Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.16.(4) 2024,pp.Comp 123–133

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Arabic word Prediction For Next and Previous Word Using Bert & CBOW Algorithms

Hawraa Ali Taher [a] *

[a]*Department of Computer Science, Faculty of Education for Girls, University of Kufa, Najaf, Iraq,Email: hawraa.alshimirty@uokufa.edu.iq*

ARTICLE INFO

Article history:
Received: 20 /09/2024
Rrevised form: 09 /10/2024
Accepted : 31 /10/2024
Available online: 30 /12/2024

Keywords:

Arabic Language

Word Prediction

CBOW

Bert Algorithm

F-Measure

ABSTRACT

One application of Natural Language Processing (NLP) is Next Word Prediction, also known as Language Modeling. This process involves predicting the most likely word to follow in a given sentence based on the preceding context. It has numerous widely used applications, like auto-correct, which is mostly used in emails and messages. It can also be used in Microsoft Word or Google searches to predict the next word based on past searches or global queries. The goal of Natural Language Generation (NLG) is to create language that is human-interpretable and natural. Users find text generation, and next-word prediction in particular, convenient as it makes typing faster and error-free. Consequently, an essential analysis topic for all languages is a personalized text prediction system. This paper suggests a novel approach for predicting the following word in a Arabic sentence. It is possible to minimize the total number of keystrokes a user makes by anticipating the next word in a sequence. In this work, Bert algorithm and Continuous Bag of Words(CBOW) are proposed to predict the next word in Arabic language, and predict the previous word. The Bert Algorithm is achieved the best accuracy , 90% for next word prediction, and 80% for previous word prediction . And, Continuous Bag of Words(CBOW) is achieved the best accuracy , 100% for next word prediction, and 100% for previous word prediction.

MSC..

https://doi.org/ 10.29304/jqcsm.2024.16.41778

1. Introduction

Physically disabled people today use computer programs to communicate, even though texting is still challenging for them. An illustration of a word prediction problem is this one. The next word that is related to the current word is suggested by the word prediction model. Next word prediction systems are designed to reduce the number of keystrokes and spelling errors, save time, and enter a whole word with a single keystroke. In this field, natural language processing, or NLP, has gained popularity as a means of improving text comprehension. One of the earliest NLP techniques was the N-gram model, in which each gram stands for a word and is trained to extract language information and produce a set of features [1].

The field of word prediction is faced with an extremely challenging task: it must solve several common complex problems that arise when working with natural language. A computer cannot effectively handle the inherent quantity of

*Corresponding author

Email addresses:

Communicated by 'sub etitor'

arising ambiguities, which include pragmatic, cultural, phonetic, and structural ambiguities for speech in addition to lexical, structural, and semantic ambiguities. Numerous studies have been conducted, and a number of fundamental NLP tasks have been used, including language modeling, parsing, lemmatization, and part-of-speech (POS) tagging[2].

The area of machine learning that creates natural language from data input is called natural language generation. Data translation into natural language is done with it. NLG has many uses, including text production, summarization, picture auto-captioning, machine translation, etc. Using deep learning techniques on a given text, automatic text generation creates new text. Predicting the word or words that will come after the provided word sequence is known as "next word prediction," and it involves a high degree of probability [3].

Over the past ten years, next word prediction has become a hot topic in natural language processing (NLP). Previously, next word prediction was done using Markov models or Support Vector Machines. As technology advances, NLP models are transformed into deep learning algorithms [4].

Our lives with text input now revolve around a personal predictive text input system, or T9. For individuals with physical limitations, word prediction is frequently suggested by therapists as a way to increase typing speed. Increased speed could not be mentioned as a benefit of word suggestion, even though papers assert that it influences writing skills, when using a standard keyboard. In order to verify predicted words, users should turn away from the keyboard. This could be a factor in word prediction failure as it could slow down the user more than the word suggestion's acceleration. This effect is unrelated in situations where the typist must divert their attention from the document. Finding out if word prediction software would increase typing speed when a user should turn away from the document was the aim of the study. According to the experiment, word prediction increased typing speed for seven out of ten participants. Thus, the outcome suggests that word prediction may be beneficial to users[5].

