1-1-2021

The Impact of Arabic Part of Speech Tagging on Sentiment Analysis: A New Corpus and Deep Learning Approach

Abdul Munem Nerabie
British University in Dubai

Manar AlKhatib
British University in Dubai

Sujith Samuel Mathew
Zayed University

May El Barachi
University of Wollongong in Dubai

Farhad Oroumchian
University of Wollongong in Dubai

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 184 (2021) 148–155

Procedia Computer Science

www.elsevier.com/locate/procedia

The 12th International Conference on Ambient Systems, Networks and Technologies (ANT)
March 23-26, 2021, Warsaw, Poland

The Impact of Arabic Part of Speech Tagging on Sentiment Analysis: A New Corpus and Deep Learning Approach

Abdul Munem Nerabie[a], Manar AlKhatib[a], Sujith Samuel Mathew[b*], May El Barachi[c], and Farhad Oroumchian[c]

[a]The British University in Dubai, Dubai, UAE, [b]Zayed University, Abu Dhabi, UAE, and [c]University of Wollongong Dubai, Dubai, UAE

Abstract

Sentiment Analysis is achieved by using Natural Language Processing (NLP) techniques and finds wide applications in analyzing social media content to determine people's opinions, attitudes, and emotions toward entities, individuals, issues, events, or topics. The accuracy of sentiment analysis depends on automatic Part-of-Speech (PoS) tagging which is required to label words according to grammatical categories. The challenge of analyzing the Arabic language has found considerable research interest, but now the challenge is amplified with the addition of social media dialects. While numerous morphological analyzers and PoS taggers were proposed for Modern Standard Arabic (MSA), we are now witnessing an increased interest in applying those techniques to the Arabic dialect that is prominent in social media. Indeed, social media texts (e.g. posts, comments, and replies) differ significantly from MSA texts in terms of vocabulary and grammatical structure. Such differences call for reviewing the PoS tagging methods to adapt social media texts. Furthermore, the lack of sufficiently large and diverse social media text corpora constitutes one of the reasons that automatic PoS tagging of social media content has been rarely studied. In this paper, we address those limitations by proposing a novel Arabic social media text corpus that is enriched with complete PoS information, including tags, lemmas, and synonyms. The proposed corpus constitutes the largest manually annotated Arabic corpus to date, with more than 5 million tokens, 238,600 MSA texts, and words from Arabic social media dialect, collected from 65,000 online users' accounts. Furthermore, our proposed corpus was used to train a custom Long Short-Term Memory deep learning model and showed excellent performance in terms of sentiment classification accuracy and F1-score. The obtained results demonstrate that the use of a diverse corpus that is enriched with PoS information significantly enhances the performance of social media analysis techniques and opens the door for advanced features such as opinion mining and emotion intelligence.

Keywords: Sentiment Analysis; Part of Speech Tagging; Arabic Language; Dialect Arabic; Neural Network.

* Corresponding Author: Tel: +97125993280
E-mail address: sujith.mathew@zu.ac.ae

10.1016/j.procs.2021.03.026

1 Introduction

Due to its morphological richness and complexity, the Arabic language presents a challenge when it comes to Part of Speech (PoS) Tagging. In simple terms, PoS Tagging is a process of labeling the different words in a sentence according to their grammatical categories like adjective, article, noun, or verb. Though the definition is simple, it has wide applications, including sentiment analysis and word sense disambiguation, among many others. Very early on, Habash and Rambow demonstrated the fact that Arabic morphology is indeed incredibly complicated [1]. Their research states that if Arabic was to be morphologically analyzed based on features such as primary partof-speech (noun, verb, or adjective), voice, gender, number, or clitics details, there would be around 333,000 morphological tags with only 2,200 of which are currently found in the Penn Arabic Treebank's (ATB) first 280,000 words. Whereas if the same was done to the English language, it would only have morphological tag-sets of roughly 50 tags for all morphological variations. Kübler and Mohamed also affirmed that the language's morphological complexity may imply difficulties for Arabic PoS Tagging, since root words may refer to thousands of different word, which in turn may contribute to problems related to data scattering [2]. Furthermore, since Arabic words are complex, the PoS tags containing morphological information may refer to segments instead of whole words.

