Exploring the Performance of Farasa and CAMeL Taggers for Arabic Dialect Tweets

Aseel Alfaidi
Department of Computer Science and Artificial Intelligence, University of Jeddah, Saudi Arabia
aalfaidi0005.stu@uj.edu.sa

Hajer Alwadei
Department of Computer Science and Artificial Intelligence, University of Jeddah, Saudi Arabia
halwadei.stu@uj.edu.sa

Areej Alshutayri
Department of Computer Science and Artificial Intelligence, University of Jeddah, Saudi Arabia
aoalshutayri@uj.edu.sa

Shahd Alahdal
Department of Computer Science and Artificial Intelligence, University of Jeddah, Saudi Arabia
saalahdal@uj.edu.sa

Abstract: *In Natural Language Processing (NLP), Part Of Speech (POS) tagging is an important step; it is a fundamental requirement for many applications, such as information extraction, machine translation, and grammar checking. Successful POS taggers have been developed for many languages, including Arabic. Currently, the spread of social media has increased the diversity of dialects as people use them in their online communications. Therefore, it has become more difficult for researchers to classify some words that are understood by humans but not computers. In addition, most Arabic POS research focuses on Modern Standard Arabic (MSA), while Dialect Arabic (DA) receives less attention. This paper aims to evaluate the performance of two Arabic taggers when used on dialect Arabic tweets and determine which tagger is the appropriate one, which will accordingly help to improve the existent taggers for dialect Arabic tweets. We used the Farasa and CAMeL taggers, which are commonly used to analyze Arabic texts and are considered the best taggers for Arabic. The results indicate that CAMeL tagger performed better than Farasa tagger, with accuracies of 92% and 83% respectively. In other words, a hybrid POS tagger trained with MSA and DA returns better results than the one trained on MSA.*

Keywords: *Dialect arabic tweets, POS, POS tagging, MSA tagger, farasa tagger, CAMeL tagger.*

Received August 31, 2021; accepted November 27, 2022
https://doi.org/10.34028/iajit/20/3/7

1. Introduction

The rapid development of information technology has led to the development of many algorithms that enable machine learning and perform functions similar to those of the human mind. These algorithms, though, have often proved effective in processing natural languages, which are a means of communicating between people to convey information and disseminate experiences among themselves. Arabic is one of the most important natural languages in the world, as it is the language of Arabs and Muslims internationally and the language of the Qur'an [2]. Therefore, the study and development of programs and applications in Arabic have become a matter of great importance for finding effective solutions to deal with the Arabic language.

Part-Of-Speech (POS) tagging belongs to the field of computational linguistics. This field is considered a branch of artificial intelligence specialties and deals with the logical modeling of natural language from a computational perspective [12]. POS combines two fields:

1. Computer science.
2. Natural Language Processing (NLP).

Computational linguistics focuses on proving linguistics theories using computers [2].

POS tagging can be defined as the process of automatically assigning or the capacity to computationally determine to a POS tag any word or other syntactic marker to all the included words in the corpus [12, 23]. Recent developments in NLP in Arabic have achieved levels of superiority and strength that support language and increase interest in language.

Twitter is a useful and powerful data source for many studies; many studies of NLP, in particular employ Twitter data [8, 13]. Despite the availability of Twitter-based POS taggers and NLP tools for the English language [14], the development of Twitter-based POS taggers for the Arabic language has lagged [8]. Some studies have proposed building Twitter-based POS taggers tailored to the Egyptian dialect only [9]. As a result, we test the effectiveness of POS taggers for modern standard Arabic on dialect Arabic. The contributions of this paper are as follows:

- To use two common POS taggers (originally proposed for Modern Standard Arabic (MSA) for dialect Arabic tweets.
- To identify and analyze errors that occur when tagging dialect Arabic tweets.

- To identifying the best POS taggers for dialect Arabic tweets.

This paper is structured as follows. Section 2 presents related work. Section 3 discusses the background and explains the Arabic language with regard to POS tagging. Section 4 introduces the dataset of dialect Arabic tweets. Section 5 evaluates POS taggers on dialectal Arabic tweets and analyzes the performance of competing POS taggers. Section 6 discusses the analysis, and we mention some of the points experienced with regard to Arabic tweets. Finally, we conclude the paper and mention some suggestions.

