As shown in Figure 3.22 SVM achieved the highest classification performance with a significant difference in terms of precision, recall and accuracy.

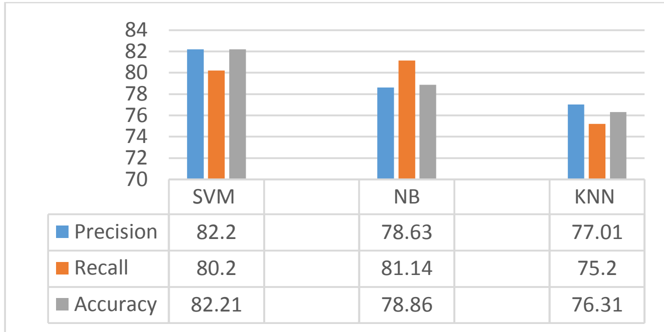


Figure 3.22 Deep InterSent performances with different classifiers

Performance comparison with related works

More experiments are conducted to compare the classification performance of InterSent against Hmeidi et al. Hmeidi, Al-Ayyoub, Mahyoub, et al., 2016 which proposed a lexicon based system for Arabic multi-label text categorization, and Al-anzi et al. Al-Anzi and AbuZeina, 2017 which investigated the performance of cosine similarity measure as well as different ML classifiers including NB, SVM, RF, etc for Arabic text classification. As shown in Figure 3.23, InterSent achieved the best classification performance over the other approaches.

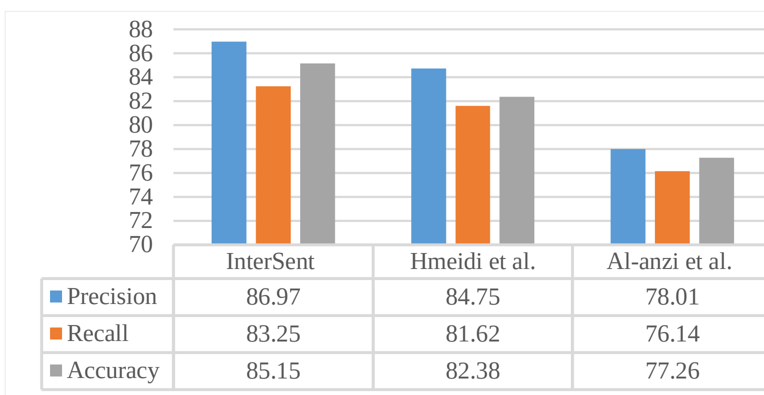


Figure 3.23 Deep InterSent performance against the state of the arts

3.7.1 Results of Sentiment Analysis

The obtained confusion matrix of this experiment is presented in Figure 3.24 where 92.67% of the positive reviews are correctly classified as positive with only 7.33% mis-classified as negative. 84.44% of the negative reviews are correctly classified as negative with only 15.56% mis-classified as positive by Deep InterSent. Table 3.12 presents the classification performance of Deep InterSent on Arabic text which reached 91.02%, 85.12%, 88.58% in terms of precision, Recall, and accuracy.

Figure 3.24 Arabic SA onfusion Matrix

Table 3.12 Overall classification performance in Arabic SA

Measure	Value
Precision	91.02
Recall	85.12
F1 score	87.96
Accuracy	88.58

Best performing embeddings model

The classification performance of Deep InterSent is examined using two other pre-trained word representations model: word2Vec which introduced by Mikolov, Sutskever, et al., 2013 it uses on two-layers Neural Network to perform distributed representation of the words, it contains 3 million words represented in 300-dimensional vector space. AraVec which introduced by Soliman, Eissa, and El-Beltagy, 2017, its distributed word representation model for Arabic language, it provides two architectures: CBOW and Skip-gram with 300 dimension vector space. Table 3.13 and Fig. 3.25 show the obtained classification accuracy of

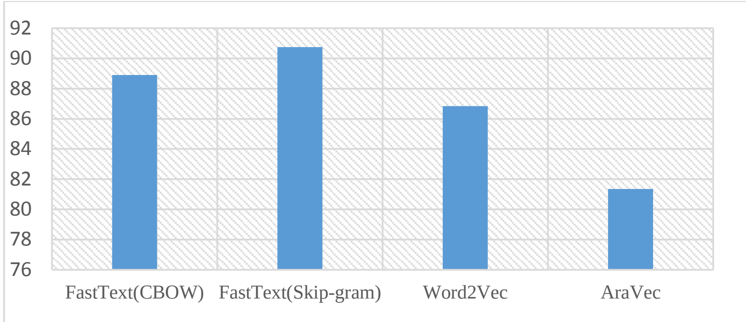


Figure 3.25 performances using different embeddings models

Word2Vec-InterSent and AraVec-InterSent compared against accuracy of FastText-InterSent models, according to these results, FastText (Skip-gram and CBOW) achieved the superior performance with accuracy of 90.75% and 88.90% respectively which is better than Word2Vec and AraVec with +3.3 and +8.8 accuracy rise respectively, at the other hand FastText Skip-gram model achieved the highest classification accuracy. This is consistent with Bojanowski et al., 2017 that FastText skip-gram produces high quality words vectors representations since it incorporates the semantic and syntactical information from the texts with the learned words vectors, also it can generate vectors representation for out-of-vocabulary words.

Table 3.13 Classification accuracy using different embeddings models

Embeddings model	Accuracy (%)
FastText(CBOW)-InterSent	88.90
FastText(Skip-gram)-InterSent	**88.58**
Word2Vec-InterSent	86.84
AraVec-InterSent	81.35

Comparison with Arabic SA related works

To validate the performance of Deep InterSent with the existing approaches in arabic SA we performed different experiments on several datasets: Large Scale Arabic Book Reviews

(LABR) dataset constructed by Aly and Atiya, 2013 it contains 63,000 of book reviews from Goodreads. Arabic Sentiment Tweets Dataset (ASTD) collected by Nabil, Aly, and Atiya, 2015 it contains 10.000 Arabic tweets. Arabic sentiment analysis Twitter dataset collected by N. a. Abdulla et al., 2013 it contains 2000 tweets classified into positive and negative. The performance of Deep InterSent is compared against: Dahou et al., 2016 which used CNN of one convolutional layer Architecture over Word2Vec. Altowayan, 2017 which experimented FastText with SVC and Logistic Regression classifiers on LABR and ASTD datasets respectively. Altowayan and Tao, 2016 which incorporated POS tags and word stemming features with Logistic Regression on both LABR and ASTD datasets. ElSahar and El-Beltagy, 2011 which utilized three feature representation techniques: Delta-TF-IDF, TF-IDF and Count, with Linear SVM for features selection and classification. N. a. Abdulla et al., 2013 which proposed lexicon-based approach, and supervised approach based on SVM classifier. And Nabil, Aly, and Atiya, 2015 which used token counts and the TF-IDF with SVM classifier. We deliberately used the same training/testing sets size of each individual study. Notably, Deep InterSent achieved the highest accuracy over state-of-the-art on LABR, ASTD, and Ar-Twitter datasets with accuracy of 90.20%, 89.72%, and 88.52% respectively, Table 3.14 and Figure 3.26 show the classification accuracy of Deep InterSent in each dataset against the other approaches listing their best classification accuracy.

Table 3.14 Accuracy comparison with the state of the art

Study	LABR	ASTD	Ar-Twitter
(Dahou et al., 2016)	89.60	79.07	85.01
(Altowayan, 2017)	84.97	87.01	-
(Altowayan and Tao, 2016)	78.60	80.21	-
(Nabil, Aly, and Atiya, 2015)	-	69.01	-
(ElSahar and El-Beltagy, 2011)	80.20	-	-
(N. a. Abdulla et al., 2013)	-	-	87.20
Deep InterSent	**90.20**	**89.72**	**88.52**

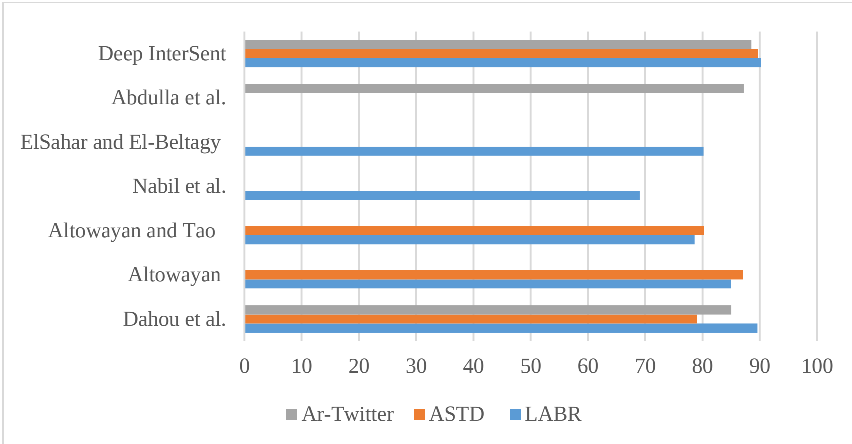


Figure 3.26 Accuracy comparison with the state of the art

It's observed that many factors can affect the performance of DL model on topics classification and sentiment analysis tasks, such as the dataset size, words representation model, the extracted features and the type of the classifier. The experimental results show that using two-layer LSTMs on the top of a CNN architecture can effectively minimize the number of convolutional layers required to capture long-term dependencies. Furthermore, the proposed architecture achieve high performance on users interests discovery and sentiment analysis for both Arabic and English languages, however, the performance on English is superior than on Arabic text, thus, CNN can perform language independent features extraction.