Research by Hadni, Lachkar, Meknassi, and Ouatik [3] presented the three approaches typically used for part of speech tagging, namely: (1) The rule-based approach; (2) The statistical approach; and (3) The hybrid approach. The rule-based approach involves establishing a set of knowledge-based rules created by linguists to determine how and where to designate POS tags specifically. The statistical approach requires building an adaptable model and calculating its parameters using the previously-tagged collection of texts. Once this is completed, previously established patterns are used to decide the tagger for certain texts. In general, efficient statistical taggers are primarily identified based on Hidden Markov Models (HMM). Lastly, the hybrid approach consists of integrating the rule-based approach with the statistical one. The latter is being used by most of the PoS Taggers lately, as better results are observed. The integration of PoS Tagging and Neural Networks is also finding its place among researchers. The concept of Neural Networks or Artificial Neural Networks (ANN) is defined as a machine learning approach that allows the processing of training data so that the computer learns a new function. Its learning is quite a long and complex process as the idea is based on the biological neural networks made up of nerve cells that naturally process information in most living organisms. The learning in these neural networks starts with the training data or the *input layer* that is then processed through many different *hidden layers* generating a *weight* represented by a number as it passes each one. As the weights pass through each subsequent layer, they are calculated in various and intricate ways until the output is derived. Throughout the training period, the weights are continually tested and adjusted until the training data that belongs to the same category produces similar results [4]. Though challenging, the usage of neural networks increases by the day because the applications are invaluable, especially with the increased use of the Arabic language in social media.

Sentiment analysis techniques extract sentiments associated with specific subjects from an online document. Analysis of opinions is a task requiring a deep understanding of the textual context, domain knowledge, and linguistics [5]. Most Natural Language Processing (NLP) methods used for sentiment analysis require an accurate Part-of-Speech (PoS) tag information for a given text. This information is provided by automatic PoS tagging. While morphological analyzers and PoS taggers for the Arabic language were proposed [6], those existing solutions do not support Dialect Arabic texts, such as the ones found on social media. Indeed, Arabic social media content is predominantly expressed in Dialect Arabic, which differs from Modern Standard Arabic (MSA) in its vocabulary, semantics, and grammatical structure. Those fundamental differences require the adaptation of PoS tagging methods to the nature of social media texts. In addition, the lack of social media text corpora that are large enough to train a tagging model contributes to the fact that automatic PoS tagging of social media texts has been rarely studied.

Indeed, existing corpora contain MSA terms that were extracted from newspapers' content.

Enabling the automatic analysis of social media content to extract sentiment and opinion intelligence is of great importance to governments and businesses. While, businesses are interested in learning about the people's opinions of their brands, their products, and their reputation, they are also interested in early recognition of new trends (e.g. fashion, sports, and wellness trends). This information can be used to refine product offerings and gain a competitive edge in the market. On the other hand, governments monitor public opinion to understand prevalent views about the current policies and events, as well as identify extreme views and trends in public views that may represent problematic situations. For instance, a negative public opinion that keeps building up about a certain topic/situation over some time may lead to riots and civic disorder. To the best of our knowledge, there exists no standard method for the annotation and PoS tagging of Arabic social media texts in an accurate manner. In this work, we address this challenge by: 1) Proposing a novel Arabic social media text corpus containing more than 5 million annotated tokens and 238,600 MSA and Dialect Arabic words collected from online content; 2) Defining custom annotation rules that are suitable for social media text characteristics; and 3) Proposing a deep learning approach for sentiment analysis that demonstrates the value of PoS tagging in enabling the achievement of superior sentiment classification accuracy.

The rest of the paper is organized as follows, in Section 2, the related work is discussed. This is followed by a presentation of our system architecture in Section 3, and a discussion of our corpus development and experimental results in Section 4. Finally, we conclude in Section 5.

2 Related Work

Part of Speech Tagging is the process of labeling or tagging a word in a body of text (or corpus) according to its corresponding part of speech. A particular word may be tagged as a noun or a verb, but in some cases, a word may hold another meaning contextually and, therefore, be tagged differently. In this section, we discuss some of the relevant works concerning Arabic PoS tagging.