2. Related Work

POS tagging is widely studied for several natural languages. For the Arabic language, researchers developed many tagger systems based on different methods years ago. We will focus here on the work relevant to the Arabic language.

Mohammed Albared *et al*. [3] designed a tagger experiment using the Bigram Hidden Markov Model (HMM) for POS tagging in Arabic. The dataset includes two sorts of Arabic texts (classical Arabic text and trendy "normal Arabic"). The training was conducted on a corpus of size 23,146 words and a test set with size 3,485 words. The tagger contains 23 tags such as conjunction, adverb, possessive pronoun, and particle. The performance of the written word HMM was evaluated using various smoothing techniques, including Laplace and Kneser-Ney and also the changed Kneser-Ney. The best result for the model was the modified Kneser-Ney approach. The typical overall accuracy for this tagger was 95.8.

While Sawalha and Atwell [21] designed a detailed syntax set to capture the long established morphological features of the Arabic language while ex- plaining all the features and signs in detail, they developed a POS tagger for annotating a wide range of Arabic text formats and applied them to a sample of the Arabic text from the Qur'an, using the Arabic morphological analyzer algorithm. They applied a part of it to the Arabic web group. It consisted of 100 million words.

Using the high strength of National Grid Services (NGS), they added conducting a spell check to detect and correct spelling errors; they developed the POS tagger tool to be able to reuse texts in a wide range of applications and to define the gold standard for comparative evaluation. Comprehensive coverage of the lexical resource showed that about 85% of the words processed with a lemmatizer recovered all the correct words previously analyzed when referring to a broad coverage dictionary.

Kadim and Lazrek [15] proposed their tagging system would consist of two Hidden Markov Models working in parallel in order to enhance the model. They restructure the Nemlar Arabic corpus and extract a new tag set from diacritics and grammatical rules. They used 40 sentences, containing 845 words. They arrived average accuracy of Tagger1 at 98.22%, an average accuracy of Tagger2 at 75.12%, and average accuracy of Tagger2 with Parallel HMM at 75.38%.

By contrast, Albogamy and Ramsay [4] evaluated three POS taggers for Arabic AMIRA, Morphological Analysis and Disambiguation for Arabic (MADA), and Stanford on a sample size of 390 tweets (5454 words); the accuracy of the three taggers, respectively, were 60.2%, 65.8%, and 49.0%. Based on the errors they encountered, the researchers introduced approaches for avoiding the noisiness of domain by pre and post processing on Arabic tweets and used agreement-based bootstrapping to create labeled training data from large amounts of unlabeled data. They selected the Stanford tagger after measuring the speeds of the products and trained it on the bootstrapped training data. The achieved accuracy ranged from 49% to 74%.

Likewise, Alharbi *et al*. [7] tested the MSA tagger for the Gulf dialect on the Gulf dialect in general. They achieved an accuracy of 75% and were motivated to design the Gulf Arabic (GA) tagger using two methods. First, they used a Support Vector Machine (SVM) tagger, and when adding some features significantly affected the accuracy. Second, they examined the performance of the bi- directional long short-term memory (Bi-LSTM) tagger. It became clear to them that the accuracy was higher. The Bi-LSTM taggers performance was 91.2%, while the best SVM tagger achieved was 85.96% when trained on the Gulf ++ dataset.

But in a different way, Darwish *et al*. [12] presented POS taggers for the four dialects for Arabic tweets based on Conditional Random Fields (CRF). The dataset includes 350 tweets for four major Arabic dialects: Egyptian, Levantine, Gulf, and Maghrebi. The researchers used 18 tags for MSA POS tagging, and they added six dialect-specific tags, two to dialects word (PROG PART and NEG PART), and four tweet-specific tags (HASH, EMOT, MENTION, and URL). They trained a joint model using data from all the dialects to train a POS tagger and analyzed some system errors. The tagger achieved 89.3% average tagging accuracy for all dialects.

These studies represent the efforts to develop POS taggers for MSA or DA. In this work, we will evaluate the Farasa and CAMeL taggers on dialect Arabic tweets, with the intention of seeing the strength of POS tagger toward Arabic tweets, instead of building a separate tagger for Arabic tweets. Also, we will use preprocessing modules to improve the accuracy of the taggers.

3. Background

This section will cover POS tagging for the Arabic

language. We mention examples of POS taggers for Arabic and taggers used to evaluate dialect Arabic tweets.