Model/Measure Precision(%) Recall(%) F1-score(%) Accuracy(%)

Chapter Four

Results interpretation

To complete this study properly, it is necessary to analyze and interpreting the obtained results in order to evaluate the hypothesis and the research questions. As already indicated in the preceding chapter. We have carefully reported missed predicted reviews, and we have addressed that the performance of the proposed Deep-Intersent model have been influenced by many issues and factors. The incorrect assigned labels of reviews, the proposed models able to provide the right classes of those reviews, although the actual labels were incorrect. Also, some terms appear only in the training while not existed or existed as synonym in the testing set which can influence the training process. In several datsets, We have noticed that a portion of the missed classified sentences contain both positive and negative sentiments orientation. This mixing can confuse the proposed model in predicting the overall polarity of the sentence. We have also noticed that some terms are used in both the modern standard Arabic to give particular meanings while the same terms are used in Arabic dialects to express different meaning and feelings intensity. In the following section we present the conclusive answers to the research questions:

What are the limitations of the existing Deep Learning algorithms? How to overcome these limitation?

Text mining and NLP approaches are commonly used to infer user interests, they depend on using keywords (explicit indicators) each assigned weights, however keywords vagueness and terms ambiguity (one word expresses several emotions) can reduce the predictions accuracy. Besides, semantic relationships of the words are not considered. On the other side, some existing studies such as Seo and B. T. Zhang, 2000 focused on tracing user's browsing behaviors (implicit indicators) (e.g duration) and user's browsed contents(e.g type of web pages viewed) to detect the users interests and orientations. However, only the previous users' interests can be obtained, predictions are not actual and not precise because users may view web pages but he is not interested in the contents. Meanwhile, these methods confined the users interest to the viewed web pages. Unsupervised approaches are commonly used in SA. However, keywords vagueness and ambiguity can decrease the accuracy of predictions. These approaches cannot consider the semantic relationships between words in the sentences. For Arabic sentiment analysis, unsupervised approaches cannot be effective due to the numerous words from several dialects to be included in the lexicons. Also, it is observed that using only CNN or using only LSTM is inadequate to achieve the desired results on Arabic sentiment analysis (Q. Huang et al., 2017), this is because CNN fails to maintain long-term dependencies, and LSTM is weak to capture local features.

How to best perform topics classification and sentiment analysis for individual user in social networks based on textual data?

As presented in section three, deep learning architecture have proved to be effective in topics classification and sentiment analysis. Experiments have proved that DL architecture combining CNNs and RNNs stacked over semantic layer that combines FastText model to obtain words vectors representation of the input sentences and WordNet knowledge bases can outperforms all of the other existing approaches. It's observed that many factors can affect the performance of DL model on topics classification and sentiment analysis tasks, such as the dataset size, words representation model, the extracted features and the type of the classifier. The experimental results show that using two-layer LSTMs on the top of a CNN architecture can effectively minimize the number of convolutional layers required to capture long-term dependencies. Furthermore, the proposed architecture achieve high performance on users interests discovery and sentiment analysis for both Arabic and English languages, however, the performance on English is superior than on Arabic text, thus, CNN can perform language independent features extraction.

How to improve the state of the art Deep learning to perform text classification semantically?

Experiments have proved that capturing long short terms dependencies and contextual information using LSTM can improve the semantic understanding of textual data, Figure 3.22, Table 3.10, Figure 3.25 proved this conclusion that CNN stacked over LSTM layer can significantly enhance the terms dependency processing thus, improving the semantic understanding of the model, furthermore, LSTM perform better with the same architecture due to the ability of capturing the contextual information, standard LSTM that has recently become an effective algorithm for sequences modeling because it can capture past contextual features using many hidden layers.

Which is the best performing words representation model for Arabic text contents?

according our verification in Table 3.18 and Table 3.13 we have proved that FastText skip-gram produces high quality words vectors representations since it incorporates the semantic and syntactical information from the texts with the learned words vectors, also it can generate vectors representation for out-of-vocabulary words

FastText and Word2Vec attained relatively comparable performances. FastText has the advantage of representing each word by breaking it into its character n-grams and the sum of its n-grams vectors gives the final word vector. That quality makes FastText representations a perfect fit for morphologically rich languages such as Arabic. Whereas Word2Vec represent each word by a distinct vector. As a result, the FastText model with SG architecture was proved to attain the best results on several datasets, FastText (Skip-gram and CBOW) achieved the superior performance with accuracy of 90.75% and 88.90% respectively which is better than Word2Vec and AraVec with +3.3 and +8.8 % accuracy rise respectively, at the other hand FastText Skipgram model achieved the highest classification accuracy. This is consistent with Bojanowski et al., 2017 that FastText skip-gram produces high quality words vectors representations since it incorporates the semantic and syntactical information from the texts with the learned words vectors, also it can generate vectors representation for out-of-vocabulary words.

Research hypotheses

Hypothesis 1: Integrating different Deep learning algorithms can provide better classification performance in a complex morphological language such as Arabic.

This hypothesis have proved to be correct, every algorithms have been selected carefully in order to perform specific function meanwhile supporting in overcoming the limitation of the other algorithms this have been proved, in section 3.10.1 The experimental results show that using two-layer LSTMs on the top of a CNN architecture can effectively minimize the number of convolutional layers required to capture long-term dependencies. Furthermore, the proposed architecture achieve high performance on users interests discovery and sentiment analysis for both Arabic and English languages, however, the performance on English is superior than on Arabic text, thus, CNN can perform language independent features extraction. Also, to increase the coverage and the initialization of the embedding space, we proposed to incorporate WordNet and ConceptNet just before FastText to perform semantic information injection. If FastText fail to generate the corresponding word vector for particular word, then WordNet and ConceptNet can identify a similar word (synonyms) for that word as shown in Algorithm.1. This integration have improved the performance of Arabic text representation. Traditional convolutional neural network fails to capture and maintain the terms dependency in long sequences of the Arabic text, therefore integrating LSTM provides effective mechanism to maintain terms dependency, so the extracted features are fed into two layers LSTM this integration have improved the classification performance of up to + 13.14% over the only CNN model reported in Dahou et al. (2016) and up to + 14.59% and +27.16% accuracy improvements over the CNN architecture with one stacked LSTM layer reported in (Abdulaziz M. Alayba et al., 2018) and (Heikal, Torki, and El-Makky, 2018) respectively also, significant influence of integrating the BiLSTM and the attention mechanism on the Arabic features extraction and processing which attained +4.82% and +2.63% of accuracy improvement over Dahou et al. (2016) and Abdulla et al. (2013) because attention mechanism allows to extract only the important feature from the features sequence generated from the BiLSTM. SVM achieved the superior performance over NB and KNN with a significant difference in terms of precision, recall and accuracy because it performs optimally in cases where there is a distinct margin of separation among classes. It is more robust i.e. due to optimal margin gap between separating hyper planes, it could do predictions better with test data. It is computationally more efficient. This is because of using Kernel trick in dual problem unlike the other classifiers even Softmax function

Hypothesis 2 :Using WordNet lexical database enhances the quality of the representation, thus improving the overall performance.

This hypothesis have proved to be correct, as shown in in Algorithm.1. incorporating WordNet with the embedding layer can effectively enhances the quality of the representation. Indeed, WordNet standardization allows interoperability between WordNets and facilitates interchange with other standardized resources. In this research, to increase the coverage and the initialization of the embedding space, we proposed to incorporate WordNet and ConceptNet just before FastText to perform semantic information injection. If FastText fail to generate the corresponding word vector for particular word, then WordNet and ConceptNet can identify a similar word (synonyms) for that word.

Hypothesis 3: Stacking LSTM over CNN provide more substantial contextual information extraction.

This hypothesis have proved to be correct, In this architecture, one convolutional layer has been used to capture n-gram features from the input sentence. A stacked Bi-directional LSTM has been integrated to further extract contextual information form the feature sequences obtained by the convolutional layer. However, standard LSTM fails to utilize the semantic dependency from the future sequence (words after the current word) when predicting the semantic meaning of the input text sequence. As presented in section 3.9, 4.5.1 LSTM have provide substantial improvement in the classification performance of the proposed DL architectures 89.83% of accuracy with +1.73% improvement over the CNN with one LSTM layer reported in (Abdulaziz M. Alayba et al., 2018), in addition to +4.82% and +2.63% of accuracy improvement over Dahou et al. (2016) and Abdulla et al. (2013) respectively. On Main-AHS dataset, Standard LSTM is a type of neural network that processes sequences in temporal order, however it fails to utilize the semantic dependency from the future sequence. LSTM based on Nadam and L2 models have improved in terms of convergence speed and accuracy. Because the LSTM based on L2 and Nadam has a stronger ability to prevent over-fitting and improve the model's generalization and convergence speed. We have proved that the integration of CNN, stacked BiLSTM and attention units improves the semantic understanding of Arabic expressions.

Chapter Five

Conclusion and Further Work

5.1 Conclusion

Recently, social media have witnessed exponential growth in user-generated content which contains enormously valuable information for different applications. Text classification is concerned with analyzing social data to identify the inclinations of the public audience. It is challenging to perform sentiment analysis regardless of deep considerations of semantic and syntactic rules, in addition to terms dependencies of the input sentence. Therefore, this study has proposed a novel deep learning system for user interest discovery and sentiment analysis. The proposed system has skillfully joint a CNN architecture stacked over two-layers LSTMs for features extraction and contextual information extraction respectively.

In this system, the quality of words representation has been improved using FastText and WordNet lexical database at the input layer. The proposed system has demonstrated a remarkable performance in predicting the topics of interest and the sentiment orientation over the predicated topics for the individual user based on textual contents of English and Arabic languages. The obtained results have shown the significant performance of the proposed system. Also, extensive experiments have been conducted to validate the performance of this system using Word2Vc model, in addition to KNN and NB classifiers. Results have proved that FastText model and SVM classifier are more relevant alternatives to learn semantic and syntactic information for both English and Arabic languages, also these experiments have demonstrated that CNN can handle multi-language features.

Different experiments have been conducted to validate the contribution of each component on the classification performance of this system. Experimental results have shown the remarkable performance of the proposed system in Arabic text sentiment analysis. Also, results have shown that LSTM and CNN have great impact on improving the classification performance. This system has achieved superior performance and has outperformed different existing approaches on the same task.