2.1 Arabic Language PoS Tagging

PoS Tagging can be complicated for languages other than English, especially those with different scripts and intricate grammar rules, such as the Arabic language. Addressing this challenge, several academics have taken up the Arabic language as a subject for researches on PoS Tagging.

Demonstrating a few of the earliest Arabic part-of-speech tagger results was how the APT tagger came into existence, marking the start of Arabic PoS Tagging development [7]. The challenges posed by the development in Arabic PoS taggers were defined when developing Brill's "rule-based" approach for Arabic PoS [8]. In its early stages, an architecture was developed for a Web-based Arabic tagger [9]. A supervised learning algorithm, Support Vector Machines (SVM), attained a 95.49% accuracy [10]. Similarly, another SVM-based algorithm was developed, in which it uses the Viterbi algorithm in decoding [1]. A reliability of 97% was accomplished for modern standard Arabic with HMM (Hidden Markov Model) PoS tagging for Arabic [11]. Whereas, when assessed on Egyptian Colloquial Arabic and the Levantine Arabic, it only achieved an accuracy of 69.83% [12] (Duh and Kirchhoff, 2005). Earlier on, the utilization of memory-based learning was investigated for morphological study and PoS tagging in written Arabic [13]. In addition to this, tagging Arabic words through an integration of rule-based methods on machine learning techniques was also explored [14]. A PoS tagging method was developed and implemented through the incorporation of a multi-agent architecture for vowel-marked Arabic texts [15]. The precision of the Arabic Morphosyntactic Tagger (AMT), which utilizes both pattern-based, and lexical and contextual methods, was identified to be 91% [16]. A research on PoS Tagger for Arabic Text [17] has included a summary of Arabic PoS Tagger development throughout the years.

As is evident by the current literature, the development of Arabic PoS taggers has come a long way, with each new tagger being more effective and more elaborate than the previous one.

2.2 Neural Networks based PoS Tagging

Using Neural Networks to build and improve the performance of PoS tagging is a relatively new and an area that is not well investigated. Research by Zheng explored the use of deep learning for Chinese word segmentation and PoS tagging [18]. In this work, neural networks were trained with a perceptron-style algorithm (PSA). PSA-based neural networks are simple to execute and have an advantage in terms of speed over the maximum likelihood scheme, and the lower chance for mal-performance. The use of such neural network resulted in character representations for large unlabelled corpus and achieved an advanced output [18].

Abumalloh et. Al., [19] presented an approach for an Arabic PoS tagger, based on neural network modelling. As they concluded their research, they stated that Artificial Neural Networks (ANNs) have been widely recognized in recent years as common models, especially in the field of computational linguistics. One of the many advantages this technique offers is that it is useful in cases when the linguistic and grammar laws of a language are not understood due to the complexity of the language itself. Therefore, to gather knowledge, ANNs must be initially trained to learn so that expected or consistent outcomes are obtained for the given input data. The weights of the neuronal connections preserve the information acquired by this method. The ANN sequentially presents the patterns taken from the training vectors, and the weights and then be modified to "remember" the information represented by those vectors. Every input vector included in the training set is usually presented more than one time. This is done in order to generate desired results as the training algorithm must evaluate the convergence of the weights to those values. After training, the ANN became the developed PoS Tagger for the Arabic language. The authors of this work concluded that they were able to obtain encouraging results, and this may therefore bring more opportunities for others to explore this area for future projects.

The studies discussed above have all expressed that PoS tagging based on neural networks have turned out to be extremely effective and have shown promising results. Our aim is to build on those results and build a PoS-enriched sentiment analyser that suits the characteristics of the Arabic language utilized in social media fora.

3 System Architecture

In order to build a performant sentiment analyzer for Arabic social media content, we follow five main steps, as depicted in Fig 1. The first step consists in the extraction of social media posts to build the needed dataset for our framework. Once collected, the data passes through a pre-processing and cleaning phase to remove unwanted symbols and tokens, then prepared to be used for training our Long short-term memory (LSTM) deep learning model. After training the model with different hyper-parameters, we select the configuration yielding the highest F1 score and use it to classify the sentiments of social media posts. In the coming sub-sections, we discuss in more details the pre-processing stages and the design of the LSTM model employed.