3.1. Dialectal Arabic

The Arabic language has three variants: Classical Arabic (CA), MSA, and Dialect Arabic (DA). CA is the language used in Quranic and Historical texts, while MSA is an official language of news and has its own standard of writing and grammar. Finally, DA is the most used language, and it is increasingly used in social media.

Many characteristics of the Arabic language make it difficult to use in NLP tasks [23, 24]. Moreover, there is a lack of language resources such as tagged corpora, which is considered essential for research and development in statistical and computational linguistics. This increases the difficulty of developing any NLP system for Arabic [19]. Thus, from various perspectives, Arabic differs from other languages, so taggers that have been developed for other languages may not be suitable for the Arabic language [2].

There are three primary categories of Arabic POS tags: noun, verb, and particle (that is, a preposition or conjunction) [5]. Also, have various subcategories, such as a person, number, gender, case, mood, and other morphological and grammatical features of Arabic [24].

Over the last few years, there has been a significant increase in Twitter usage in the Arabic-speaking world. According to a study conducted by Semiocast and published in 2012, Arabic turned out to be the most popular language on Twitter [5].

Because of the nature of the text platform, tweets are not always written in the official grammar or with correct spelling. Many abbreviations are used. With the appearance of many dialects, different words may have the same meaning, such as (انظر) in MSA, which has the same meaning as (شوف) in dialects. Also, in cases where diacritics are absent, two words may have the same letters but with different meanings for example, the words (ذهبُ), which a verb means "went," and (ذهبَ), which is a noun meaning "gold". These are some is the challenges facing POS taggers in the Arabic language.

3.2. POS Tagging and Tagger for Arabic

Many NLP algorithms rely on POS tagging. The POS tagging falls during the syntactic analysis stage, which specifies the words into their proper POS tag [19, 20]. POS tagging is an essential part of the pre-processing process in a wide range of applications, such as knowledge extraction, machine translation, and sentiment analysis [5]. According to Khoja [16], the first stage for the tagger is the initial tagging, in which the word is searched for in the lexicon and given the appropriate tags for the word. In addition, the first requirement in annotating Arabic text is to create a tag set that can accurately describe and address all of the language's information [16, 23]. POS may be trained using supervised or unsupervised approaches. These include SVM, rule-based approach, Markov model approach, or maximum entropy approach [22].

Researchers and companies had developed Arabic language different taggers. Some of these companies are RDI, Sakhr, and Xerox [2]. The Arabic Part-of-Speech Tagger (APT) is considered the first tagger system for the Arabic language. It contains a tag set from 131 tags. Khoja [16] derived it as an initial tag set based on the grammar of the Arabic language.

Also, the Stanford POS tagger was developed in English, and it was later expanded to other languages, including Arabic. Also, Madamira, is a system for morphological analysis and disambiguation of Arabic. It is the combination of two tools, MADA and AMIRA [18], and performs many tasks relevant to Arabic processing, including POS tagging. To the best of our knowledge, the latest taggers existing for the Arabic language are Farasa and CAMeL. As they will be used in this work, they will be explained in the next section in detail.

3.3. Farasa Tagger

Farasa is an Arabic NLP toolkit serving many tasks as POS tagging. It is fast, accurate, and outperforms state- of-the-art Arabic taggers like MADAMIRA. The Arabic Language Technologies Group has developed the Farasa tagger at Qatar Computing Research Institute (QCRI). Darwish and Mubarak [11] described this tagger as ("insight" in Arabic). It is an SVM-based segmented system that ranks potential word segmentations based on a variety of features and lexicons. The features include stems, prefixes, and suffixes, and their combinations likelihoods, inclusion in lexicons containing valid stems or named persons, and underlying stem models [1] It is easy to use and install and is freely available for academic and research purposes. It also supports all platforms, which proves its superiority over the rest.

Table 1. Tags used in the Farasa tagger.

Tag	Description	Tag	Description
Adv	Adverb	JUS	justification attached to verbs
Conj	Conjunction	FOR- EIGN	non-Arabic or non-MSA words
Noun	Noun	FUT	future particle "s"
VSUFF	Verb suffix	PART	prefix and swf
NUM	Number	CASE	alef of tanween fatha
PREP	Preposition	ADJ	Adjective
PUNC	Punctuation	ADET	Determiner
PART	Particles	ABREV	Abbreviation
PRON	Pronoun	NSUFF	Noun suffix
V	Verb		

It was implemented by Maged Saeed as Farasapy, which is an implementation of Farasa toolkit in the Python programming language. Python is a general purpose language that is popular for many NLP tasks.

