Tec Empresarial
P-ISSN: 1659-2395; E-ISSN: 1659-3359

EMPLOYING EXPLAINABLE ARTIFICIAL INTELLIGENCE FOR ADAPTING MULTI-DIALECTICAL ARABIC TEXTIN DECISION-MAKING SCENARIOS

Youmna Abdelwahab[1], Mohamed Kholief[2], Ahmed Ahmed Hesham Sedky[3]
[1,2,3]College of Computing, Arab Academy for Science, Technology, and Maritime Transport. Alexandria, Egypt.
[1]youmna.Abdelfattah5@student.aast.edu; [2]kholief@aast.edu; [3]ahsedkyy@aast.edu;

Abstract
As machine learning has gained prevalence across multiple fields, its complexity, as well as that of deep learning (DL), has continuously increased. Over the last span of years, the use of deep learning across several fields in the aim of advancing, improving and accelerating the classification in general or improving the sentiment analysis in specific, came with it a lot of ambiguity of how these algorithms works for non-experts. which in return researchers started to implement the XAI Algorithms in order to proof and clarify the basic process of how the sentiments classified. While most preexistent studies have utilized the XAI across English and other Latin-based languages for extended reasons, this particular study was utilized to attempt explaining Attention-based long short-term memory results used across the process of Arabic multi-dialect dataset. With the use of local interpretable model-agnostic explanations, our objective was to further demonstrate and simplify the LSTM-led prediction of sentiment polarity process. We attempt to explain how the outcome of attention-based LSTM reaches that of sentiment analysis by applying XAI, thereby yielding potential insights into the study of complex DL models across domains. And finalizing it with a simplified comparison of the probabilistic weights represented on the shown examples with the occurrence of words across the dataset.
Keywords: Arab sentiment analysis, Text-Mining, LSTM, Deep Learning, Multi-dialectical Arabic.

1 Introduction

Machine Learning has been used across several multiple applications, such as medical diagnostics and other domain-specific areas. And with the increase of availability and fast development come along with-it a higher complexity. Despite the increasing prevalence of ML, these models continue to lack explain ability. Because the model outputs must convey information to stakeholders, such descriptions should be written in a human-interpretable language. Consequently, stakeholders would interpret and respond to forecasts more efficiently and with greater confidence. Throughout this study it was mentioned that in [1], some of the main purposes and usage of explaina-ble AI is to make it easier to detect usability, reliability and build trust and fairness. Accordingly, XAI was utilized across several fields covering prior studies in the sole aim of clarifying the proposed classification of models within specific domain fea-tures and confidence in the context of deep learning (DL) and ML. In this particular study,[2] they have proposed SA model for the aim of polarity classification of cus-tomer reviews on a China-based e-commerce website, for approximately 100,000 cus-tomer reviews were collected for testing and training. SA is a field of study that classi-fies opinions and expressions as positive, negative, or neutral [3]. Although there are multiple definitions in the literature, SA is best defined as the analytics used to extract data based on user sentiment [4]. Mood and emotion analysis, also known as opinion mining, involves the study of opinions, thoughts, experiences, feelings, and actions in text form [4]. In previous studies, SA using ML and DL has been applied for accurate polarity classification across various domains and languages. In [5], SA was employed to measure the feasibility of targeting specific European cities, where data were gathered from TripAdvisor online reviews. Although the vast majority of prior studies have focused on English data, several studies have been conducted using Arabic text SA. For instance, in [6], the study applied an SA approach across the Emirati dialect to construct manual annotations pertaining to user reviews. The study was completed by examining the performance of the dataset across ML classifiers. In [7], the SA of various Twitter datasets pertaining to COVID-19 was implemented. The proposed model was used as a precaution rather than as a prediction tool for COVID-19. Subsequently,[8] proposed a long short-term memory (LSTM)-RNN-based DL model with attention layers was proposed. When the model was used to perform an SA on COVID-19-related tweets, it achieved a 20% improvement in performance.

In the previously mentioned studies, the researcher's goal was to create a variegated model that accurately classified the polarity of textual content on platforms. However, they lacked a simplified explanation of how the classification of polarities occurred for non-experts. Therefore, some studies have employed XAI-based models as justifi-cation tools. In previous studies, multiple experiments were performed using multiple DL models at different sentiment levels. One attention-based LSTM approach per-formed best on the entire Arabic word-level SA dataset [9].Therefore, the cardinal con-tribution within this study is to explain and interpret the sentiment analysis process done across mutli-dialect dataset with the help of LIME with attention-based LSTM model with a multi-dialect Arabic text generic dataset. The general approach used in this study is illustrated in Fig1. Section 2 presents a review of the relevant literature, and section 3 summarizes the methods and materials used throughout the study. Finally, Section 4 presents the results and conclusions.