Fig.1. Part of Speech Tagging and Sentiment Analysis Framework

3.1 *Data Processing and Cleansing*

In this phase, we clean the social media posts to remove noise and unwanted symbols. The goal of this phase is to maximize the number of words whose embedding can be determined in the pre-trained word embedding model.

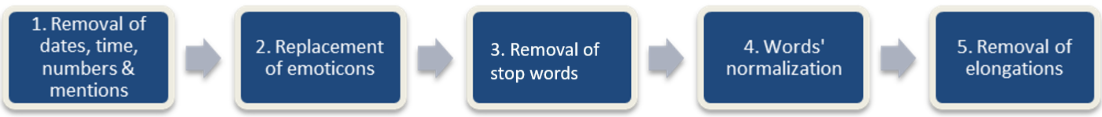

Fig.2. Social media data cleaning and pre-processing

As depicted in Fig 2, the steps we follow for cleaning the Arabic comments are as follows:

- Removing the dates, time, numbers (English and Arabic), URLs, and special symbols such as @ (mention symbol).
- Replacing emoticons encountered in the comments with their emojis.
- Removing the stop words – with the list of stop words encompassing MSA and Dialect Arabic, for example (زي like), (ده this), and) حتى until).
- Normalizing the words and removing unwanted punctuation marks and symbols.
- Removing elongation and using a single occurrence instead.

In relation to step 2, since the word embedding model does not contain representation for emoticons, we developed a mapping between known emoticons to their corresponding emojis. This way we do not lose the sentiment expressed by those emoticons. Emojis are tokenized by placing spaces between them, so that every emoji is looked-up on its own in the word embedding model. This helps when a group of emojis, with no spaces in between, appear in a comment while this combination might not have a corresponding word embedding.

3.2 Data Preparation

In order to prepare the data to be used for training our machine learning model, two operations are conducted:

1. *Post Representation:* Each social media post consists of a set of words that has variable lengths. In the data preparation phase, we replace each word in the comment by its corresponding word embedding from a pre-trained distributed word representation. We use the AraVec [20] word embedding. Each comment is represented as a 2D vector of dimension $n\ x\ d$, where n is the number of words in the comment and d is the dimensional length of the word's vector representation. We use the AraVec skip gram model of dimension 300, i.e., d = 300, to ensure that all the comments are of the same fixed size, by padding each comments representation by zeros. This way each comment will be of size $n\ x\ d$.
2. *Handling dialect words' embedding:* During the substitution of words in the comments by its vector representations, some words may not have a corresponding word embedding. There are different ways to handle these words, either by initializing their vectors with random values and considering them as dialect words (a special words) or by considering them as unknown words with no meaning which are then removed.

3.3 Long Short-Term Memory (LSTM) Model

To enable the accurate classification of sentiments in social media texts, we relied on a LSTM machine learning model, which was configured to the problem at hand. LSTM is a special kind of Recurrent Neural Network (RNN), capable of learning long-term dependencies [21]. LSTMs are explicitly designed to avoid the long-term dependency problem that was encountered with the standard RNN. A bidirectional-LSTM remembers the context of the previous words and the next words context [22].

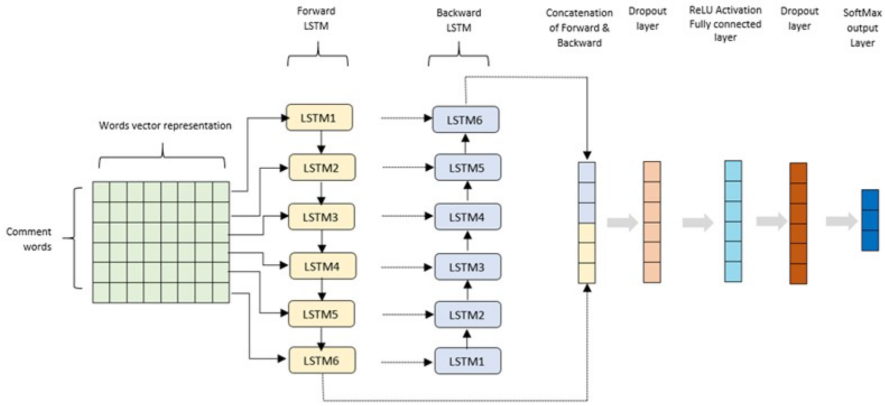

Fig 3: Simplified representation of our LSTM deep learning model for sentiment analysis