The tagger used the simplified Penn Arabic Tree- bank (ATB) tag set proposed by Darwish *et al*. [12] and shown in Table 1 the POS tag set used in this tagger.

3.4. CAMeL Tagger

CAMeL is a tool for Arabic NLP in Python. It was developed by the CAMeL lab at New York University in Abu Dhabi. CAMeL provides many functionalities, one of which is the POS tags which are the union of tags in the MSA and the dialects. CAMeL POS is inspired by the ARZATB tag set and guidelines [17]. Ossama Obeid et al. discussed CAMeL tools, stating that the goal of CAMeL is to facilitate research on adaptation between MSA and other Arabic dialects [17]. The tag set is divided into three categories:

- Proclitics (14 tags)
- Enclitics (2 tags)
- Basewords (39 tags).

Table 2 shows the POS tag set used in this tagger.

Table 2. Tag used in the CAMeL tagger.

Proclitics tags		Baseword tags		Enclitics tags	
Tag	**Description**	**Tag**	**Description**	**Tag**	**Description**
PART_DET	أداة_تعريف	NOUN	اسم	PART_NEG	أداة_نفي
CONJ	عطف_حرف	NOUN_NUM	عدد_اسم	PRON	ضمير
PREP	جر_حرف	NOUN_QUANT	كم_اسم		
PART_NEG	نفي_أداة	ADJ	صفة		
PART_FUT	أداة_استقبال	ADJ_NUM	عدد_صفة		
PART_PROG	مضارعة_أداة	ADJ_COMP	مقارنة_صفة		
CONJ_SUB	ربط_أداة	ADV	ظرف		
PRON_DEM	إشارة_ضمير	VERB	فعل		
PRON_INTERROG	استفهام_ضمير	PUNC	ترقيم_علامة		
PART	أداة	FOREIGN	أجنبي		
PART_CONNECT	ربط_حرف	ABBREV	اختصار		
PART_EMPHATIC	توكيد_أداة	DIGIT	رقم		
PART_RC	شرط_جواب	NOUN_PROP	اسم_علم		
PART_VOC	نداء_أداة				

4. Data

This section presents the datasets that were used for testing the two POS taggers: CAMel and Farasa for each of the taggers used, along with a description of data cleaning and preprocessing.

4.1. Data Description

We used the dialectal Arabic dataset (ASAD dataset) that has been constructed by Alharbi *et al*. [6]. It includes a set of 30,000 tweets comprising 256,221 words. The selected dataset was around 69% of the tweets in 2020, 30% were tweeted in 2019, and the remaining 1% were between 2012 and 2018. The dataset covered many dialects; every accent was specified and represented based on the models developed by Lucidya. Thirty-six percent of the tweets are written in Modern Arabic, 31% of the tweets used the Khaleeji dialect, 22% used the Hijazi dialect, and 10% tweeted in Egyptian dialect.

4.2. Data Preprocessing

To ensure that words were tagged correctly and increase the proposed tagger results, data cleaning was performed. This was necessary because the tweets contain non-Arabic symbol, such as English words which are used in hashtags and usernames for example, (#فايبر_STC) or (@mhrsd_care). Also, some tweets has words with repeated letters and emojis to express feelings or emotions for example, in repeating a character (ه) for (هههههههههههههه) which makes it difficult for the taggers to label the tweets. Two preprocessing tools were used: (Preprocessed Arabic text) by Motaz Saad for removing diacritics, punctuation, and repeating characters; and Tnkeeh, an Arabic preprocessing library for Python, as shown in Figure 1, for removing some tweet mentions, links, hashtags, and English alphabets and digits.