In this paper, we have divided the content into several main sections, Section 2, Related Work, reviews the existing literature, discussing related research and explaining how this study builds upon and contributes to previous work. Section 3, System Architecture, focuses on the proposed system's structure, providing details about the design and methodology used to construct the studied system or model. Section 4, Results and Discussion, presents the findings obtained from experiments, followed by an in-depth analysis and discussion of their significance and implications. Finally, Section 5, Conclusion, summarizes the key takeaways from the research and suggests future directions for further development.

2. Related Work

Sharma, and et.,(2019) [3] emphasizes the creation of language that is natural and understandable to humans. This paper suggests a novel approach for predicting the following word in a Hindi sentence. It is possible to minimize the amount of keystrokes a user makes by anticipating the next word in a sequence. For the task of predicting the next word, two deep learning techniques—Long Short Term Memory (LSTM) and Bi-LSTM—have been investigated. For LSTM and Bi-LSTM, respectively, accuracy of 59.46% and 81.07% was noted. This method can be applied to a number of NLG tasks, such as sentence and story auto-completion. Ganai, & Khursheed, (2019) [6] employed a tree-based generative language model for document and part ranking. Nodes in the tree represent various sections, paragraphs, and titles in a document. Each document tree node has a clearly defined language model at it. A leaf node's language model can be inferred directly from the text in the node's document portion. The inner nodes of the tree can be predicted by using a linear interpolation between the different juvenile nodes. The paper also explains how this model would satisfactorily describe a few popular structural queries. Atçili, and et.(2021) [4] To select the optimal algorithms for the next word prediction, Turkish dataset was used. The open sources will be used to create the corpus. Only articles and comments about football, basketball, volleyball, and tennis will be included in the corpus. The study was carried out as much as possible in a restricted area due to time, dataset, and hardware constraints. Based on the outcomes to be obtained from this, it is thought to be advantageous for choosing the most successful model on broader subjects and even on a dataset with no constraints at all. Shakhovska, and et. (2021)[5] proposed, LSTM, Markov chains, and their hybrids were selected for next-word prediction for the Ukrainian language,. Their sequential structure, which relies on past output to produce current output, aids in effectively completing the next-word prediction task. The Markov chains gave the most rapid and satisfactory results. The hybrid model operates slowly but produces satisfactory results. Unlike T9, the user can generate multiple words or sentences using the model instead of just one. Tiwari, and et. (2022) [1] suggested to Utilize deep learning methods to anticipate the upcoming Hindi word. builds upon the fundamental neural network architecture of Bidirectional Long Short Term Memory and Long Short Term Memory. The model created in this work outperformed earlier techniques and had the highest accuracy among neural network-based methods using the IITB English-Hindi parallel corpus. Agarwal, and et.. (2022) [7] the BERT (Bidirectional Encoder Representations from Transformers) and ML (Masked Language) models are used to predict the next word based on the previous words for the Hindi language.

Also, for comparison of different algorithms and techniques on word prediction in related work with we work, as show in table (1):

Table 1: Comparison of different algorithms and techniques on Word Prediction

Author	Algorithm	Data set Language	Different
Sharma, and et.,(2019)[3]	Use Long Short Term Memory (LSTM) & Bi-LSTM	Hindi data set	Prediction for next word
Ganai, & Khursheed, (2019) [6]	Use a tree-based generative language model for ranking documents and parts	General	Prediction for next word
Atçili, and et.(2021)[4]	Use Recurrent Neural Networks (RNN) and Long Short-Term Memory Networks (LSTM).	Turkish dataset	Prediction for next word
Shakhovska, and et. (2021) [5]	Use LSTM and Markov chains	Ukrainian Data set	Prediction for next word
Tiwari, and et. (2022)[1]	Uses “Long Short Term Memory” and “Bidirectional Long Short Term Memory” as the base neural network architecture	Hindi data set	Prediction for next word
Agarwal, and et.. (2022) [7]	BERT (Bidirectional Encoder Representations from Transformers) Model and ML	Hindi data set	Prediction for next word

By reviewing previous works, we find that our scientific paper differs from theirs. We predict the next and previous words using the BERT algorithm and CBOW algorithm and applied it to a database to predict words, which is considered one of the most important applications of natural language processing.