As depicted in Fig 3, in our deep learning model, the words vector representation of each social media comment is first passed to each of the LSTMs (forward and backward), both with hidden layer size *h*. The final output of each LSTM is concatenated to yield a vector of length *2h*. This vector is then passed to a fully connected layer with Rectified Linear Unit (ReLU) activation function. A dropout layer is placed after the LSTM layer and another dropout layer is placed after the fully connected layer. Finally, a softmax layer is added to classify the comments' according to the sentiment class as illustrated in 3. We performed many hyper-parameters tuning experiments to select the best LSTM model. The selected model is the one achieving the highest F1-score among the tested models.

4 Corpus Development and Experimental Results

One of the key contributions of this work consists in the development of a new corpus for Arabic social media text analysis that is enriched with PoS information and was used for the training of our deep learning model. We now describe the process of development of this corpus and its unique features, before presenting the experimental results obtained.

4.1 Arabic Social Media Text Corpus Enriched with PoS Information

In this work, we focused on the sentiment analysis of movie reviews, and developed two corpora for training purposes:

- ***Stage 1 – WebCom:*** In the first stage, we developed *WebCom*, a newspaper text corpus that contains web comments collected from ida2at.com[1] – a popular Arabic website treating topics related to Arabic art and literature. The corpus contains MSA web texts and comments written by 56 film reviewers and critics. Comments about 50 movies were collected as raw data, and a subset was selected for manual PoS annotation. The selection of comments is carried out randomly from different users according to their posting frequencies. Each token is annotated with validated PoS tags. Four different types of social media texts were represented, namely: 1) Instagram comments; 2) YouTube comments; 3) Tweets; and 4) Facebook comments.
- ***Stage 2 – WebTrain:*** In the second stage, we developed WebTrain – a social media-based corpus. The corpus was developed using comments from 65,550 users on social media. The selection of comments was conducted randomly from different users based on their posting frequencies, to obtain a corpus where different types of sentiments were represented. Furthermore, each token was manually annotated with validated PoS tags and synonyms, as mean to obtain a customized tagger for social media texts. The resulting corpus contains complete PoS information in the form of manually validate PoS tags, lemmas, and Arabic synonyms based on the WordNet [23]. To the best of our knowledge, WebTrain is currently the largest Arabic text corpus enriched with PoS information.

Further statistical corpus information is given in Table 1.

Table 1: Statistical information of developed corpus

	WebCom	WebTrain	Total
#Comments	520	330,235	330,755
#Tokens	45,730	5,080,896	5,126,626
#Words	8,600	230,000	238,600
#users	56	65,550	65,606

In order to provide a sufficiently large ***training dataset*** for our deep learning model, we combined WebTrain with the WebCom newspaper text corpus. This combination represents the largest manually annotated Arabic corpus, including more than 5 million tokens, 238 thousand words, 330 thousand comments and data collected from 65 thousand users. As for the ***test dataset***, it consisted of 12,519 collected comments and reviews of three popular Arabic movies, namely: “الفلوس’money”, “Les Baghdad لص بغداد " ,and “كازابلانكا Casablanca " . An example of a test comment along with its PoS information is shown in Fig 4 and Table 2.

Fig 4. Sample test social media comment

Table 2: Analysis of sample social media comment

Word	Synonym	PoS	lemmas	Meaning
بجد	فعلاً	Noun	جد	Really
فلم	فيلم	Noun	فلم	Movie
شكلو	من الواضح	Noun	شكل	Its
فاجر	رائع	Adjective	فجر	It is clear, Obviously
شغل	عمل	Noun	شغل	Work
عالي	مرتفع	Adjective	عالي	High
فشخ	عظيم	Adjective	فشخ	Great
معودنا	متعود	Noun	عود	Used to do

4.2 Experimental Results

In order to obtain the optimal configuration for our LSTM model, we varied each hyper-parameter value and calculated the accuracy and F1-score. Table 3 presents the optimal configuration obtained for our model.