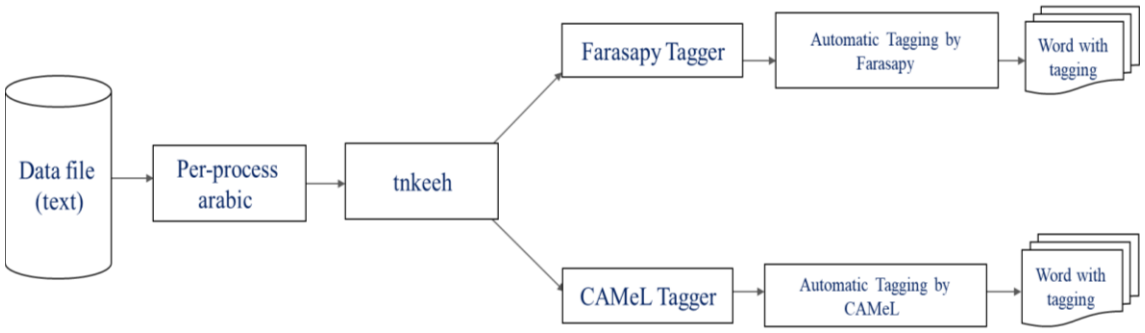

Figure 1. Preprocessing data and tagging data.

5. Results and Discussion

This section explains the evaluation phase of the taggers used in a dialect Arabic tweet and highlights some errors common to both taggers.

5.1. Evaluating POS Tagger

In our assessment of the publicly available POS taggers. The two taggers, Farasa and CAMeL, were used. We compared the performance of each tagger on 400 tweets

and with 3, 623 words extracted from the ASAD dataset. Figure 1 illustrates the method used for tagging data. The performance of these taggers is evaluated based on the output of each tag on words of tweets. Table 3 shows a comparison between Farasa and CAMeL taggers for a single tweet. The similarity between taggers is mostly in particles and sometimes in nouns; they differ most with regard to verbs. The results showed a slight difference in the success rates of the Farasa and CAMeL taggers. The accuracy for Farasa and CAMeL on Arabic tweets is 83% and 92%, respectively. The accuracy was calculated by the percentage of tags on the dataset that are correctly tagged. The formula used to calculate the accuracy is as follows:

$$\text{Accuracy} = \frac{\text{Number of words correctly tagged}}{\text{Total number of Words}} \quad (1)$$

Table 3. Tag Arabic tweets (Farasa, CAMeL tagger).

Tagger	من اروع و اجمل الفعاليات لي حضرتها حقت موسم جدة									
	جدة	موسم	حقت	حضرتها	لي	الفعاليات	اجمل	و	أروع	من
Farasapy	NOUN+NSUFF-FS	NOUN-MS	NOUN+NSUFF-FD	V+PRON+PRON	NOUN-MS	DET+NOUN+NSUFF-FP	NOUN-FP	CONJ	ADJ-MS	PREP
CAMeL	noun	noun	verb	noun	prep	noun	noun	conj	noun	prep

These estimates are almost convergent to the accuracies for the taggers used, where accuracy is 97% for CAMeL [17] and Farasa is 96.2% [11]. Perhaps the reason for the proximity to the accuracy is the dataset, comprising as it does 36% tweets in MSA. Moreover, many tweets have misspelled words as well as abbreviations which may have caused a slight difference between the accuracies, as we will present examples in detail in the following sections.

5.2. Results Analysis

Here, we explain and analyze the outcomes and highlight the different forms of errors of each tagger when tagging the dialect Arabic tweets. Following are error classifications and examples for each.

5.2.1. Dialectal Arabic Words

The CAMeL tagger correctly identifies the dialect words, some word for example (اي) classified as (NOUN_QUANT), which is Noun quantifiers that express either quantity or approximation in a tagger [20]. At the same time, Farasa tagger misclassified it as a subordinate part of the word and may not classify most of the Arabic accent words correctly.

Also, the word such as (راح) is classified as (NOUN-FP) by Farasa, while it is, in fact, a verb, which was correctly classified in the CAMeL tagger.

The word (برا) was classified by both taggers as on the tag (NOUN). But Farasa tagger split the word (بر NOUN-MS+ا CASE). The tag CASE means (alef of tanween fatha). Farasa tagger uses word segmentation, which is especially necessary for a formally rich language such as Arabic. The division of words into their component prefixes, their stem, and their suffixes in Arabic is known as word segmentation [12].

5.2.2. Modern Standard Arabic Words

Farasa was good at tagging MSA, but there were some errors when used with the colloquial dialect. For example, (ربي) was classified as (VERB), while the word (شيء) was tagged as (ADJ). The CAMeL tagger, by contrast, classified both words correctly. That being said, CAMeL does tag some words in MSA incorrectly, though such occurrences are minimal. We found that the majority of the tokens that were mislabeled were so categorized because of morphological annotation, which is used by this tagger for example, (كورونا) is tagged (VERB) instead of (NOUN).