5.2 Future Work

These results give further encouragement for future research directions, it is worth investigating: The application of deep learning architectures including CNN and RNN in users modeling and recommendations systems. The application of the attention-based model on user interest categorization based on multi-sources data such as (images and the associated text). It is worth incorporating the proposed systems with social networks APIs, particularly Tweeter API which provide rich semantics text contents. Expanding the coverage and diversity of the constructed dataset to contain Arabic dialects, particularly Sudanese dialects.

REFERENCES

Abbas, M., K. Smaili, and D. Berkani (2011). "Evaluation of topic identification methods on arabic corpora". In: *Journal of Digital Information Management* 9.5, pp. 185–192. ISSN: 09727272.

Abbas, Mourad and Daoud Berkani (2006). "A Topic Identification Task for Modern Standard Arabic". In: 2006, pp. 1092–1096.

Abbas, Mourad, Kamel Smaïli, Daoud Berkani, Mourad Abbas, Kamel Smaïli, Daoud Berkani Tr-classifier, Mourad Abbas, Kamel Smaili, and Daoud Berkani (2017). "TR-Classifier and kNN Evaluation for Topic Identification tasks To cite this version : TR-Classifier and kNN Evaluation for Topic Identification tasks". In:

Abbasi, Ahmed, Hsinchun Chen, and Arab Salem (2008). "Sentiment analysis in multiple languages". In: *ACM Transactions on Information Systems* 26.3, pp. 1–34. ISSN: 10468188. DOI: 10.1145/1361684.1361685. URL: http://portal.acm.org/citation.cfm?doid=1361684.1361685.

Abd-Elhamid, Laila, Doaa Elzanfaly, and Ahmed Sharaf Eldin (2016). "Feature-based sentiment analysis in online Arabic reviews". In: *Proceedings of 2016 11th International Conference on Computer Engineering and Systems, ICCES 2016*, pp. 260–265. DOI: 10.1109/ICCES.2016.7822011.

Abdulla, Nawaf a, Nizar a Ahmed, Mohammed a Shehab, and Mahmoud Al-ayyoub (2013). "Arabic Sentiment Analysis". In: *Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT13)* 6.12, pp. 1–6. ISSN: 16113349. DOI: 10.1109/AEECT.2013.6716448.

Abdulla, Nawaf A., Mahmoud Al Ayyoub, and Mohammed Naji Al Kabi (2014). "An extended analytical study of Arabic sentiments". In: *International Journal of Big Data Intelligence* 1.1/2, p. 103. ISSN: 2053-1389. DOI: 10.1504/ijbdi.2014.063845.

Aggarwal, Charu C., Xiangnan Kong, Quanquan Gu, Jiawei Han, and Philip S. Yu (2014). "Active learning: A survey". In: *Data Classification: Algorithms and Applications*, pp. 571–605. DOI: 10.1201/b17320.

Aggarwal, Charu C. and Cheng Xiang Zhai (2013). "Mining text data". In: *Mining Text Data* 9781461432, pp. 1–522. DOI: 10.1007/978-1-4614-3223-4.

Ain, Qurat Tul, Mubashir Ali, Amna Riaz, Amna Noureen, Muhammad Kamran, Babar Hayat, and A Rehman (2017). “Sentiment Analysis Using Deep Learning Techniques: A Review”. In: 8.6.

Alayba, Abdulaziz M., Vasile Palade, Matthew England, and Rahat Iqbal (2017). “Arabic language sentiment analysis on health services”. In: February, pp. 114–118. DOI: 10.1109/asar.2017.8067771.

— (2018). “A combined CNN and LSTM model for Arabic sentiment analysis”. In: *arXiv:1807.02911v3* 11015 LNCS, pp. 179–191. ISSN: 16113349. DOI: 10.1007/978-3-319-99740-7{_}12.

Alayba, Abdulaziz M, Vasile Palade, Matthew England, and Rahat Iqbal (2018). “Improving Sentiment Analysis in Arabic Using Word Representation”. In: *2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)*, pp. 13–18.

Alexander Clark, Chris Fox, and Shalom Lappin (2013). *The Handbook of Computational Linguistics and Natural Language Processing*. WILEY-BLACKWELL, ISBN: 9781405155816. URL: http://www.ghbook.ir/index.php?name=%D9%81%D8%B1%D9%87%D9%86%DA%AF%20%D9%88%20%D8%B1%D8%B3%D8%A7%D9%86%D9%87%20%D9%87%D8%A7%DB%8C%20%D9%86%D9%88%DB%8C%D9%86&option=com_dbook&task=readonline&book_id=13650&page=73&chkhashk=ED9C9491B4&Itemid=218&lang=fa&tmpl=component.

Alowaidi, Sana, Mustafa Saleh, and Osama Abulnaja (2017). “Semantic Sentiment Analysis of Arabic Texts”. In: *International Journal of Advanced Computer Science and Applications* 8.2, pp. 256–262. ISSN: 2158-107X.

Alsmearat, Kholoud, Mohammed Shehab, Mahmoud Al-ayyoub Riyad, and Ghassan Kanaan (2015). “Emotion Analysis of Arabic Articles and Its Impact on Identifying the Author ’ s Gender”. In:

Altowayan, A Aziz (2017). “Improving Arabic Sentiment Analysis with Sentiment-Specific Embeddings”. In: *IEEE International Conference on Big Data (BIGDATA) Improving*, pp. 4314–4320.

Altowayan, A Aziz and Lixin Tao (2016). “Word Embeddings for Arabic Sentiment Analysis”. In: *IEEE International Conference on Big Data (Big Data) Word*, pp. 3820–3825. URL: http://tanzil.net.

Altrabsheh, Mazen El-masri Nabeela (2017). “Successes and challenges of Arabic sentiment analysis research : a literature review”. In: *Social Network Analysis and Mining*. ISSN: 1869-5469. DOI: 10.1007/s13278-017-0474-x.

Aly, Mohamed and Amir Atiya (2013). “LABR: A Large Scale Arabic Book Reviews Dataset”. In: *Proceedings ofthe 51st Annual Meeting ofthe Association for Computational Linguistics*, pp. 494–498. DOI: 10.13140/2.1.3960.5761. URL: https://www.aclweb.org/anthology-new/P/P13/P13-2088.pdf.

Al-Anzi, Fawaz S. and Dia AbuZeina (2017). “Toward an enhanced Arabic text classification using cosine similarity and Latent Semantic Indexing”. In: *Journal of King Saud University - Computer and Information Sciences* 29.2, pp. 189–195. ISSN: 22131248. DOI: 10.1016/j.jksuci.2016.04.001. URL: http://dx.doi.org/10.1016/j.jksuci.2016.04.001.

Al-anzi, Fawaz S and Dia Abuzeina (2017). “Toward an enhanced Arabic text classification using cosine similarity and Latent Semantic Indexing”. In: *Journal of King Saud University - Computer and Information Sciences* 29.2, pp. 189–195. ISSN: 1319-1578. DOI: 10.1016/j.jksuci.2016.04.001. URL: http://dx.doi.org/10.1016/j.jksuci.2016.04.001.

Appear, To, American Society, and Information Science (2000). “To Appear in the Journal of the American Society for Information Science (JASIS) 1”. In: *Journal of the American Society for Information Science*, pp. 1–52.

Assiri, Adel, Ahmed Emam, and Hmood Aldossari (2015). “Arabic Sentiment Analysis: A Survey”. In: *(IJACSA) International Journal of Advanced Computer Science and Applications* 6.12, pp. 75–85. ISSN: 16113349. DOI: 10.1007/978-3-319-20367-6.

Attia, Mohammed, Younes Samih, Ali Elkahky, and Laura Kallmeyer (2019). “Multilingual multi-class sentiment classification using convolutional neural networks”. In: *Proceedinds: 11th International Conference on Language Resources and Evaluation*, pp. 635–640.

Al-Ayyoub, Mahmoud, Abed Allah Khamaiseh, Yaser Jararweh, and Mohammed N. Al-Kabi (2019). “A comprehensive survey of arabic sentiment analysis”. In: *Information Processing and Management* 56.2, pp. 320–342. ISSN: 03064573. DOI: 10.1016/j.ipm.2018.07.006. URL: https://doi.org/10.1016/j.ipm.2018.07.006.

Al-ayyoub, Mahmoud and Aya Nuseir (2016). “Hierarchical Classifiers for Multi-Way Sentiment Analysis of Arabic Reviews”. In: 7.2, pp. 531–539.

El-Beltagy, Samhaa R. and Ahmed Ali (2013). “Open issues in the sentiment analysis of Arabic social media: A case study”. In: *2013 9th International Conference on Innovations in Information Technology, IIT 2013*, pp. 215–220. DOI: 10.1109/Innovations.2013.6544421.

Bhargava, Preeti, Oliver Brdiczka, and Michael Roberts (2015). “Unsupervised Modeling of Users ’ Interests from their Facebook Profiles and Activities”. In: pp. 191–201.

Birjali, Marouane, Abderrahim Beni-hssane, and Mohammed Erritali (2017). “ScienceDirect ScienceDirect Machine Learning and Semantic Sentiment Analysis based Algorithms for Suicide Sentiment Prediction in Social Networks”. In: *Procedia Computer Science* 113, pp. 65–72. ISSN: 1877-0509. DOI: 10.1016/j.procs.2017.08.290. URL: http://dx.doi.org/10.1016/j.procs.2017.08.290.

Black, William, Sabri Elkateb, Adam Pease, Horacio Rodriguez, and Musa Alkhalifa (2006). “Introducing the Arabic WordNet Project”. In: *Word Journal Of The International Lin-*

guistic Association 22, pp. 295–299. URL: http://vossen.info/docs/2006/arabic.pdf%5Cnhttp://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.128.1517.