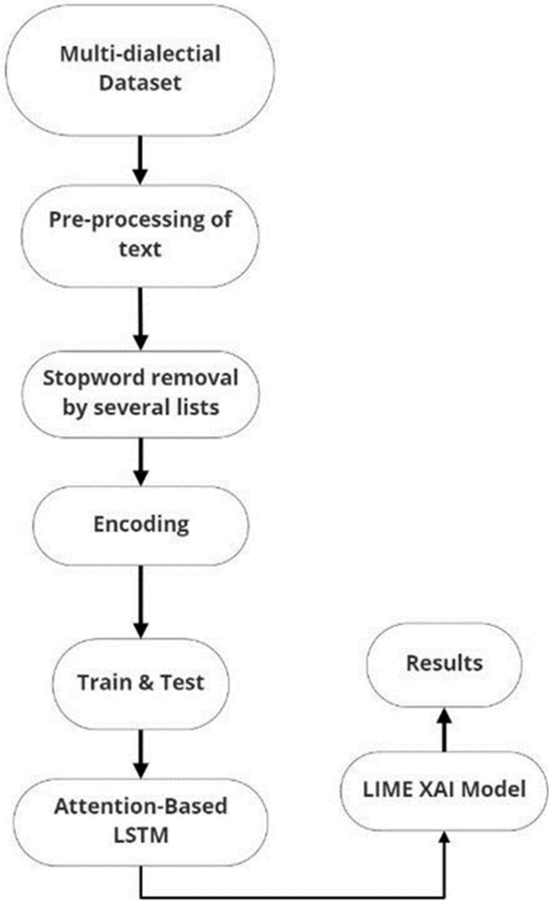

Figure 1. General Approach

2 Literature Review

In prior studies on SA, such as (S. S. Aljameel et al. 2020), which were conducted on COVID-19-related tweets, thedesigned DL models lacked XAI as an interpretation tool. XAI has been employed across a varietyof domains, as in (S. Gite et al. 2020), where XAI methods were used with LSTM-based ML to predict stocks andexplain sentiments associated with headlines, thereby allowing LIME users to improve stockpredictions.Meanwhile,(A. J. I. Alaff, et al. 2021)theysuggested aXAI-basedNB modelto estimatenumberofinfected people and predict prospective and possible future out-breaks from COVID-19 symptomsdisclosed in Turkish Twitter data. In (A. Adak et al. 2022), the features were used to defend a sentiment polarity onLSTM, as well as certain hybrid LSTM-based models, on customer reviews mainly specified onfoodreviewsduringthepandemicusingLIMEandSHapleyAdditiveexPlanations(SHAP).

Owing to resource constraints, (A.A.Aporna et al. 2022) used XAI to classify offensive topics in Bangla's textual data.They provided a graphical presentation showing the association between political and offensivetexts. In (Polley,S.2022), XAI was used for the main of explaining the legal text for lawyers especially withinthe similarities in text with respect to specific aspects. When (I.H.Choi et L. 2020) attempted allocating importantkeywords specially in the IT field using attention-based LSTM, and finalizing their work bycorrelatingtheobtainedwordfrequencywiththecorrespondingLIMEprediction,anddiscovereda new method for identifying job descriptions. One study aimed to further explain Twitter users'sentiments by applying LIME to the proposed BI-LSTM model, enabling the interpretation ofpublic perception across multiple domains (K. R. Chowdhury et al. 2021). Owing to its underlying complexity, the Englishtext dataset was used in the NLP domain to detect sarcasm to aggregate a supervised learningalgorithms (KumarA et al. 2021). While (R. Sharma et al. 2021), the study attempted to use an explainable AI approach on Airbnb datawith the objective of facilitating decision-making processes in the context of large marketingdatasets. Throughout (G.Tang et al. 2021) this study they have utilizes the LIME XAI algorithms across the sourcecodevulnerabilitydetectionfieldtoclearlyexplainandhowtheMLandDLis usedwithinlines ofcodeandtotesthowaccuratetheclassificationofvulnerabilitywithcomplexlinesofcodes.