5.2.3. Connected Words

Due to the limited number of characters on Twitter, users may join words, either on purpose or by mistake. CAMeL tags all connected words, such as the word (لوسمحت), as (NOUN_PROP); Farasa tagged the same word as (NOUN-MS), which indicates a similarity between the two.

5.2.4. Abbreviations in the Word

Some Twitter users chose to employ shortened forms of certain words. For instance, they might render the preposition (في) as simply (ف). This word was tagged (ABBREV) in both taggers. Accordingly, it is difficult to define such a word (PREP).

5.2.5. Words with Spelling Mistakes

Some words in Arabic tweets contain spelling mistakes, and the majority of these words were misclassed. It was more difficult for the CAMeL tagger to tag such words correctly. For example, the word (بسم اله) was tagged (اله) as (VERB), while in the original the word (الله) should be tagged as (NOUN). Despite the misspellings, though, Farasa was able to correctly tag this word as (NOUN-MS).

5.2.6. Transliterated Words

Arabic-speaking Twitter users sometimes use untranslated English words and write the words in Arabic letters for example, (online) is rendered as (اون لاين) and (sport) as (سبورت). Interestingly, in such cases, CAMeL Tagger tagged those words as PROP_NOUN, while Farasa tagged them NOUN-MS.

The error rate of both taggers was also analyzed

manually. Table 4 shows that Farasa has a higher error rate (17%) when compared to CAMeL tagger (8%); the single most common errors by Farasa was confusing VERBS with NOUNS, by a ratio of 6%. However, the ratio was very small, with the CAMeL tagger 0.7%. It is possible that this was owing to the lack of diacritics, given that only Farasa was trained on data with diacritics. The most common error in the CAMeL tagger was classifying PREP+NOUN as NOUN_PROP, which reached 6.20%. This is attributable to the connection of words in tweets.

Table 4. Analysis errors in each of the taggers.

CAMeL Tagger		Farasa Tagger	
Error Type	**Rate**	**Error Type**	**Rate**
Verb → Noun	0.7%	Verb → Noun	6%
Noun →Verb	0.06%	Noun →Verb	3%
Noun → Adj	0.1%	Noun→ Adj	2%
Prep → Adv	0.06%	Prep →Noun	1.7%
Noun → noun_prop	0.7%	Adj →Noun	0.4%
Prep+ noun → noun_prop	6.2%	Adv →Noun	3%
		Conj →Noun	0.8%
Total	8%		17%

6. Conclusions and Recommendations

This paper presented clear and simple testing and evaluation of two Arabic language taggers' performance on dialectal Arabic tweets. The results showed that both taggers presented high success rates albeit with some slight errors on Arabic tweets. We used both Farasa and CAMeL taggers on a group of tweets in the colloquial dialect and measured the accuracy rate for each of them. Overall, we achieved an accuracy of 83% in the Farasa tagger and 92% in CAMeL tagger. Moreover, the results showed that the CAMeL tagger had superior results, making it best for the Arabic tweets. The results also indicate that a POS tagger trained on both MSA and DA, such as CAMeL, is highly efficient when applied to Arabic tweets; hence, it is not necessarily to create a dialect-specific POS tagger. Our future work includes evaluating taggers to a wider range of dialects, using data found in [10], which includes five main dialects: Gulf, Iraqi, Egyptian, Levantine, and North African. We suggest some solutions when using MSA taggers for Arabic tweets, including solutions to some errors in section 5, from two aspects.

- The first aspect is to improve the dataset in the processing stage

1. Automatic Arabic spelling error detection and correction can be used in the dataset such as arcorrector, which is a library in Python, before feeding the taggers with the data.
2. The dataset can further be refined, and concatenation words can be separated e.g., (لوسمحت), which is mentioned in the error analysis section). When the words were separated, it resulted in the correct classification in the taggers in the case of the example (لو) was classified (CONJ) and (سمحت) as (VERB).