Blair, Stuart J, Yaxin Bi, and Maurice D Mulvenna (2017). "Unsupervised Sentiment Classification : A Hybrid Sentiment-Topic Model Approach". In: DOI: 10.1109/ICTAI.2017.00076.

Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomas Mikolov (2017). "Enriching Word Vectors with Subword Information". In: *arXiv:1607.04606* 5.3, pp. 729–734. ISSN: 17599679. DOI: 10.1039/c2ay25919b. URL: http://arxiv.org/abs/1607.04606.

Boudad, Naaima, Rdouan Faizi, Rachid Oulad Haj Thami, and Raddouane Chiheb (2018). "Sentiment analysis in Arabic: A review of the literature". In: *Ain Shams Engineering Journal* 9.4, pp. 2479–2490. ISSN: 20904479. DOI: 10.1016/j.asej.2017.04.007. URL: http://dx.doi.org/10.1016/j.asej.2017.04.007.

Boudad, Naaima, Rdouan Faizi, Rachid Oulad, Haj Thami, and Raddouane Chiheb (2017). "Sentiment analysis in Arabic : A review of the literature". In: *Ain Shams Engineering Journal*. ISSN: 2090-4479. DOI: 10.1016/j.asej.2017.04.007. URL: http://dx.doi.org/10.1016/j.asej.2017.04.007.

Cabreira, Ariel G., Martín Tripode, and Adriáln Madirolas (2009). "Artificial neural networks for fish-species identification". In: *ICES Journal of Marine Science* 66.6, pp. 1119–1129. ISSN: 10543139. DOI: 10.1093/icesjms/fsp009.

Chang, Victor (2018). "A proposed social network analysis platform for big data analytics". In: *Technological Forecasting and Social Change* 130.October, pp. 57–68. ISSN: 00401625. DOI: 10.1016/j.techfore.2017.11.002. URL: http://dx.doi.org/10.1016/j.techfore.2017.11.002.

Chen, Shu, Guang Chen, and Wei Wang (2016). "The joint effect of semantic and syntactic word embeddings on sentiment Analysis." In: *Proceedings of NIDC2016* 17, pp. 5–9.

Chen, Tao, Ruifeng Xu, Yulan He, and Xuan Wang (2017). "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN". In: *Expert Systems with Applications* 72, pp. 221–230. ISSN: 09574174. DOI: 10.1016/j.eswa.2016.10.065. URL: http://dx.doi.org/10.1016/j.eswa.2016.10.065.

Chen, Yuling and Zhi Zhang (2018). "Research on text sentiment analysis based on CNNs and SVM". In: *Proceedings of the 13th IEEE Conference on Industrial Electronics and Applications, ICIEA 2018*, pp. 2731–2734. DOI: 10.1109/ICIEA.2018.8398173.

Cheng, Hua Li and Cheol Park Soon (2006). "High performance text categorization system based on a novel neural network algorithm". In: *Proceedings - Sixth IEEE International Conference on Computer and Information Technology, CIT 2006*. DOI: 10.1109/CIT.2006.98.

Chiu, Jason P. C. and Eric Nichols (2015). "Named Entity Recognition with Bidirectional LSTM-CNNs". In: 2003. ISSN: 2307-387X. DOI: 10.3115/1119176.1119204. URL: http://arxiv.org/abs/1511.08308.

Choi, Eunsol, Daniel Hewlett, Jakob Uszkoreit, Illia Polosukhin, Alexandre Lacoste, and Jonathan Berant (2017). "Coarse-to-fine question answering for long documents". In: *ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)* 1.January 2018, pp. 209–220. DOI: 10.18653/v1/P17-1020.

Cinar, Yagmur Gizem, Susana Zoghbi, and Marie Francine Moens (2015). "Inferring User Interests on Social Media from Text and Images". In: *Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW 2015*, pp. 1342–1347. DOI: 10.1109/ICDMW.2015.208.

Claypo, Niphat and Saichon Jaiyen (2014). "Opinion mining for Thai restaurant reviews using neural networks and mRMR feature selection". In: *2014 International Computer Science and Engineering Conference, ICSEC 2014*, pp. 394–397. DOI: 10.1109/ICSEC.2014.6978229.

Dahou, Abdelghani, Shengwu Xiong, Junwei Zhou, Mohamed Houcine Haddoud, and Pengfei Duan (2016). "Word embeddings and convolutional neural network for Arabic sentiment classification". In: *Proceedings of the COLING 2016, 26th International Conference on Computational Linguistics :Technical Papers*, pp. 2418–2427. URL: https://www.aclweb.org/anthology/C/C16/C16-1228.pdf.

Darabi, Majid and Nasseh Tabrizi (2017). "An ontology-based framework to model user interests". In: *Proceedings - 2016 International Conference on Computational Science and Computational Intelligence, CSCI 2016*, pp. 398–403. DOI: 10.1109/CSCI.2016.0082.

Deerwester, Scott, George W Furnas, Thomas K Landauer, and Richard Harshman (2015). "Cara Hidup Susah.pdf". In: *Kehidupan* 3.12, p. 34. DOI: 10.1017/CBO9781107415324.004.

Du, Changshun and Lei Huang (2018). "Text classification research with attention-based recurrent neural networks". In: *International Journal of Computers, Communications and Control* 13.1, pp. 50–61. ISSN: 18419844. DOI: 10.15837/ijccc.2018.1.3142.

Du, Jiachen, Lin Gui, Ruifeng Xu, and Yulan He (2018). "A convolutional attention model for text classification". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10619 LNAI, pp. 183–195. ISSN: 16113349. DOI: 10.1007/978-3-319-73618-1{_}16.

Duwairi, Rehab and Mahmoud El-Orfali (2014). "A study of the effects of preprocessing strategies on sentiment analysis for Arabic text". In: *Journal of Information Science* 40.4, pp. 501–513. ISSN: 17416485. DOI: 10.1177/0165551514534143.

Eirinaki, Magdalini, Shamita Pisal, and Japinder Singh (2012). "Feature-based opinion mining and ranking". In: *Journal of Computer and System Sciences* 78.4, pp. 1175–1184. ISSN: 00220000. DOI: 10.1016/j.jcss.2011.10.007.

Elawady, R. M., S. Barakat, and N. M. Elrashidy (2014). "Different Feature Selection for Sentiment Classification". In: *International Journal of Information Science and Intelligent System* 3.1, pp. 137–150.

Elhawary, Mohamed and Mohamed Elfeky (2010). "Mining Arabic business reviews". In: *Proceedings - IEEE International Conference on Data Mining, ICDM*, pp. 1108–1113. ISSN: 15504786. DOI: 10.1109/ICDMW.2010.24.

Elouardighi, Abdeljalil, Mohcine Maghfour, Hafdalla Hammia, and Fatima-zahra Aazi (2017). "Analysis in the Standard or Dialectal Arabic". In: *2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech)*.

ElSahar, Hady and Samhaa R. El-Beltagy (2011). "Building large arabic multi-domain resources for sentiment analysis". In: *Springer-Verlag Berlin Heidelberg* 9042, pp. 23–34. ISSN: 16113349. DOI: 10.1007/978-3-319-18117-2{_}2.

Elsahar, Hady and Samhaa R El-beltagy (2015). "Building Large Arabic Multi-domain Resources". In: pp. 23–34. DOI: 10.1007/978-3-319-18117-2.

Fernández-Gavilanes, Milagros, Tamara Álvarez-López, Jonathan Juncal-Martínez, Enrique Costa-Montenegro, and Francisco Javier González-Castaño (2016). "Unsupervised method for sentiment analysis in online texts". In: *Expert Systems with Applications* 58, pp. 57–75. ISSN: 09574174. DOI: 10.1016/j.eswa.2016.03.031.

Fornaia, Andrea, Christian Napoli, Giuseppe Pappalardo, and Emiliano Tramontana (2015). "Using AOP neural networks to infer user behaviours and interests". In: *CEUR Workshop Proceedings* 1382, pp. 46–52. ISSN: 16130073.

Fouad, Mohammed M., Tarek F. Gharib, and Abdulfattah S. Mashat (2018). "Efficient Twitter Sentiment Analysis System with Feature Selection and Classifier Ensemble". In: *Springer International Publishing AG* 4.April, pp. 516–527. DOI: 10.1007/978-3-319-74690-6.

G.A., Miller, R.Beckwith, C.Fellbaum, D.Gross, and K.Miller. (1993). "Introduction to WordNet: An On-line Lexical Database". In: August.

Ghosh, Rahul, Kumar Ravi, and Vadlamani Ravi (2016). "A novel deep learning architecture for sentiment classification". In: *3rd Int'l Conf. on Recent Advances in Information Technology | RAIT-2016 | A* 27, pp. 1102–1111. ISSN: 22129413. DOI: 10.1007/978-3-319-68195-5{_}122. URL: http://arxiv.org/abs/1707.05809.

Glorot, Xavier, Antoine Bordes, and Yoshua Bengio (2011). "Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach". In: *Proceedings of the 28th*

International Conference on Machine Learning 1, pp. 513–520. URL: http://www.icml-2011.org/papers/342_icmlpaper.pdf.

Graves, Alex, A r. Mohamed, and Geoffrey E. Hinton (2013). "Speech recognition with deep recurrent neural networks". In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* 6, pp. 6645–6649. ISSN: 1520-6149. DOI: 10.1109/ICASSP.2013.6638947.

Graves, Alex, Abdel-rahman Mohamed, and Geoffrey Hinton (2013). "Speech recognition with deep recurrent neural networks". In: *arXiv:1303.5778v1 [cs.NE]* 3.

Graves, Alex and Jürgen Schmidhuber (2005). "Framewise phoneme classification with bidirectional LSTM and other neural network architectures". In: *Neural Networks* 18.5-6, pp. 602–610. ISSN: 08936080. DOI: 10.1016/j.neunet.2005.06.042.