Although they determined LIME to be an effective tool for vulnerability detection, they discovered a limitation in which the second IF condition in the code samples was not detected. In, (Tay,G et al. 2023) theyhaveconductedastudyfortheaim ofcomparingthe2mainXAIalgorithms inrespectto4aspectswithin the software testing datasets. While in (N.P et al. 2022) the have utilized the XAI to facilitate thedetermination process of key attributes utilized across the diabetes to help in the creation ofdiabetes predictor model for a enhanced classification process. Finally, (G.I.Pérez-Landa et al.2021) used it to understandwhy text within tweets would be considered racist to prevent racism. As listed in Table 1, mostprior studies have employed the LIME approach across several domains within the Englishlanguage owing immense quantity of data and accessible corpus. Compared to English basedstudies lesser studies have applied this method to lower-resource with high-complex morphologylanguagesuchasArabic.Asstatedpreviouslybasedonthe conductedperformance ofLIMEacross prior studies it showcased a promising as an interpretability tool, the present study appliedLIME to Arabic textual data. More specifically, the objective of this study was to applymultidialectal Arabic texts. To further justify the SA classification, LIME was employed todetermine why certain features were specified for particular polarities. Therefore, this studycontributestotheapplicabilityofXAItoTwitter-basedmultidialectalArabicSA.

2.1 Exploratory DataAnalysis

Exploratory data analysis (EDA) is a type of analysis that promotes fine interpretation of data byhelping to understand different attributes and their contributions to the target variable. As EDAreveals conflicting or incomplete data (Pearson RK. 2018) and helps reconcile assumptions and intuitions withreality, it can be used to interpret complex DL models, as is the case with XAI methods. Theforemost XAI algorithms are the LIME and SHAP. LIME is an open-supply framework that wasfirst used by (RibeiroMT, et al. 2016)to explain the predictions of a device. This framework focuses on the selectionprocess of complex ML algorithms and subjective predictions. The method is local in the sensethat the framework analyses selected observations, and interpretable in the sense that its outputmust be apprehended manually. Conversely, SHAP strategies are used to explain the impact ofevery function and enable local and global evaluation of datasets and problems. This approachprimarily employs game theory to explain the output of device-study models (LundbergSM et L. 2017). The LIME waschosen based on a previous study

(R.Guidotti et al. 2020); to suggest an approach for measuring how well served isthelocalexplanationarecorrectinaccordancetosyntheticgroundtruthtoexplanation.
Theexperimental results demonstrate how the proposed approach easily assesses local explanations ofsites and characterizes the quality of local explanation methods. During the valuation of resultsrepresent Local in-text explanation results for word importance explanations showed that LIMEextracted more robust explanations with higher recall and precision compared to SHAP. Inaddition, it returns the best descriptions depending on the identified words in accordance to thewords size used as vocabulary. In addition, the Arabic dialects used in social media platformscontain many words; therefore, a local explainer would work well with the variations used across.In prior studies, LIME produced satisfactory results with English textual content throughoutexceptional domains. In this study, it was employed to analyze multi-dialect Arabic textual contentand to pro-vide explanations regarding how attention-based LSTM classifies the polarity of ASAtextual content. This resulted into promising results of clearer explanation of the Deep learningmodelthatcanbeutilizedtobuildtrustfordecision-makerswhenusingtheblack-boxmodel.

3 Methodology

3.1 Pre-processing

The data used throughout the experiment were obtained from a public dataset compiled by (Boujou E, Chataoui et al. 2021),containing approximately 50,000 tweets from Algeria, Egypt, Lebanon, Tunisia, and Morocco. Nodatelimitwas mentionedintheirwork,andthetextwaslabelledwithnegative,positive,orneutrallabels. Removing punctuation and resource locators: Because all the data were collected fromTwitter, they included certain French and English words. These words and numbers were removedduring the cleaning process. Furthermore, punctuation and URLs within the tweets were removedusing the re.sub method. Removing emoticons and pictographs: Many users employ emoticonswithintweetsto emphasizeorhintattheir emotions.Astheseemoticonsgenerallydo notaffectthetextualinformationitself,theywereremovedusingthere.submethod.
Removing stop words: Stop words appear within the text but do not have a significant effect on theoverall meaning. Two lists of stop words were employed at this stage: a text file provided by (Boujou E, Chataoui et al. 2021)comprising 751 stop words and a list provided by Python's NLTK library. A final cumulative listwas compiled manually after creating a counter for most words within the dataset. Despiteremoving and excluding hashtags, certain Arabic words were not removed automatically but wereremoved manually. Furthermore, be-cause Twitter users generally use informal or slang languages,repetitivecharacters areoftenusedtoemphasizeaparticularfeeling.Whilecharacterrepetitionisa significant contributor for emphasizing a feeling written astride the slang used within social plat-forms. Forexample, theword"ليه"canbewrittenas"ليييييه" toemphasizedisbelief. These words were not removed.