- The second aspect of the Tagger

1. Add an inventory of words commonly written as abbreviations in DA in Table 5 and training the tagger as a PREP instead of ABBREV.
2. Possibly adding CAMeL tagger with tag sets (e.g., the Stanford English tag set) in case of written foreign words rendered in Arabic letters, as shown in Table 6. Thus, instead of tagging words as NOUN, such words could be classified to their correct/relevant tag in their original language.
3. Adding the tags for special characters in Twitter to the training dataset e.g., the hashtag (#), URLs, emoji, and the (@) sign may give better results where preprocessing step may be skipped.

Table 5. Abbreviations of prepositions in Arabic dialect.

Preposition	Abbreviations
على	ع
في	ف
ما	م
يا	ي
لي	ل
كي	ك

Table 6. Transliterations in Arabic dialect.

Transliteration	English word	Tag
اونلاين	online	Noun
قو	go	Verb
وات	what	Pronoun
از	is	Verb
بلوك	block	Noun
نو	No	Determiner

References

[1] Abdelali A., Darwish K., Durrani N., and Mubarak H., "Farasa: A Fast and Furious Segmenter for Arabic," *in Proceedings of the Conference of The North American Chapter of The Association for Computational Linguistics: Demonstrations*, San Diego, pp. 11-16, 2016.

[2] Abumalloh R., Al-Sarhan H., Ibrahim O., and Abu-Ulbeh W., "Arabic Part-of-speech Tagging," *Journal of Soft Computing and Decision Support Systems*, vol. 3, no. 2, pp. 45-52, 2016.

[3] Albared M., Omar N., Aziz M., Juzaiddin A., and Nazri M., "Automatic Part of Speech Tagging For Arabic: An Experiment Using Bigram Hidden Markov Model," *in Proceedings of the 5th International Conference*, Beijing, pp. 361-370, 2010.

[4] Albogamy F. and Ramsay A., "Fast and Robust POS Tagger For Arabic Tweets Using Agreement-Based Bootstrapping," *in Proceedings of the 10th International Conference on Language Resources and Evaluation*, Portorož, pp. 1500-1506, 2016.

[5] Albogamy F. and Ramsay A., "POS tagging for Arabic tweets," *in Proceedings of the*

International Conference Recent Advances in Natural Language Processing, Hissar, pp. 1-8, 2015.

[6] Alharbi B., Alamro H., Alshehri M., Khayyat Z., Kalkatawi M., Jaber I., and Zhang X., "ASAD: A Twitter-based Benchmark Arabic Sentiment Analysis Dataset," *arXiv preprint arXiv:2011.00578*, 2020.

[7] Alharbi R., Magdy W., Darwish K., AbdelAli A., and Mubarak H., "Part-of-Speech Tagging for Arabic Gulf Dialect using Bi-LSTM," *in Proceedings of the 11th International Conference on Language Resources and Evaluation*, Japan, pp. 3925-3932, 2019.

[8] AlKhwiter W. and Al-Twairesh N., "Part-of-Speech Tagging for Arabic Tweets Using CRF and Bi-LSTM," *Computer Speech and Language*, vol. 65, pp. 101138M, 2021.

[9] Al-Sabbagh R. and Girju R., "A Supervised POS Tagger for Written Arabic Social Networking Corpora," *in Proceedings of the 11th Conf Nat Lang Process KONVENS Empir Methods Nat Lang Process-Proc Conf Nat Lang Process*, Vienna, pp. 39-52, 2012.

[10] Alshutayri A. and Atwell E., "Exploring Twitter as a Source of an Arabic Dialect Corpus," *International Journal of Computational Linguistics*, vol. 8, no. 2, pp. 37-44, 2017.

[11] Darwish K. and Mubarak H., "Farasa: A New Fast and Accurate Arabic Word Segmenter," *in Proceedings of the 10th International Conference on Language Resources and Evaluation*, Portorož, pp. 1070-1074, 2016.

[12] Darwish K., Mubarak H., Abdelali A., and Eldesouki M., "Arabic POS Tagging: Don't Abandon Feature Engineering Just Yet," *in Proceedings of The 3rd Arabic Natural Language Processing Workshop*, Valencia, pp. 130-137, 2017.

[13] Darwish K., Mubarak H., Abdelali A., Eldesouki M., Samih Y., Alharbi R., Magdy W., and Kallmeyer L., "Multi-dialect Arabic POS Tagging: A CRF Approach," *in Proceedings of The Eleventh International Conference on Language Resources And Evaluation*, Miyazaki, pp. 93-98, 2018.