3. System Architecture

Natural language processing (NLP) relies heavily on word and query prediction and Time series, which is crucial for many everyday applications. By offering precise recommendations that speed up and simplify the process of looking for information on search engines and social media platforms, this prediction helps to improve the user experience. Furthermore, word prediction shortens the time needed for text entry, increasing productivity and efficiency in auto-writing applications. Query prediction in chatbot systems helps customer service representatives provide accurate and timely answers, increasing customer satisfaction and boosting the efficiency of human-machine interaction. Word and query prediction is a crucial component of enhancing user experience and increasing productivity in information access and retrieval because, it generally improves efficiency and accuracy of search processes[8]. In this paper, we will apply word prediction on Arabic data set. In figure(1), that display steps work of the proposed method , as following :

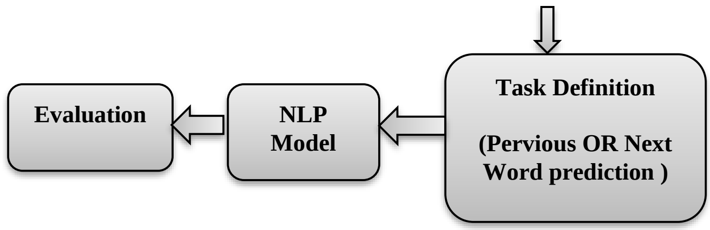

Fig 1. Generalized structure of Arabic Word Prediction

3.1 Data Collection and Preparation:

In this research, the bert 'asafaya/bert-base-arabic' model was used specifically trained to understand Arabic. This model relies on a large dataset of Arabic texts, including the Arabic Wikipedia and other text sources, to enable it to understand the context and semantics of words in Arabic more accurately.

As for the CBOW algorithm, it is trained on a database given by the user, in our research, the algorithm was trained on a special database that was collected manually, and it can be trained on any other database.

3.2 Context Selection:

To predict the next word, ascertain the context. The context could consist of a series of words or an entire sentence.

3.3 Task Definition:

Make sure the model is given the correct task definition when it comes to word prediction. This entails giving the model's input—the sentence and its preceding words . There are two types of pre-training: Next Sentence Prediction (NSP) and Masked LM. The first one involves using the "[MASK]" token to mask a predetermined percentage of the input tokens, and then predicting the tokens that have been masked. In the second, given two sentences A and B, 50% of the time B is labeled as Is Next—the actual sentence that comes after A—and 50% of the time B is a randomly selected sentence from the corpus—labeled as Not Next .

3.4 NLP Model

3.4.1 Bert Arabic Prediction Model

Advanced models for natural language processing include the Bert model. The framework of the BERT model consists of two phases: pre-training and fine-tuning. Pre-training involves masking a certain percentage of the input tokens at random (using the “[MASK]” token) and then predicting those masked tokens. The model is trained on an unlabeled large corpus.

Fine-tuning is straightforward: the model is started with the pre-trained parameters, and all the parameters are adjusted using labeled data for particular tasks. This is made possible by the Transformer's self-attention mechanism, which allows BERT to model numerous downstream tasks. All we have to do is enter the exact inputs and outputs for every task into BERT and tweak every setting. The original implementation, described in detail in, served as the foundation for the multi-layer bidirectional Transformer encoder model architecture used in BERT. This kind of encoder is made up of a stack of N = 6 identical layers. Every one of these layers has two sub-layers. A multi-head self attention mechanism and a position-wise fully connected feedforward network are the two types of networks. It uses a residual connection that surrounds every sublayer, then followed by a layer normalization [9]. In Figure (2) and Figure (3), the mechanism of the BERT natural language processing model is illustrated.