3.2 LSTM attention-based Model

As mentioned previously, the present study was conducted to investigate DL architectures andapproaches for multi-dialect Arabic SA using word-level LSTM models (AbdelwahabY et al. 2021). In the context ofdomain-specific multidialectal texts, prior studies have found that attention-based LSTM yieldsoptimal analytical performance. We extended this notion by applying LIME to an attention-basedLSTM model, thereby improving the interpretability of the emotion classification. During thisinvestigation, the dataset was divided into 80% for training and 20% for testing. The attentionLSTM model was executed on a certain sentiment level which is the word level using tokens in thefirst input layer. Figure 3 illustrates the proposed methodology. The following diagrams illustratehow each layer is within the attention-based LSTM model which includes an embedding layer asan input layer. Following 3 LSTM layers with the dense layer before flattening the outcome andthencon-catenatingtheresultsforthedenselayertoclassifythetweets.

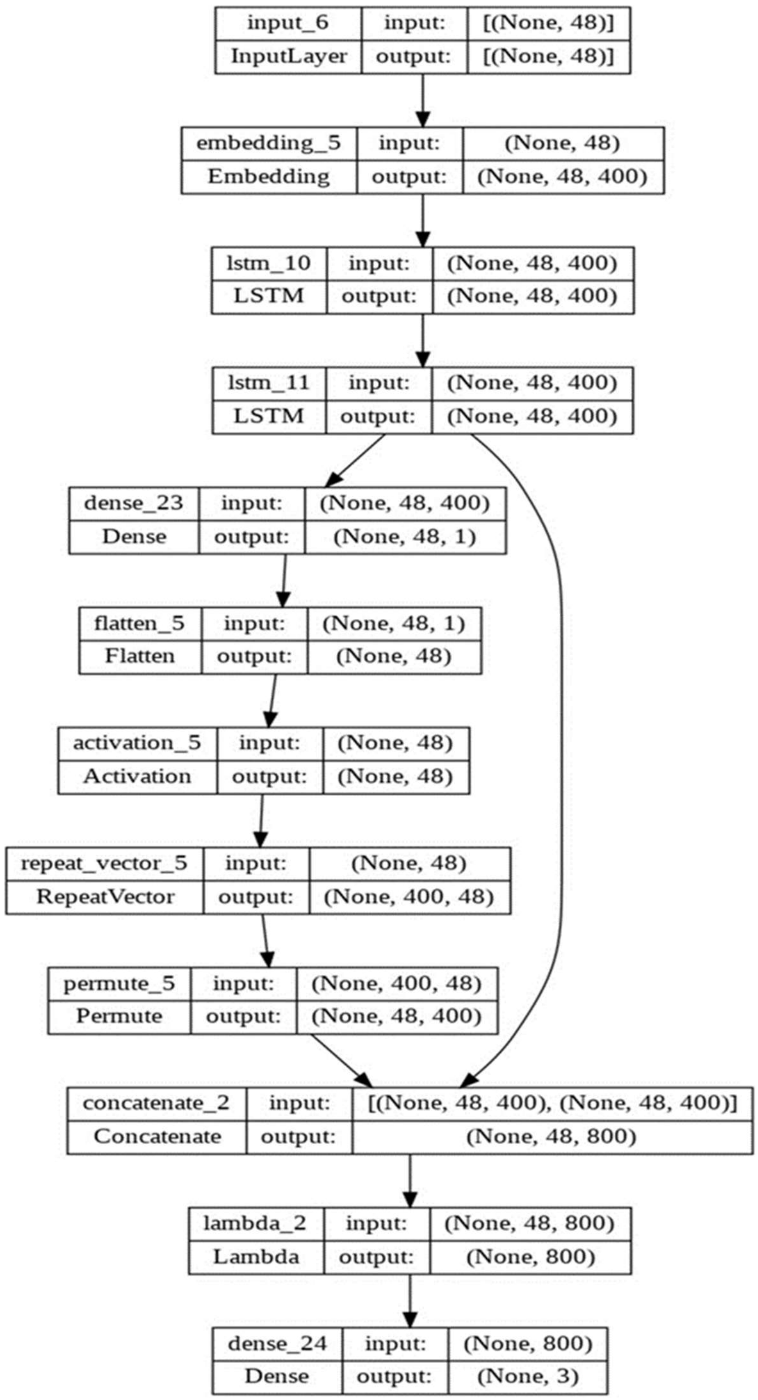

Figure 2.Proposed Model

The accuracy that was reached by the experiment was a 79% on the attention LSTM model byadding an attention layer to enhance and emphasize on the classification accuracy within theArabic text dataset. Due to its complex morphology. Throughout this study we concentrated onhowfocusingonindividualwordsbyevaluatingthewordcountsintheembeddinglayer.

Throughout our study, attention was utilized to enhance the accuracy of the results whenclassifying sentimentpolaritiesacross themodelinFigure2.Thisapproachwaspreviouslyusedby (I.H.Choi et al. 2022)toobtainanimprovementinaccuracywhenclassifyingwordsrelatedtoITjobs.Whenclassifying Bangla text (A.A.Aporna et al. 2022) also considered a low-resource language to detect hate speech, thisapproach yielded 75% accuracy using Bi-LSTM and 78% accuracy using conv-LSTM. In contrast,our attention-based LSTM approach produced a 79% accuracy in the detection of sentimentpolarity.