[14] Gimpel K., Schneider N., O'Connor B., Das D., Mills D., Eisenstein J., Heilman M., Yogatama D., Flanigan J., and Smith N., "Part-of-speech Tagging for Twitter: Annotation, Features, and Experiments," *in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics:Shortpapers*, Portland, pp. 42-47, 2011.

[15] Kadim A. and Lazrek A., " Parallel HMM-Based Approach for Arabic Part of Speech Tagging," *The International Arab Journal of Information Technology*, vol. 15, no. 2, pp 341-351, 2018.

[16] Khoja S., "APT : Arabic Part-Of-speech Tagger," *in Proceedings of the Student Workshop at NAACL*, pp. 20-25, 2001.

[17] Obeid O., Zalmout N., Khalifa S., Taji D., Oudah M., Alhafni B., Inoue G., Eryani F., Erdmann A., and Habash N., "CAMeL tools: An Open Source Python Toolkit for Arabic Natural Language Processing," *in Proceedings of the 12th Language Resources and Evaluation Conference*, Marseille, pp. 7022-7032, 2020.

[18] Pasha A., Al-Badrashiny M., Diab M., El Kholy A., Eskander R., Habash N., Pooleery M., Rambow O., and Roth R., "MADAMIRA: A Fast, Comprehensive Tool for Morphological Analysis and Disambiguation of Arabic," *in Proceedings of the Language Resources and Evaluation Conference*, Reykjavik, pp. 1094-1101, 2014.

[19] Salameh S., "A Review of Part of Speech Tagger for Arabic Language," *International Journal of Computation and Applied Sciences*, vol. 4, no. 2, pp. 344-349, 2018.

[20] Sawalha M. and Atwell E., "A standard Tag Set Expounding Traditional Morphological Features for Arabic Language Part-of-speech Tagging," *Word Structure*, vol. 6, no. 1, pp. 43-99, 2013.

[21] Sawalha M. and Atwell E., Fine-Grain Morphological Analyzer and Part-of-Speech Tagger for Arabic Text," *in Proceedings of the 7th Conference on International Language Resources and EvaluationEuropean Language Resources Association*, Valletta, pp. 1258-1265, 2010.

[22] Yousif J. and Al-risi M., "Part of Speech Tagger for Arabic Text Based Support Vector Machines : a Review Part of Speech Tagger for Arabic Text Based Support Vector Machines: a Review," *Journal on Soft Computing*, vol. 9, no. 2, pp. 1867-1873, 2019.

[23] Zeroual I., Lakhouaja A., and Belahbib R., "Towards A Standard Part of Speech Tagset for the Arabic Language," *Journal of King Saud University-Computer and Information Sciences*, vol. 29, no. 2, pp. 171-178, 2017.

[24] Zribi C., Torjmen A., and Ahmed M., "A Multi-Agent System for POS-Tagging Vocalized Arabic Texts," *The International Arab Journal of Information Technology*, vol. 4, no. 4, pp. 322-329, 2007.

Aseel Alfaidi received a BS degree in computer science from Taibah University, Medina, Saudi Arabia, in 2018. She is currently pursuing an MS degree in computer science at the Department of Computer Science and the Artificial Intelligence University of Jeddah, Jeddah, Saudi Arabia. Her research interests include machine learning and data science and voice analysis for disease diagnosis.

Hajer Alwadei received a BSc degree in computer science from King Khaled University, Abha, KSA, in 2018. She is currently pursuing an MSc degree in computer science at the Department of Computer Science and Artificial Intelligence University of Jeddah, Jeddah, KSA. From 2020 to 2021. Now, her major interests in Artificial Intelligence.

Areej Alshutayri is an Assistant Professor in the department of computer science and Artificial Intelligence at the University of Jeddah. Areej collected and created a social media Arabic dialect text corpus (SMADC) using Twitter, Facebook, and Online newspapers. Areej's research interests in using Artificial Intelligence which includes machine learning and natural language processing to understand languages especially Arabic language and its dialects.

Shahd Alahdal is an assistant professor in computer science and artificial intelligence at the University of Jeddah. Shahd’s research interests lie in the area of social computing, exploring psychological theories using machine learning algorithms. Mainly, natural language processing to understand human behavior and emotions in both languages Arabic and English.