Gridach, Mourad (2016). "Character-Aware Neural Networks for Arabic Named Entity Recognition for Social Media". In: *6th Workshop on South and Southeast Asian Natural Language Processing*, pp. 23–32. URL: http://www.aclweb.org/old_anthology/W/W16/W16-37.pdf#page=35.

Gu, Chengwei, Ming Wu, and Chuang Zhang (2017). "Chinese Sentence Classification Based on Convolutional Neural Network". In: *IOP Conference Series: Materials Science and Engineering* 261.1. ISSN: 1757899X. DOI: 10.1088/1757-899X/261/1/012008.

Guan, Hu, Jingyu Zhou, and Minyi Guo (2009). "A Class-Feature-Centroid classifier for text categorization". In: *WWW'09 - Proceedings of the 18th International World Wide Web Conference* May, pp. 201–210. DOI: 10.1145/1526709.1526737.

Guellil, Imane, Faical Azouaou, and Marcelo Mendoza (2019). "Arabic sentiment analysis: studies, resources, and tools". In: *Social Network Analysis and Mining* 9.1, pp. 1–17. ISSN: 18695469. DOI: 10.1007/s13278-019-0602-x. URL: https://doi.org/10.1007/s13278-019-0602-x.

Gupta, Pankaj, Ritu Tiwari, and Nirmal Robert (2016). "Sentiment analysis and text summarization of online reviews: A survey". In: *International Conference on Communication and Signal Processing, ICCSP 2016*, pp. 241–245. DOI: 10.1109/ICCSP.2016.7754131.

Hai, Zhen, Gao Cong, Kuiyu Chang, Peng Cheng, and Chunyan Miao (2017). "Analyzing Sentiments in One Go : A Supervised Joint Topic Modeling Approach". In: *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING,* DOI: 10.1109/TKDE.2017.2669027.

El-halees, Alaa M (2014). "Arabic Text Classification Using Maximum Entropy Alaa". In: *The Islamic University Journal* January 2007.

Harrag, Fouzi and Eyas El-Qawasmah (2009). “Neural network for Arabic text classification”. In: *2nd International Conference on the Applications of Digital Information and Web Technologies, ICADIWT 2009*, pp. 778–783. DOI: 10.1109/ICADIWT.2009.5273841.

Harris, Zellig S. (1954). “Distributional Structure”. In: *<i>WORD</i>* 10.2-3, pp. 146–162. ISSN: 0043-7956. DOI: 10.1080/00437956.1954.11659520.

Hassan, Abdalraouf and Ausif Mahmood (2017). “Deep Learning approach for sentiment analysis of short texts”. In: *2017 3rd International Conference on Control, Automation and Robotics (ICCAR)*, pp. 705–710. DOI: 10.1109/ICCAR.2017.7942788. URL: http://ieeexplore.ieee.org/document/7942788/.

— (2018). “Convolutional Recurrent Deep Learning Model for Sentence Classification”. In: *IEEE Access* 6, pp. 13949–13957. ISSN: 21693536. DOI: 10.1109/ACCESS.2018.2814818.

Haydar, Mohammad Salman, Mustakim Al Helal, and Syed Akhter Hossain (2018). “Sentiment Extraction From Bangla Text : A Character Level Supervised Recurrent Neural Network Approach”. In: *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, pp. 1–4. DOI: 10.1109/IC4ME2.2018.8465606.

He, Ruining and Julian McAuley (2016). “Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering”. In: *25th International World Wide Web Conference, WWW 2016*, pp. 507–517. DOI: 10.1145/2872427.2883037.

Heikal, Maha, Marwan Torki, and Nagwa El-Makky (2018). “Sentiment Analysis of Arabic Tweets using Deep Learning”. In: *Procedia Computer Science* 142, pp. 114–122. ISSN: 18770509. DOI: 10.1016/j.procs.2018.10.466. URL: https://doi.org/10.1016/j.procs.2018.10.466.

Hemmatian, Fatemeh and Mohammad Karim Sohrabi (2017). “A survey on classification techniques for opinion mining and sentiment analysis”. In: *Artificial Intelligence Review*, pp. 1–51. ISSN: 15737462. DOI: 10.1007/s10462-017-9599-6. URL: https://doi.org/10.1007/s10462-017-9599-6.

Hinton, Geoffrey E., Simon Osindero, and Yee Whye Teh (2006). “A fast learning algorithm for deep belief nets”. In: *Neural Computation* 18.7, pp. 1527–1554. ISSN: 08997667. DOI: 10.1162/neco.2006.18.7.1527.

Hmeidi, Ismail, Mahmoud Al-Ayyoub, Nawaf A. Abdulla, Abdalrahman A. Almodawar, Raddad Abooraig, and Nizar A. Mahyoub (2015). “Automatic Arabic text categorization: A comprehensive comparative study”. In: *Journal of Information Science* 41.1, pp. 114–124. ISSN: 17416485. DOI: 10.1177/0165551514558172.

Hmeidi, Ismail, Mahmoud Al-Ayyoub, Nizar A. Mahyoub, and Mohammed A. Shehab (2016). “A lexicon based approach for classifying Arabic multi-labeled text”. In: *International*

Journal of Web Information Systems 12.4, pp. 504–532. ISSN: 17440092. DOI: 10.1108/IJWIS-01-2016-0002.

Hmeidi, Ismail, Bilal Hawashin, and Eyas El-Qawasmeh (2008). "Performance of KNN and SVM classifiers on full word Arabic articles". In: *Advanced Engineering Informatics* 22.1, pp. 106–111. ISSN: 14740346. DOI: 10.1016/j.aei.2007.12.001.

Hochreiter, Sepp and J?urgen Schmidhuber (1997). "Long short-term memory". In: *Neural Computation* 9.8, pp. 1735–1780.

Hong, Taekeun, Chang Choi, and Juhyun Shin (2018). "CNN-based malicious user detection in social networks". In: *Concurrency Computation* 30.2, pp. 1–10. ISSN: 15320634. DOI: 10.1002/cpe.4163.

Huang, Faliang, Shichao Zhang, Jilian Zhang, and Ge Yu (2017). "Multimodal learning for topic sentiment analysis in microblogging". In: *Neurocomputing* 253, pp. 144–153. ISSN: 18728286. DOI: 10.1016/j.neucom.2016.10.086.

Huang, Qiongxia, Riqing Chen, Xianghan Zheng, and Zhenxing Dong (2017). "Deep sentiment representation based on CNN and LSTM". In: *Proceedings - 2017 International Conference on Green Informatics, ICGI 2017*, pp. 30–33. DOI: 10.1109/ICGI.2017.45.

Jain, Tejashri Inadarchand and Dipak Nemade (2010). "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis". In: *International Journal of Computer Applications* 7.5, pp. 12–21. ISSN: 09758887. DOI: 10.5120/1160-1453. URL: http://www.ijcaonline.org/volume7/number5/pxc3871453.pdf.

Jones, Karen Sparck (1972). "A statistical interpretation of term specificity and its application in retrieval". In: *Journal of Documentation* 28.1, pp. 11–21. ISSN: 00220418. DOI: 10.1108/eb026526.

Al-kabi, Mohammed N, Amal H Gigieh, Izzat M Alsmadi, and Heider A Wahsheh (2014). "Opinion Mining and Analysis for Arabic Language". In: *International Journal of Advanced Computer Science and Applications* 5.5, pp. 181–195.

Kacser, H. (1991). "A superior theory?" In: *Journal of Theoretical Biology* 149.1, pp. 141–144. ISSN: 10958541. DOI: 10.1016/S0022-5193(05)80076-1.

Kalchbrenner, Nal, Edward Grefenstette, and Phil Blunsom (2014). "A convolutional neural network for modelling sentences". In: *52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference* 1, pp. 655–665.

Kelaiaia, Abdessalem and Hayet Farida Merouani (2013). "Clustering with Probabilistic Topic Models on Arabic Texts Abdessalem". In: *Springer International Publishing Switzerland* 488, pp. 37–46. ISSN: 1860949X. DOI: 10.1007/978-3-319-00560-7. URL: http://www.scopus.com/inward/record.url?eid=2-s2.0-84893026904&partnerID=tZOtx3y1%5Cnhttp://link.springer.com/10.1007/978-3-319-00560-7.

Keyvanpour, Mohammadreza, Zahra Karimi Zandian, and Maryam Heidarypanah (2020). "OMLML: a helpful opinion mining method based on lexicon and machine learning in social networks". In: *Social Network Analysis and Mining* 10.1. ISSN: 18695469. DOI: 10.1007/s13278-019-0622-6. URL: https://doi.org/10.1007/s13278-019-0622-6.

Khan, Farhan Hassan, Usman Qamar, and Saba Bashir (2017). "A semi-supervised approach to sentiment analysis using revised sentiment strength based on SentiWordNet". In: *Knowledge and Information Systems* 51.3, pp. 851–872. ISSN: 02193116. DOI: 10.1007/s10115-016-0993-1.

Khasawneh, Rawan T., Heider A. Wahsheh, Izzat M. Alsmadi, and Mohammed N. Ai-Kabi (2015). "Arabic sentiment polarity identification using a hybrid approach". In: *2015 6th International Conference on Information and Communication Systems, ICICS 2015*, pp. 148–153. DOI: 10.1109/IACS.2015.7103218.

Kim, Jeongin, Dongjin Choi, Byeongkyu Ko, Eunji Lee, and Pankoo Kim (2014). "Extracting user interests on facebook". In: *International Journal of Distributed Sensor Networks* 2014. ISSN: 15501477. DOI: 10.1155/2014/146967.

Kim, Yoon (2014a). "Convolutional Neural Networks for Sentence Classification". In: *arXiv:1408.5882v2*. URL: http://arxiv.org/abs/1408.5882.