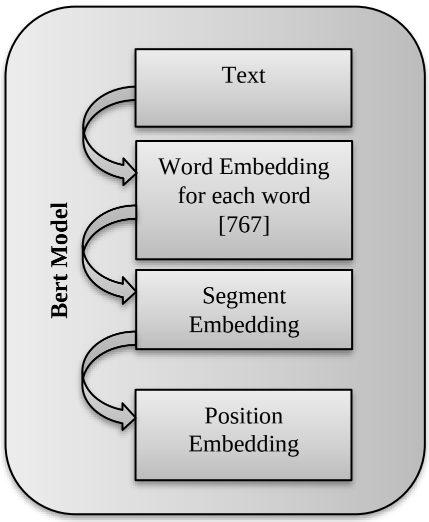

Fig 2. Bert Model

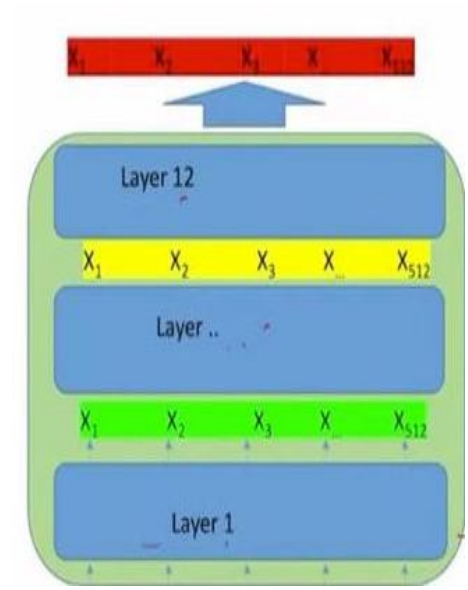

Fig 3. Bert Architecture

Each sub-layer's output is represented as LayerNorm(x + Sublayer(x)), where Sublayer(x) is the sub-layer's implementation of the function. First, we must define scaled dotproduct attention in relation to multi-head self attention. It has the following definition:

$$\text{Attention}(Q, K, V) = \text{sofmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \ldots\ldots (1)$$

where dk is the dimension of the Q and K matrices, Q is the matrix of queries, K is the matrix of keys, V is the matrix of values. Currently, multi-head attention can be described as:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h)W^o \ldots\ldots (2)$$

Where, headi = focus(KWKi, VWVi, QWQi) in this instance The multi-head attention process involves projecting the queries, keys, and values h times using different, learned linear projections to dk, dk, and dv (the dimensions of the values matrix), respectively. We then run the attention function on each of these projected queries, keys, and values in parallel, producing dv dimensional output values. These are concatenated and projected to yield the final values. All of the keys, values, and queries originate from the same location when selfattention is present. BERT uses a sequence of tokens based on the following characteristics to represent a single sentence or a pair of sentences (e.g., the pair hquestion, answeri): By using WordPiece embedding, BERT. The first token in the sequence is "[CLS]". When there are two sentences in a sequence, they are divided using the "[SEP]" token. Furthermore, an embedding indicating whether a token belongs in the first or second sentence is added to each token. The segment, position, and token embedding for a given token are added to create its input representation[10,11] .

3.4.2 Continuous Bag of Words(CBOW) Arabic Prediction Model:

There are more practical uses for the word2vec model in various NLP tasks. Machine learning text classification has found useful applications for the semantic meaning provided by word2vec for each word in vector representations[12,13]. They are used to analyze words semantically, syntactically, and by analogy. Word2vec is available in two varieties: Skip-Gram and CBOW. They used to predict a word given a context, and the opposite is also true. They have added two computational techniques—hierarchical softmax and negative sampling to maximize word2vec's efficiency[14,15].

CBOW Model , It can be described as a neural network, but its purpose is to predict the word that is missing from a given sentence—typically the last word—by calculating the embedding matrices of the words that are entered. These matrices are then processed to move on to the next layer of the network, where the softmax activation function is the last layer. to select the word that is lacking. A single hot encoded vector of size V is the input, also known as the context word. The output of the hidden layer, which has N neurons in it, is once more a vector of vector length with softmax values as its elements. The words in the image below, W\nv is the weight matrix that maps the outputs of the hidden layer to the final output layer (N*V dimensional matrix), whereas Wvn is the weight matrix that maps the input x to the hidden layer (V*N dimensional matrix). The neurons in the hidden layer simply replicate the input's weighted sum to the subsequent layer. The output layer's softmax computations account for the non-linearity of the activation function. However, the model mentioned above predicted the target using just one context word. To achieve the same goal, we can use several context words[16].