3.3 Feature Selection

After cleaning and preprocessing, all data were converted into an interpretable format for the DLmodelusinganumericalfeatureextractionmethodknownasvectorization.Inthisprocess,TermFrequency- Inverse document frequency (TF-IDF) was used to convert data into a numericalformat, where each word was represented by a matrix. This process is also known as wordembedding.Theencodingclass wasthenusedtoassignpositivewords withanumericvalueof1and negative words with a numeric value of 0 (sckitlearn,. 2022). Figure 3. presents a visualized word cloudrepresentationofcommonwordsthroughoutthedataset,asproposedin(Boujou E, Chataoui et al. 2021).

Figure 3. A word cloud Presentation of the Dataset

4 Results

4.1.Applying LIMEXAI Model for the Attention-Based LSTM Model

Throughout this study the explainable AI was implemented to explain, simplify and providetransparency for the SA applied conducted on the multi-dialect dataset. A similar approach wasapplied by [11] for COVID-19-related Twitter data to determine potential viral exposures andbreakouts.In contrast, the present study implemented XAI on ASA, and more specifically onmulti-dialectArabictext,becauseof dueitscomplexityandvariations.Furthermore,theLIME XAI model was applied. This approach works as an approximation technique for DL models usingthelocalinterpretablemodeltoexplainindividualpredictions.First,weappliedLIMEtothe multi-dialect dataset proposed in (Boujou E, Chataoui et al. 2021). The presented sentiment, shown in Fig 4, indicates a higherpotential for the loss of life as a result of suicide in comparison to war.

This information ispredictedtobeanegativesentiment. Accordingly, thewordsالحروب(war)andمستوى(level)are classified as negative throughout the figure. This illustrates the selection according to theprobability. The probability of all classes should be equal to 1, where 0.99 of the sentiment is classifiedas negativeand0.01isclassifiedasneutral. Wecanalsoseethatthewordثانية(second) was marked as neutral. Because this word is generally indicative of time, it may be used in eithernegativeorpositivesentiments.

In Fig 5, LIME was utilized on another negative sentiment sample, which indicates a correlationbetween electionsandviraloutbreaks. Throughoutthisfigure, thewordsفيروس(virus)andانتخبات(elections) are classified as negative. As shown in Figure 5, these words were considered the maincontributors to the negative sentiments based on probability. The probability of all classesshouldbeequalto1,where0.91of thesentimentisclassifiedasnegative,0.07isclassifiedasneutral,and0.02isclassifiedaspositive.Wecan observethatthewordالسواد(black)wasmarkedasnegative with a lower weight of 0.09. The lower weight assignment may be a result of the nature of theword itself, which in Arabic may refer to the color, or to something negative or hateful, dependingon the context. As shown in Table 1, we also compared certain words between their local weightsrepresented by LIME, and their overall occurrence throughout the dataset. For example, althoughمش occurred more than 2000 times, LIME assigned it a relatively low weight of 0.07, whereasmoreimpactfulwordswithinacertainsentencewereassignedweightsof0.54and0.19.

In Fig 6. The LIME was used with a positive sentiment that indicates a sentence of gratitudetowards a reporter writing a great article and thanking the reporter the sentence states “Thank youfor your great article, you have explained everything that we wanted to say thanks to you”. Words thatindicatedthepositivedirectionofthis particularsentenceareشكرا(Thankyou),بجد(seriously), وانتي(you),ليكي(toyou)withprobabilities of0.31, 0.21, 0.08and0.07respectively21, 0.08and 0.07respectively.(Thankyou)hadthehighestprobabilitiesbecauseamongthemduetothewholesentencewas beingmoreofagratitudestatementtowardsaperson. Meanwhile, thewordsبحقمحبوب (In name of someone you love) which is mainly used across the Arabic language for bothpurposes. therefore, it was given a 0.05 and 0.06 within the negative impact. According topreviously stated study (K. Fiok et al. 2021), the LIME satisfies two out of the seven purposes stated as the primarygoalsofXAIapplications.