— (2014b). "Convolutional Neural Networks for Sentence Classification". In: pp. 1746–1751. ISSN: 10709908. DOI: 10.3115/v1/D14-1181. URL: http://arxiv.org/abs/1408.5882.

Koulali, Rim, Mahmoud El-Haj, and Abdelouafi Meziane (2013). "Arabic Topic Detection using automatic text summarisation". In: *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*. ISSN: 21615322. DOI: 10.1109/AICCSA.2013.6616460.

Koulali, Rim and Abdelouafi Meziane (2014). "Feature Selection for Arabic Topic Detection Using Named Entities". In: *Proceeding of CITALA*, pp. 243–246. URL: https://pdfs.semanticscholar.org/3e13/a1d64fab36068c41145f42c5e873d67d7854.pdf.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances In Neural Information Processing Systems*, pp. 1–9. ISSN: 10495258. DOI: http://dx.doi.org/10.1016/j.protcy.2014.09.007.

Lalji, Thakare Ketan and Sachin N Deshmukh (2016). "Twitter Sentiment Analysis Using Hybrid Approach". In: *International Research Journal of Engineering and Technology (IRJET)*, pp. 2887–2890.

Lam, Savio L.Y. and Dik Kun Lee (1999). "Feature reduction for neural network based text categorization". In: *Proceedings - 6th International Conference on Database Systems for Advanced Applications, DASFAA 1999*, pp. 195–202. DOI: 10.1109/DASFAA.1999.765752.

Lee, Donghyun, Minkyu Lim, Hosung Park, Yoseb Kang, Jeong Sik Park, Gil Jin Jang, and Ji Hwan Kim (2017). "Long short-term memory recurrent neural network-based acoustic model using connectionist temporal classification on a large-scale training corpus". In: *China Communications* 14.9, pp. 23–31. ISSN: 16735447. DOI: 10.1109/CC.2017.8068761.

Lee, Kathy, Diana Palsetia, Ramanathan Narayanan, Md Mostofa Ali Patwary, Ankit Agrawal, and Alok Choudhary (2011). "Twitter trending topic classification". In: *Proceedings - IEEE International Conference on Data Mining, ICDM*, pp. 251–258. ISSN: 15504786. DOI: 10.1109/ICDMW.2011.171.

Lei (2004). "Do Deep Nets Really Need to be Deep?" In: *Informe Final* 59-60.59, p. 55. ISSN: 1659-1216. URL: http://web.catie.ac.cr/informacion/RFCA/rev59/rna59_60_p123_129.pdf.

Lei, Lianxin, Junguo Lu, and Shaohui Ruan (2019). "Hierarchical Recurrent and Convolutional Neural Network Based on Attention for Chinese Document Classification". In: *Proceedings of the 31st Chinese Control and Decision Conference, CCDC 2019*, pp. 809–814. DOI: 10.1109/CCDC.2019.8833090.

Letarte, Gaël, Frédérik Paradis, Philippe Giguère, and François Laviolette (2019). "Importance of Self-Attention for Sentiment Analysis". In: *arXiv:1703.03130v1 [cs.CL] 9 Mar 2017 Published*, pp. 267–275. DOI: 10.18653/v1/w18-5429.

Lewenberg, Yoad, Yoram Bachrach, and Svitlana Volkova (2015). "Using emotions to predict user interest areas in online social networks". In: *Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2015*. DOI: 10.1109/DSAA.2015.7344887.

Li, Si, Hao Zhang, Weiran Xu, and Jun Guo (2014). "Chinese Text Sentiment Analysis Based on Combination Model". In: 10, pp. 1–7.

Li, Xiangsheng, Yanghui Rao, Haoran Xie, Raymond Y K Lau, Jian Yin, and Fu Lee Wang (2017). "Bootstrapping Social Emotion Classification with Semantically Rich Hybrid Neural Networks". In: *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING* 3045.c, pp. 1–16. DOI: 10.1109/TAFFC.2017.2716930.

Liang, Ting Peng and Hung Jen Lai (2002). "Discovering user interests from Web browsing behavior: An application to Internet news services". In: *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 2718–2727. ISSN: 15301605. DOI: 10.1109/HICSS.2002.994214.

Lin, Zhouhan, Minwei Feng, Cicero Nogueira Dos Santos, Mo Yu, Bing Xiang, Bowen Zhou, and Yoshua Bengio (2019). "A structured self-attentive sentence embedding". In: *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, pp. 1–15.

Liu, Chi Harold, Jie Xu, Jian Tang, and Jon Crowcroft (2018). "Social-aware Sequential Modeling of User Interests: A Deep Learning Approach". In: *IEEE Transactions on Knowledge and Data Engineering* XX.XX, pp. 1–13. ISSN: 15582191. DOI: 10.1109/TKDE.2018.2875006.

Lu, Y., Y. Rao, J.Yang, and J.Yin. (2018). "Incorporating Lexicons into LSTM for Sentiment Classification". In: *2018 International Joint Conference on Neural Networks (IJCNN)*, p. 1. ISSN: 21614407. URL: http://mendeley.csuc.cat/fitxers/0f093fd3fff6230dab142add74997c48.

Lu, Yao, Xiangfei Kong, Xiaojun Quan, Wenyin Liu, and Yinlong Xu (2010). "Exploring the sentiment strength of user reviews". In: *Springer-Verlag Berlin Heidelberg* 6184 LNCS, pp. 471–482. ISSN: 03029743. DOI: 10.1007/978-3-642-14246-8{_}46.

Ma, Yukun, Erik Cambria, and Sa Gao (2016). "Label Embedding for Zero-shot Fine-grained Named Entity Typing". In: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 171–180. ISSN: 16101928. DOI: 10.3813/AAA.918956. URL: http://aclweb.org/anthology/C16-1017.

Ma, Yunfei, Yi Zeng, Xu Ren, and Ning Zhong (2011). "User Interests Modeling Based on Multi-source Personal Information Fusion and Semantic Reasoning". In: *Springer-Verlag Berlin Heidelberg* 6335, pp. 195–205. ISSN: 03029743. DOI: 10.1007/978-3-642-15470-6. URL: http://www.scopus.com/inward/record.url?eid=2-s2.0-78649817930&partnerID=tZOtx3y1.

Mahalakshmi, B and K Duraiswamy (2012). "An Overview of Categorization Techniques". In: *International Journal of Modern Engineering Research (IJMER)* 2.5, pp. 3131–3137.

Mangal, Nimita, Rajdeep Niyogi, and Alfredo Milani (2013). "Analysis of Users' Interest Based on Tweets". In: *Springer International Publishing Switzerland* 7971, pp. 12–23. DOI: 10.1007/978-3-642-39637-3. URL: http://link.springer.com/10.1007/978-3-642-39637-3.

Marinai, Simone, Marco Gori, and Giovanni Soda (2005). "Artificial neural networks for document analysis and recognition". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.1, pp. 23–35. ISSN: 01628828. DOI: 10.1109/TPAMI.2005.4.

El-masri, Mazen, Nabeela Altrabsheh, and Hanady Mansour (2017). "Successes and challenges of Arabic sentiment analysis research : a literature review Successes and challenges of Arabic sentiment analysis research : a literature review". In: *Social Network Analysis and Mining* November. ISSN: 1869-5469. DOI: 10.1007/s13278-017-0474-x.

McAuley, Julian and Jure Leskovec (2013). "Hidden factors and hidden topics: Understanding rating dimensions with review text". In: *RecSys 2013 - Proceedings of the 7th ACM Conference on Recommender Systems*, pp. 165–172. DOI: 10.1145/2507157.2507163.

McAuley, Julian, Rahul Pandey, and Jure Leskovec (2015). "Inferring networks of substitutable and complementary products". In: *Proceedings of the ACM SIGKDD Interna-*

tional Conference on Knowledge Discovery and Data Mining 2015-Augus, pp. 785–794. DOI: 10.1145/2783258.2783381. URL: http://arxiv.org/abs/1506.08839.

Medhat, Walaa, Ahmed Hassan, and Hoda Korashy (2014). "Sentiment analysis algorithms and applications: A survey". In: *Ain Shams Engineering Journal* 5.4, pp. 1093–1113. ISSN: 20904479. DOI: 10.1016/j.asej.2014.04.011. URL: http://dx.doi.org/10.1016/j.asej.2014.04.011.

Medrouk, Lisa and Anna Pappa (2017). "Deep learning model for sentiment analysis in multi-lingual corpus". In: *Springer International Publishing* 10634 LNCS, pp. 205–212. ISSN: 16113349. DOI: 10.1007/978-3-319-70087-8{_}22.

Merri, Bart Van and Cifar Senior Fellow (2014). "Learning Phrase Representations using RNN Encoder – Decoder for Statistical Machine Translation". In: *Proceedings ofthe 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1724–1734,* pp. 1724–1734.

Mesleh, Abdelwadood Moh d.A. (2007). "Chi square feature extraction based svms arabic text categorization system". In: *ICSOFT 2007 - 2nd International Conference on Software and Data Technologies, Proceedings* PL.DPS/KE//-, pp. 235–240.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean (2013). "Efficient Estimation of Word Representations in Vector Space". In: *arXiv:1301.3781*, pp. 1–12. URL: http://arxiv.org/abs/1301.3781.

Mikolov, Tomas, Edouard Grave, Piotr Bojanowski, Christian Puhrsch, and Armand Joulin (2017). "Advances in Pre-Training Distributed Word Representations Tomas". In: *arXiv:1712.09405 [cs.CL]* 28.7, pp. 2114–2118. ISSN: 09155287. DOI: 10.1589/jpts.28.2114. URL: http://arxiv.org/abs/1712.09405.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean (2013). "Distributed Representations of Words and Phrases and their Compositionality". In: *arXiv:1310.4546v1*, pp. 1–9. URL: http://arxiv.org/abs/1310.4546.