Table 3: Optimal LSTM model parameters

Hyper-parameter	Value
LSTM hidden state dimension	200
Number of units in fully connected layer	30
Dropout rate	0.2
Learning rate	0.001
Number of epochs	10
Batch Size	50

To determine the performance of our system in relation to sentiment analysis, we conducted three main experiments. In the first experiment, we used our MSA newspaper text corpus (WebCom) as training dataset for our model and calculated the sentiment classification accuracy and F1-score for three sentiment categories *(positive, negative, and neutral)* when applied to our test dataset. The average classification accuracy obtained in this case was low (34.2%) due to the fact that our test dataset contained mainly Dialect Arabic text, which cannot be tagged and classified accurately by an MSA classifier and PoS tagger. In the second experiment, we trained our model on our social media text corpus (WebTrain), without including PoS information in the training corpus. The results obtained improved significantly when compared to the first experiment, reaching an accuracy of 63.78%. This improvement in performance is since the WebTrain dataset contained Dialect Arabic text, and thus was able to classify most of the test dataset text. In the third experiment, we combined the WebCom and the WebTrain datasets and included PoS information for Dialect Arabic as enrichment to the corpus used as training dataset. The results obtained were significantly higher than the previous two experiments, reaching an average accuracy of 89.31%. This improvement in performance is due to the richness of the corpus employed, which combines MSA with Dialect Arabic words and tokens, and is enriched with PoS information, thus demonstrating the value of PoS tagging for enhancing the results of sentiment analysis. The summary of the experiments' results is shown in Table 4 and Table 5 details the classification accuracies of the three models, with respect to the positive, neutral, and negative sentiment categories. It should be noted that the negative sentiment category yielded higher classification accuracy for all three models, when compared to the positive sentiment category.

Table 4: Model's performance, using different training datasets

Training Dataset	Average Classification Accuracy (%)	Average F1-Score (%)
WebCom, with MSA PoS Tagging	34.2%	33.6%
WebTrain, without PoS tagging	63.78%	61.75%
WebCom with MSA PoS tagging + WebTrain with Dialect Arabic PoS tagging	89.31%	88.26%

Table 5: Sentiment analysis results for the three categories of sentiment

Training Dataset	Classification Accuracy (%)		
	Positive Sentiment	Neutral Sentiment	Negative Sentiment
WebCom, with MSA PoS Tagging	29.4%	35.60%	34.81%
WebTrain, without PoS tagging	60.72%	60.08%	63.11%
WebCom with MSA PoS tagging + WebTrain with Dialect Arabic PoS tagging	85.7%	91.24%	88.65%

5 Conclusion

With the proliferation of social media platforms, there is a vast amount of user-generated content that can be used as source of information and provide an indication of people's sentiments in relation to occurring events. A growth in the population of social media users is observed in the Arab region, which constantly witnesses social, economic, and political changes. In this paper we proposed a new Arabic social media text corpus that is enriched with PoS information and showed the impact of using such corpus along with a deep learning model on the performance of sentiment analysis. Our approach was verified using social media content for Arabic movie reviews and comments.

This research was funded by Zayed University, UAE, from the research incentive fund R19099.

References

[1] Habash, N. and Rambow, O. (2005) 'Arabic tokenization, part-of-speech tagging and morphological disambiguation in one fell swoop', in *ACL-05 - 43rd Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, pp. 573–580.

[2] Mohamed, E. and Kübler, S. (2010) 'Arabic part of speech tagging', in *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010*, pp. 2537–2543.

[3] Hadni, M. et al. (2013) 'Improving Rule-Based Method for Arabic POS Tagging Using HMM Technique', in airccj.org, pp. 257–269.

[4] Medhat, W., Hassan, A. and Korashy, H. (2014) 'Sentiment analysis algorithms and applications: A survey', *Ain Shams Engineering Journal*. Ain Shams University, 5(4), pp. 1093–1113.