Table 1.Wordcount and LIME-probability weights

Words	LIME-weight	WordsCount
فيروس/Virus	0.54	149
دا/This	0.19	163
السواد/Black	0.09	68
انتخبات/Elections	0.19	84

مش/Not	0.07	2325

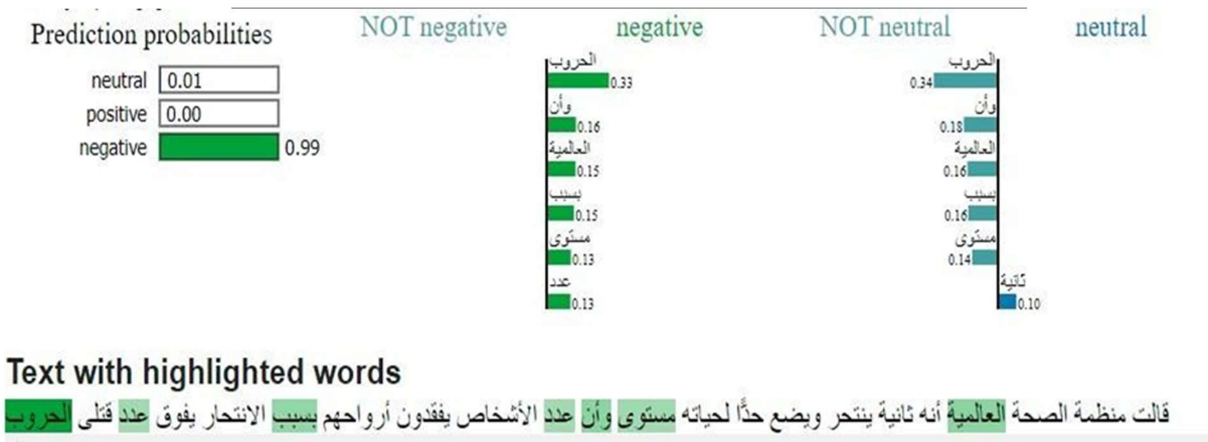

Figure 4. Negative sentiment

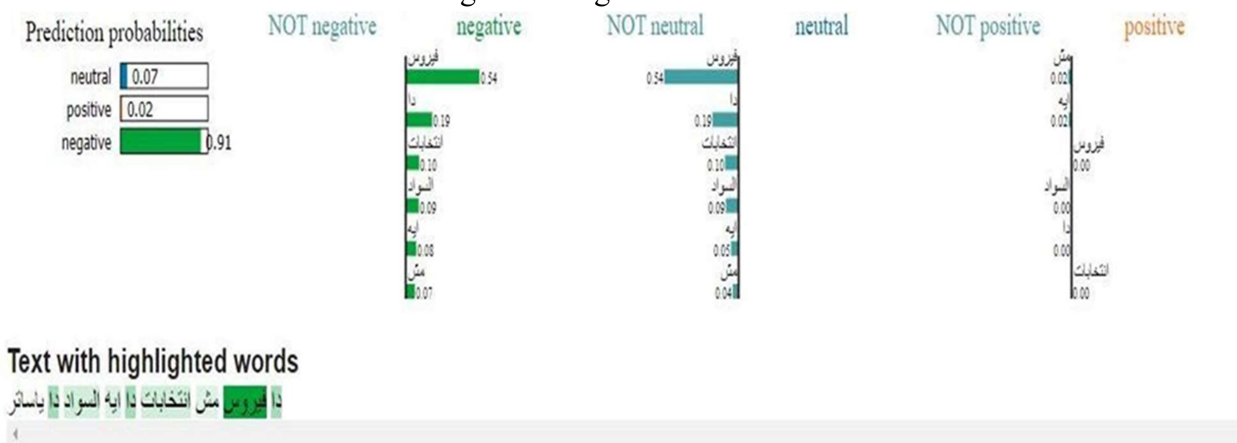

Figure 5. NegativeSentiment

Figure 6. Positive Sentiment

4. Discussion

4.1. AcomparisonwithState-of-ArtAlgorithms

Throughout this study we conducted experimental trials on BI-LSTM, attention-based LSTM andML state of-art algorithms and state-of-art Transformers and obtained results that varied higher

orlower than to those of the proposed attention-based LSTM. Throughout the experiments thevariationofthestate-of-artclassifierswereusedwiththeirsequentialnatureofthealgorithms.
These experiments were done to show the variation of accuracies across a low resource Arabiclanguage with comparison of the attention LSTM model. While throughout the transformer had ahigher accuracy than the attention-based model. The main goal for this concise is to understand,clarify the sentiment analysis classification process across multi-dialectical generic dataset with itsslang variations too. Due to the size and unbalanced nature of the multi-dialect dataset (Boujou E, Chataoui et al. 2021), acomparison with the State- of-Art classifiers were done across Table 2. Where it illustrated avariety of result across the classifiers with the lowest accuracy of 69 across both random forest andK-Neighbors and highest accuracy across SVC with an accuracy of 76 and a close result of BI-LSTM with an accuracy of 78 while the attention-based LSTM model achieved a 79%. This studyalsofocusedontheinterpretabilityofthe classificationprocess,whichwasdemonstratedinFigures4and5.