Miller, George A (1995). "WordNet1.pdf". In: 38.11, pp. 39–41. ISSN: 00993964. DOI: 10.1016/0099-3964(69)90010-6.

Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller (1990). "Introduction to wordnet: An on-line lexical database". In: *International Journal of Lexicography* 3.4, pp. 235–244. ISSN: 09503846. DOI: 10.1093/ijl/3.4.235.

mohamed, BINIZ (2018). *DataSet for Arabic Classification", Mendeley Data, v2*. DOI: 10.17632/v524p5dhpj.2. URL: https://data.mendeley.com/datasets/v524p5dhpj/2.

Mohammed, Ammar and Rania Kora (2019). "Deep learning approaches for Arabic sentiment analysis". In: *Social Network Analysis and Mining* 9.1, pp. 1–12. ISSN: 18695469. DOI: 10.1007/s13278-019-0596-4. URL: https://doi.org/10.1007/s13278-019-0596-4.

Mostafa, Ayman Mohamed (2017). "An Evaluation of Sentiment Analysis and Classification Algorithms for Arabic Textual Data". In: *International Journal of Computer Applications* 158.3, pp. 975–8887. URL: http://www.ijcaonline.org/archives/volume158/number3/mostafa-2017-ijca-912770.pdf.

Muhammad, Aminu, Nirmalie Wiratunga, and Robert Lothian (2016). "Contextual sentiment analysis for social media genres". In: *Knowledge-Based Systems* 108, pp. 92–101. ISSN: 09507051. DOI: 10.1016/j.knosys.2016.05.032. URL: http://dx.doi.org/10.1016/j.knosys.2016.05.032.

N., Mohammed, Izzat M., Amal H., and Heider A. (2014). "Opinion Mining and Analysis for Arabic Language". In: *International Journal of Advanced Computer Science and Applications* 5.5, pp. 181–195. ISSN: 2158107X. DOI: 10.14569/ijacsa.2014.050528.

Nabil, Mahmoud, Mohamed Aly, and Amir Atiya (2015). "ASTD: Arabic Sentiment Tweets Dataset". In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* September, pp. 2515–2519. ISSN: 18770509. DOI: 10.18653/v1/D15-1299. URL: http://aclweb.org/anthology/D15-1299.

Niu, Xiaolei, Hou, Yuexian, Wang, and Panpan (2017). "Bi-Directional LSTM with Quantum Attention Mechanism for Sentence Modeling". In: *Springer International Publishing AG 2017* October, pp. 178–188. DOI: 10.1007/978-3-319-70096-0.

Odeh, Ashraf, Aymen Abu-Errub, Qusai Shambour, and Nidal Turab (2014). "Arabic Text Categorization Algorithm Using Vector Evaluation Method". In: *International Journal of Computer Science and Information Technology* 6.6, pp. 83–92. ISSN: 09754660. DOI: 10.5121/ijcsit.2014.6606.

Ombabi, Abubakr H., Onsa Lazzez, Wael Ouarda, and Adel M. Alimi (2017). "Deep Learning Framework based on Word2Vec and CNN for Users Interests Classification". In: *2017 Sudan Conference on Computer Science and Information Technology (SCCSIT)*, pp. 1–7.

Ombabi, Abubakr H., Wael Ouarda, and Adel M. Alimi (2020). "Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks". In: *Social Network Analysis and Mining* 10.1, pp. 1–13. ISSN: 18695469. DOI: 10.1007/s13278-020-00668-1. URL: https://doi.org/10.1007/s13278-020-00668-1.

Al-Osaimi, Salha and Muhammad Badruddin (2017). "Sentiment Analysis Challenges of Informal Arabic Language". In: *International Journal of Advanced Computer Science and Applications* 8.2, pp. 278–284. ISSN: 2158107X. DOI: 10.14569/ijacsa.2017.080237.

Ouyang, Xi, Pan Zhou, Cheng Hua Li, and Lijun Liu (2015). "Sentiment Analysis Using Convolutional Neural Network". In: *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, pp. 2359–2364.

DOI: 10.1109/CIT/IUCC/DASC/PICOM.2015.349. URL: http://ieeexplore.ieee.org/document/7363395/.

Pal, Subarno, Soumadip Ghosh, and Amitava Nag (2018). “Sentiment Analysis in the Light of LSTM Recurrent Neural Networks”. In: *International Journal of Synthetic Emotions* 9.1, pp. 33–39. ISSN: 1947-9093. DOI: 10.4018/ijse.2018010103.

Pang, Bo and Lillian Lee (2002). “A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts”. In: *Proceedings ofthe ACL*.

Pasquet, J., M. Chaumont, G. Subsol, and M. Derras (2016). “Speeding-up a convolutional neural network by connecting an SVM network”. In: *Proceedings - International Conference on Image Processing, ICIP* 2016-Augus, pp. 2286–2290. ISSN: 15224880. DOI: 10.1109/ICIP.2016.7532766.

Patel, Ankit B., Tan Nguyen, and Richard G. Baraniuk (2016). “A probabilistic framework for deep learning”. In: *Advances in Neural Information Processing Systems* Nips, pp. 2558–2566. ISSN: 10495258.

Pawar, Kishori K and R R Deshmukh (2015). “Twitter Sentiment Classification on Sanders Data using Hybrid Approach”. In: 17.4, pp. 118–123. DOI: 10.9790/0661-1741118123.

Pennacchiotti, Marco and Ana-Maria Popescu (2011). “A Machine Learning Approach to Twitter User Classification Marco”. In: *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* 75.5, pp. 326–327. ISSN: 08949115. DOI: 10.1097/00002060-199609000-00002.

Plagas, Mapa De and Listado De Plagas (2019). “B C G B C G”. In: pp. 1–14.

Prabowo, Rudy and Mike Thelwall (2009). “Sentiment analysis: A combined approach”. In: *Journal of Informetrics* 3.2, pp. 143–157. ISSN: 17511577. DOI: 10.1016/j.joi.2009.01.003.

Preethi, G and P Venkata Krishna (2017). “Application of Deep Learning to Sentiment Analysis for Recommender System on Cloud”. In:

Quispe, Oscar, Alexander Ocsa, and Ricardo Coronado (2018). “Latent semantic indexing and convolutional neural network for multi-label and multi-class text classification”. In: *2017 IEEE Latin American Conference on Computational Intelligence, LA-CCI 2017 - Proceedings*, pp. 1–6. ISBN: 9781538637340. DOI: 10.1109/LA-CCI.2017.8285711.

R. Collobert, J. Weston, L. Bottou, and et al (2011). “Natural Language Processing (Almost) from Scratch Ronan”. In: *Journal ofMachine Learning Research 12 (2011) 2493-2537* 2017-Janua.November, pp. 328–338. ISSN: 1532-4435. DOI: 10.1109/CIC.2017.00050.

Ravi, Kumar and Vadlamani Ravi (2015). *A survey on opinion mining and sentiment analysis: Tasks, approaches and applications*. Vol. 89. June. Elsevier B.V., pp. 14–46. ISBN: 9140235351. DOI: 10.1016/j.knosys.2015.06.015. URL: http://dx.doi.org/10.1016/j.knosys.2015.06.015.

Ravuri, Suman and Andreas Stoicke (2016). "A comparative study of neural network models for lexical intent classification". In: *2015 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2015 - Proceedings* 2, pp. 368–374. DOI: 10.1109/ASRU.2015.7404818.

Saad, Motaz and Wesam Ashour (2010). "OSAC: Open Source Arabic Corpora". In: *6th International Conference on Electrical and Computer Systems (EECS'10), Nov 25-26, 2010, Lefke, Cyprus.* November, pp. 118–123. DOI: 10.13140/2.1.4664.9288. URL: http://site.iugaza.edu.ps/msaad/files/2010/12/mksaad-OSAC-Open-Source-Arabic-Corpora-EECS10-rev8.pdf.

Sahu, Tirath Prasad and Sanjeev Ahuja (2016). "Sentiment analysis of movie reviews: A study on feature selection and classification algorithms". In: *International Conference on Microelectronics, Computing and Communication, MicroCom 2016*. DOI: 10.1109/MicroCom.2016.7522583.

Sainath, Tara N., Oriol Vinyals, Andrew Senior, and Hasim Sak (2015). "Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* 2015-Augus, pp. 4580–4584. ISSN: 15206149. DOI: 10.1109/ICASSP.2015.7178838.

Sasmita, Dhanang Hadhi, Alfan F. Wicaksono, Samuel Louvan, and Mirna Adriani (2017). "Unsupervised aspect-based sentiment analysis on Indonesian restaurant reviews". In: *Proceedings of the 2017 International Conference on Asian Language Processing, IALP 2017* 2018-Janua, pp. 383–386. DOI: 10.1109/IALP.2017.8300623.

Senge, Robin, Juan José del Coz, and Eyke Hüllermeier (2019). "Rectifying Classifier Chains for Multi-Label Classification". In: *arXiv:1906.02915v1*, pp. 151–158. URL: http://arxiv.org/abs/1906.02915.

Seo, Young Woo and Byoung Tak Zhang (2000). "Learning user's preferences by analyzing Web-browsing behaviors". In: *Proceedings of the International Conference on Autonomous Agents* January, pp. 381–387. DOI: 10.1145/336595.337546.

Shen, Tao, Tianyi Zhou, Guodong Long, Jing Jiang, and Chengqi Zhang (2018). "Bi-directional block self-attention for fast and memory-efficient sequence modeling". In: *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, pp. 1–18.

Shen, Yelong, Xiaodong He, Jianfeng Gao, Li Deng, and Grégoire Mesnil (2014). "Learning semantic representations using convolutional neural networks for web search". In: *WWW 2014 Companion - Proceedings of the 23rd International Conference on World Wide Web*, pp. 373–374. DOI: 10.1145/2567948.2577348.