[5] Hardesty, L. (2017) *Explained: Neural networks | MIT News*, *MIT News*. Available at: http://news.mit.edu/2017/explained-neural-networksdeep-learning-0414 (Accessed: 2 August 2020).

[6] Pasha, A. *et al.* (2014) 'MADAMIRA: A fast, comprehensive tool for morphological analysis and disambiguation of Arabic', in *Proceedings of the 9th International Conference on Language Resources and Evaluation, LREC 2014*, pp. 1094–1101.

[7] Khoja, S. (2001) *APT: Arabic Part-of-speech Tagger*, *pdfs.semanticscholar.org*. Available at: https://pdfs.semanticscholar.org/4072/d185e733726fca0861398f23b03d84eaf2a8.pdf (Accessed: 2 August 2020).

[8] Freeman, A. (2001) Brill's POS tagger and a Morphology parser for Arabic', *ACL 2001 Workshop on Data-Driven Machine Translation,* p.7.

[9] HM Harmain (2004) 'Arabic part-of-speech tagging', in *The Fifth Annual UAE University Research, 2004*. The Fifth Annual UAE University Research.

[10] Diab, M., Hacioglu, K. and Jurafsky, D. (2004) 'Automatic Tagging of Arabic Text: From Raw Text to Base Phrase Chunks', *HLT-NAACL 2004: Short Papers*, pp. 149–152. Available at: http://cnts.uia.ac.be/conll2003/ner/bin/conlleval. (Accessed: 29 July 2020).

[11] Shamsi, F. Al and Guessoum, A. (2006) 'A Hidden Markov Model –Based POS Tagger for Arabic', *Proceeding of the 8th International Conference on the Statistical Analysis of Textual Data*, pp. 31–42. Available at: http://eserver.org/langs/qalam.txt (Accessed: 29 July 2020).

[12] Duh, K. and Kirchhoff, K. (2005) 'POS tagging of dialectal Arabic', in *aclweb.org*, p. 55. doi: 10.3115/1621787.1621798.

[13] Marsi, E., van den Bosch, A. and Soudi, A. (2005) 'Memory-based morphological analysis generation and part-of-speech tagging of Arabic', in *aclweb.org*, p. 1.

[14] Tlili-Guia, Y. (2006) 'Hybrid Method for Tagging Arabic Text', *Journal of Computer Science*, 2(3), pp. 245–248.

[15] Zribi, C. B. O., Torjmen, A. and Ahmed, M. Ben (2006) 'An efficient multi-agent system combining POS-taggers for arabic texts', in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 121– 131.

[16] Alqrainy, S. (2008) *A morphological-syntactical analysis approach for Arabic textual tagging*.

[17] Sawalha, M. and Atwell, E. (2010) 'Fine-grain morphological analyzer and part-of-speech tagger for Arabic text', in *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010*, pp. 1258–1265.

[18] Zheng, X., Chen, H. and Xu, T. (2013) 'Deep learning for Chinese word segmentation and POS tagging', in *EMNLP 2013 - 2013 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*. Association for Computational Linguistics, pp. 647–657.

[19] Ali Abumalloh, R. *et al.* (2018) 'Arabic Part-of-Speech Tagger, an Approach Based on Neural Network Modelling', *International Journal of Engineering & Technology*, 7(2.29), p. 742. doi: 10.14419/ijet.v7i2.29.14009.

[20] Soliman, A. B., Eissa, K. and El-Beltagy, S. R. (2017) 'AraVec: A set of Arabic Word Embedding Models for use in Arabic NLP', in *Procedia Computer Science*. Elsevier B.V., pp. 256–265.

[21] Hochreiter, S. and Schmidhuber, J. (1997) 'Long Short-Term Memory', *Neural Computation*. MIT Press Journals, 9(8), pp. 1735–1780.

[22] Heikal, M., Torki, M. and El-Makky, N. (2018) 'Sentiment Analysis of Arabic Tweets using Deep Learning', in *Procedia Computer Science*. Elsevier B.V., pp. 114–122.

[23] Alkhatib, M., Monem, A. A. and Shaalan, K. (2017) 'A Rich Arabic WordNet Resource for Al-Hadith Al-Shareef', in *Procedia Computer Science*, pp. 101–110.