Classifiers	Accuracy
LogisticRegression	72
DecisionTree	70
LinearSVC	76
RandomForest	69
MultinomialNB	71
K-Neighbors	69
BERT	85
BI_LSTM	78
Attention–BasedLSTM	79

Table 2. StateofArtcomparison

Prior studies have demonstrated several experiments (AbdelwahabY et al. 2023) that reveal the attention-based model asthe highest-performing LSTM model across word-level ASA. In this present study, we applied anLSTM methodology with an attention layer and word count in the embedding layer. The XAIapproach, LIME, was applied to the LSTM model and achieved an accuracy of 79%. The primaryobjective was to interpret the classification of sentiments used within the DL model. We tested theattention-based LSTM model to examine how sentiments are classified into polarities on textualfeatures. We also analyzed how the output of these features, along with their correspondingprobabilities, may help select more appropriate keywords across specific domains within the multi-dialectArabictextcorpus. WeconcludethatLIMEis appropriateformulti-dialectArabictextsduetoitsabilitywithconcentratingcheckingwordswithinalocalsentiment.

Throughout this study a couple of limitations have faced the proposed methodology, first notreaching to the intended accuracy due to the Arabic Morphology complexity, second would bewithin the dataset used which did not contain all of the dialect variations used across the

Arabicspeaking countries and the availability of such datasets which concentrates more across multi-dialectical spoken text across social media services. Third, within the application of SHAP XAIthere was a limitation within the visualization in the manner of the features representation due toitsuniquewritingsystem.Fourth,theBERTworksmoreefficiently withSHAPvaluewhichmadeus reconsider the outcome for the following transformers. As the main contribution with thisincisive study is to clarify, ease and elucidate how can explainable AI be a contributor forclarifying the classification process for non-experts. Finally, our future work will focus onapplying the XAI technique while comparing the results with an Arabic text dictionary to furtherenhance model performance. We intend to compare words with alternate meanings to a database,and also across other sentiment levels other than word level such as sentence and documents level,Furthermore, we intend to curate a dataset that encompasses all Arabic dialects with the help ofspecializedannotatorsforeachdialect.

REFRENCE

1. Fiok, Krzysztof, Farzad V. Farahani, Waldemar Karwowski, and Tareq Ahram.(2022) "Explainable artificial intelligence for education and training." The Journal of Defense Modeling and Simulation 19, no. 2: 133-144. https://doi.org/10.1177/1548512921102.
2. A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser et al.(2020) "Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, vol. 58, pp. 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.
3. Ronen Feldman. 2013. Techniques and applications for sentiment analysis. Commun. ACM 56, 4 (April 2013), 82–89. https://doi.org/10.1145/2436256.2436274ACM Digital Library. https://dl.acm.org/doi/abs/10.1145/2436256.2436274 .
4. L. Yang, Y. Li, J. Wang and R. S. Sherratt,(2020) "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," in IEEE Access, vol. 8, pp. 23522-23530, 2020, 10.1109/ACCESS.2020.2969854
5. Oksana Tokarchuk, Jacob Charles Barr, Claudia Cozzio.(2022)How much is too much? Estimating tourism carrying capacity in urban context using sentiment analysis,Tourism Management,Volume 91,2022,104522, ISSN0261 5177, https://doi.org/10.1016/j.tourman.2022.104522
6. Aljameel, S.S.; Alabbad, D.A.; Alzahrani, N.A.; Alqarni, S.M.; Alamoudi, F.A.; Babili, L.M.; Aljaafary, S.K.; Alshamrani, F.M.(2021) A Sentiment Analysis Approach to Predict an Individual's Awareness of the Precautionary Procedures to Prevent COVID-19 Outbreaks in Saudi Arabia. Int. J. Environ. Res. Public Health 2021, 18, 218. https://doi.org/10.3390/ijerph18010218
7. Abdelwahab Y, Kholief M, Sedky A. (2022) An Experimental Survey of ASA on DL Classifiers Using Multi-Dialect Arabic Texts. Accepted. Future of Information and Communication Conference San Francisco, March,2023.
8. Alaff, A.J.I., Mukhairez, H.H.A., Kose, U. (2021). An Explainable Artificial Intelligence Model for Detecting COVID-19 with Twitter Text Classification: Turkey Case. In: Maji, A.K., Saha, G., Das, S., Basu, S., Tavares, J.M.R.S. (eds) Proceedings of the International Conference

on Computing and Communication Systems. Lecture Notes in Networks and Systems, vol 170. Springer, Singapore. https://doi.org/10.1007/978-981-33-4084-8_9

9. Rathore, R.K., Kolonin, A. (2022). Explorative Study of Explainable Artificial Intelligence Techniques for Sentiment Analysis Applied for English Language. In: Dua, M., Jain, A.K., Yadav, A., Kumar, N., Siarry, P. (eds) Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-5747-4_73.