Shickel, Benjamin, Patrick James Tighe, Azra Bihorac, and Parisa Rashidi (2018). "Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health

Record (EHR) Analysis”. In: *IEEE Journal of Biomedical and Health Informatics* 22.5, pp. 1589–1604. ISSN: 21682194. DOI: 10.1109/JBHI.2017.2767063.

Al-Smadi, Mohammad, Bashar Talafha, Mahmoud Al-Ayyoub, and Yaser Jararweh (2018). “Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews”. In: *International Journal of Machine Learning and Cybernetics*, ISSN: 1868-8071. DOI: 10.1007/s13042-018-0799-4. URL: http://dx.doi.org/10.1007/s13042-018-0799-4.

Sohangir, Sahar, Dingding Wang, Anna Pomeranets, and Taghi M. Khoshgoftaar (2018). “Big Data: Deep Learning for financial sentiment analysis”. In: *Journal of Big Data* 5.1. ISSN: 21961115. DOI: 10.1186/s40537-017-0111-6. URL: https://doi.org/10.1186/s40537-017-0111-6.

Soliman, Abu Bakr, Kareem Eissa, and Samhaa R. El-Beltagy (2017). “AraVec: A set of Arabic Word Embedding Models for use in Arabic NLP”. In: *Procedia Computer Science* 117, pp. 256–265. ISSN: 18770509. DOI: 10.1016/j.procs.2017.10.117. URL: https://doi.org/10.1016/j.procs.2017.10.117.

Soucy, Pascal and Guy W. Mineau (2005). “Beyond TFIDF weighting for text categorization in the vector space model”. In: *IJCAI International Joint Conference on Artificial Intelligence*, pp. 1130–1135. ISSN: 10450823.

Sze, Vivienne, Yu Hsin Chen, Tien Ju Yang, and Joel S. Emer (2017). “Efficient Processing of Deep Neural Networks: A Tutorial and Survey”. In: *Proceedings of the IEEE* 105.12, pp. 2295–2329. ISSN: 00189219. DOI: 10.1109/JPROC.2017.2761740.

Taj, Soonh, Baby Bakhtawer Shaikh, and Areej Fatemah Meghji (2019). “Sentiment analysis of news articles: A lexicon based approach”. In: *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies, iCoMET 2019*, pp. 1–5. DOI: 10.1109/ICOMET.2019.8673428.

Tang, Yichuan (2013). “Deep Learning using Linear Support Vector Machines”. In: *arXiv:1306.0239v4*. URL: http://arxiv.org/abs/1306.0239.

Tarwani, Kanchan M and Swathi Edem (2017). “Survey on Recurrent Neural Network in Natural Language Processing”. In: 48.6, pp. 301–304.

Vateekul, Peerapon and Thanabhat Koomsubha (2016). “A study of sentiment analysis using deep learning techniques on Thai Twitter data”. In: *2016 13th International Joint Conference on Computer Science and Software Engineering, JCSSE 2016*, pp. 1–6. DOI: 10.1109/JCSSE.2016.7748849. URL: http://ieeexplore.ieee.org/document/7748849/.

Verma, Shitanshu and Pushpak Bhattacharyya (2009). “Incorporating semantic knowledge for sentiment analysis”. In: *Proceedings of ICON*. URL: http://www.cse.iitb.ac.in/pb/papers/icon09-sa.pdf.

Vo, Quan Hoang, Huy Tien Nguyen, Bac Le, and Minh Le Nguyen (2017). "Multi-channel LSTM-CNN model for Vietnamese sentiment analysis". In: *Proceedings - 2017 9th International Conference on Knowledge and Systems Engineering, KSE 2017* 2017-Janua, pp. 24–29. DOI: 10.1109/KSE.2017.8119429.

Wang, Jian and Zewen Cao (2017). "Chinese Text Sentiment Analysis Using LSTM Network Based on L2 and Nadam". In: pp. 1891–1895.

Wint, Zar Zar, Yuki Manabe, and Masayoshi Aritsugi (2018). "Deep Learning Based Sentiment Classification in Social Network Services Datasets". In: *2018 IEEE International Conference on Big Data, Cloud Computing, Data Science & Engineering (BCD)*, pp. 91–96. DOI: 10.1109/BCD2018.2018.00022. URL: https://ieeexplore.ieee.org/document/8530698/.

Xu, Haotian, Alexander Kotov, Ming Dong, April Idalski Carcone, Dongxiao Zhu, and Sylvie Naar-King (2016). "Text classification with topic-based word embedding and Convolutional Neural Networks". In: *7th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, pp. 88–97. DOI: 10.1145/2975167.2975176.

Xu, Ke, Xushen Zheng, Yi Cai, Huaqing Min, Zhen Gao, Benjin Zhu, Haoran Xie, and Tak Lam Wong (2018). "Improving user recommendation by extracting social topics and interest topics of users in uni-directional social networks". In: *Preprint submitted to Knowledge-Based Systems* 140, pp. 120–133. ISSN: 09507051. DOI: 10.1016/j.knosys.2017.10.031.

Xue, Di Xiu, Rong Zhang, Hui Feng, and Ya Lei Wang (2016). "CNN-SVM for Microvascular Morphological Type Recognition with Data Augmentation". In: *Journal of Medical and Biological Engineering* 36.6, pp. 755–764. ISSN: 21994757. DOI: 10.1007/s40846-016-0182-4.

Yang Y., Pedersen J.O (1997). "A comparative study on feature selection in text categorization". In: *International Conference on Machine Learning, Nashville-Tenn-USA, July 1997*,

Yang, Hsin Chang, Chung Hong Lee, and Chun Yen Wu (2018). "Sentiment Discovery of Social Messages Using Self-Organizing Maps". In: *Cognitive Computation* 10.6, pp. 1152–1166. ISSN: 18669964. DOI: 10.1007/s12559-018-9576-7.

Yang et al. (2017). "Attention-Based LSTM for Target-Dependent Sentiment Classification". In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pp. 5013–5014.

Yenter, Alec and Abhishek Verma (2018). "Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis". In: *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2017* 2018-Janua, pp. 540–546. DOI: 10.1109/UEMCON.2017.8249013.

Yin, Wenpeng, Hinrich Schütze, Bing Xiang, and Bowen Zhou (2016). "Erratum: "ABCNN: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs"". In: *Transactions of the Association for Computational Linguistics* 4, pp. 566–567. ISSN: 2307-387X. DOI: 10.1162/tacl{_}a{_}00244.

Young, Tom, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria (2018a). "Recent trends in deep learning based natural language processing [Review Article]". In: *IEEE Computational Intelligence Magazine* 13.3, pp. 55–75. ISSN: 15566048. DOI: 10.1109/MCI.2018.2840738.

— (2018b). "Recent trends in deep learning based natural language processing [Review Article]". In: *IEEE Computational Intelligence Magazine* 13.3, pp. 55–75. ISSN: 15566048. DOI: 10.1109/MCI.2018.2840738.

Yuan, Shuhan, Xintao Wu, and Yang Xiang (2018). "Incorporating pre-training in long short-term memory networks for tweet classification". In: *Social Network Analysis and Mining* 8.1, ISSN: 18695469. DOI: 10.1007/s13278-018-0530-1. URL: http://dx.doi.org/10.1007/s13278-018-0530-1.

Yue, Lin, Weitong Chen, Xue Li, Wanli Zuo, and Minghao Yin (2018). "A survey of sentiment analysis in social media". In: *Knowledge and Information Systems*, pp. 1–47. ISSN: 02193116. DOI: 10.1007/s10115-018-1236-4. URL: https://doi.org/10.1007/s10115-018-1236-4.

Zaghoul, Fawaz A.L. and Sami Al-Dhaheri (2013). "Arabic text classification based on features reduction using artificial neural networks". In: *Proceedings - UKSim 15th International Conference on Computer Modelling and Simulation, UKSim 2013*, pp. 485–490. DOI: 10.1109/UKSim.2013.135.

Zarrinkalam, Fattane, Hossein Fani, Ebrahim Bagheri, Mohsen Kahani, and Weichang Du (2016). "Semantics-enabled user interest detection from Twitter". In: *Proceedings - 2015 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2015* 1, pp. 469–476. DOI: 10.1109/WI-IAT.2015.182.

Zhang, Qiuyue and Ran Lu (2019). "A Multi-Attention Network for Aspect-Level Sentiment Analysis". In: *Future Internet*.

Zhang, Shuai, Lina Yao, Aixin Sun, and Yi Tay (2017). "Deep Learning based Recommender System: A Survey and New Perspectives". In: *arXiv:1707.07435v6* 1.1, pp. 1–35. URL: http://arxiv.org/abs/1707.07435.

Zhang, Y., D. Song, X .li, and P. Zhang (2014). "Unsupervised Sentiment Analysis of Twitter Posts Using Density Matrix Representation Yazhou". In: *Springer-Verlag* 68.5, pp. 316–329. ISSN: 0022-0418. DOI: 10.1108/jd.2012.27868eaa.002.

Zhang, Yangsen, Jia Zheng, Yuru Jiang, Gaijuan Huang, and Ruoyu Chen (2019). "A text sentiment classification modeling method based on coordinated CNN-LSTM-attention model". In: *Chinese Journal of Electronics* 28.1, pp. 120–126. ISSN: 10224653. DOI: 10.1049/cje.2018.11.004.

Zhou, Yujun, Jiaming Xu, Jie Cao, Bo Xu, Changliang Li, and Bo Xu (2017). "Hybrid Attention Networks for Chinese Short Text Classification". In: *Computacion y Sistemas* 21.4, pp. 759–769. ISSN: 20079737. DOI: 10.13053/CyS-21-4-2847.

Zrigui, Mounir, Rami Ayadi, Mourad Mars, and Mohsen Maraoui (2012). "Based on Latent Dirichlet Allocation". In: *Journal of Computer Information Systems* 20.2, pp. 125–140. ISSN: 13301136. DOI: 10.2498/cit.1001770.