10. Gite, S., Khatavkar, H., Srivastava, S., Maheshwari, P., Pandey, N. (2021). Stock Prices Prediction from Financial News Articles Using LSTM and XAI. In: Singh, P.K., Wierzchoń, S.T., Tanwar, S., Ganzha, M., Rodrigues, J.J.P.C. (eds) Proceedings of Second International Conference on Computing, Communications, and Cyber-Security. Lecture Notes in Networks and Systems, vol 203. Springer, Singapore. https://doi.org/10.1007/978-981-16-0733-2_11.

11. Adak A, Pradhan B, Shukla N, Alamri A.(2022)Unboxing Deep Learning Model of Food Delivery Service Reviews Using Explainable Artificial Intelligence (XAI) Technique. Foods.11(14):2019. https://doi.org/10.3390/foods11142019

12. Chowdhury, K.R., Sil, A., Shukla, S.R. (2021). Explaining a Black-Box Sentiment Analysis Model with Local Interpretable Model Diagnostics Explanation (LIME). In: Singh, M., Tyagi, V., Gupta, P.K., Flusser, J., Ören, T., Sonawane, V.R. (eds) Advances in Computing and Data Sciences. ICACDS 2021. Communications in Computer and Information Science, vol 1440. Springer, Cham. https://doi.org/10.1007/978-3-030-81462-5_9

13. Kumar, A., Dikshit, S., & Albuquerque, V. H. C. (2021). Explainable artificial intelligence for sarcasm detection in dialogues. Wireless Communications and Mobile Computing, 2021. https://doi.org/10.1155/2021/2939334.

14. In Hyeok Choi, Yang Sok Kim, and Choong Kwon Lee. (2020). A Study of the Classification of IT Jobs Using LSTM and LIME. In The 9th International Conference on Smart Media and Applications (SMA 2020). Association for Computing Machinery, New York, NY, USA, 248–252. https://doi.org/10.1145/3426020.3426083

15. Tang, G. et al. (2021). Interpretation of Learning-Based Automatic Source Code Vulnerability Detection Model Using LIME. In: Qiu, H., Zhang, C., Fei, Z., Qiu, M., Kung, SY. (eds) Knowledge Science, Engineering and Management. KSEM 2021. Lecture Notes in Computer Science(), vol 12817. Springer, Cham. https://doi.org/10.1007/978-3-030-82153-1_23

16. Aporna, A.A., Azad, I., Amlan, N.S., Mehedi, M.H.K., Mahbub, M.J.A., Rasel, A.A. (2022). Classifying Offensive Speech of Bangla Text and Analysis Using Explainable AI. In: Singh, M., Tyagi, V., Gupta, P.K., Flusser, J., Ören, T. (eds) Advances in Computing and Data Sciences. ICACDS 2022. Communications in Computer and Information Science, vol 1613. Springer, Cham. https://doi.org/10.1007/978-3-031-12638-3_12

17. Pérez-Landa, Gabriel Ichcanziho, Octavio Loyola-González, and Miguel Angel Medina-Pérez. (2021) "An Explainable Artificial Intelligence Model for Detecting Xenophobic Tweets." Applied Sciences 11, no. 22: 10801. https://doi.org/10.3390/app112210801.

18. Marco Ribeiro, Sameer Singh, and Carlos Guestrin.(2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, pages 97–101, San Diego, California. Association for Computational Linguistics. doi:10.18653/v1/N16-3020.https://doi.org/10.1145/2939672.2939778.

19. Scott M. Lundberg and Su-In Lee. (2017). A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 4768–4777. https://dl.acm.org/doi/10.5555/3295222.3295230.

20. K. Boyd, “LASIK — laser eye surgery,” American Academy of Ophthalmology, Aug. 23, 2022. Available via https://www.aao.org/eye-health/treatments/lasik (accessed Aug. 23, 2022.

21. Youmna Abdelwahab, “Master_dataset/README.md at main youmna Abdelwahab/Master_dataset,” GitHub. Available https://github.com/youmnaAbdelwahab/Master_dataset/blob/main/README.md. (Accessed: 24-Sep-2022)

22. “TensorFlow Core v2.9.1,” TensorFlow. Available via: https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/text/Tokenizer(accessed Jun. 07, 2022).

23. Basim Mahbooba, Mohan Timilsina, Radhya Sahal, Martin Serrano,(2021) "Explainable Artificial Intelligence (XAI) to Enhance Trust Management in Intrusion Detection Systems Using Decision Tree Model", Complexity, vol. 2021, Article ID 6634811, 11 pages. https://doi.org/10.1155/2021/6634811.