In Figure (4) and Figure (5), the mechanism of the CBOW natural language processing model is illustrated.

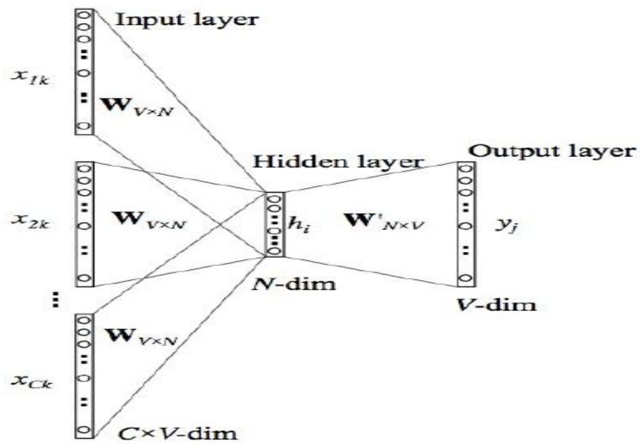

Fig 4. CBOW Mechanism

We calculate their embedding values, find the average of each word, then we rotate this matrix and multiply it by the training values h and we can infer the next word using softmax. As in equation following:

$$h = \frac{1}{|C|}\sum_{c\in C} W^{(1)}\, c \quad \ldots\ldots (3)$$

$$p(y|C) = \mathrm{softmax}(W^{(2)T}\, h) \quad \ldots\ldots (4)$$

In summary, the neural network learns to represent words in a way that preserves semantic meaning, and CBOW focuses on predicting a target word based on its surrounding context. The process of training entails modifying the weights of the network to enhance its precision in predicting target words.

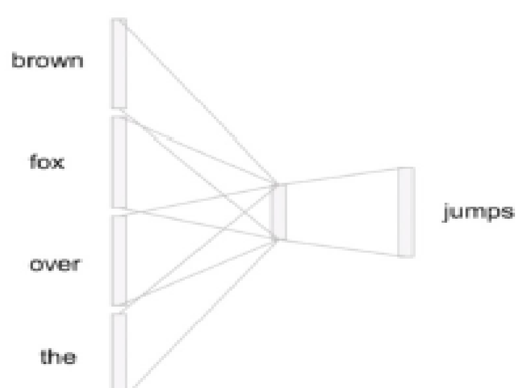

Fig 5. CBOW prediction Word

4. The Results and Discussion :

The results of our experiments showed promise for both previous and next position. When it came to Arabic word prediction, BERT proved resilient and had a high accuracy rate. BERT's ability to predict Arabic words is demonstrated by its success, which highlights the complexity of linguistic structures it can handle. Still, dealing with particular dialects and semantic subtleties presents difficulties. This work provides a basis for future developments in NLP for Arabic language processing by demonstrating the efficacy of BERT in Arabic word prediction.

The word prediction results of the BERT and CBOW algorithms are as follows:

Table 2. Results of Prediction for Next Word Using Bert Algorithm

The Query	The Query After Replace Word with Mask	Predicted Index	Prediction for Next Word
أني مريض جدا احتاج الى الذهاب الى المستشفى	أني مريض جدا احتاج الى الذهاب الى [mask]	8724	المستشفى
الهاتف المحمول أحد تطبيقات الذكاء الاصطناعي	الهاتف المحمول أحد تطبيقات الذكاء [mask]	19441	الاصطناعي
لغة بايثون هي لغة برمجة	لغة بايثون هي لغة[mask]	13966	برمجة
الحاسوب مهم ويستخدم كثيرا في الحياة اليومية	الحاسوب مهم ويستخدم كثيرا في الحياة [mask]	7318	اليومية
السيارة وسيلة نقل من مكان الى مكان اخر	السيارة وسيلة نقل من مكان الى مكان [mask]	1947	اخر

Table 3. Results of Prediction for Pervious Word Using Bert Algorithm

The Query	The Query After Replace Word with Mask	Predicted index	Prediction for Pervious Word
أني مريض جدا احتاج الى الذهاب الى المستشفى	أني مريض جدا احتاج الى الذهاب [mask] المستشفى	1774	الى
الهاتف المحمول أحد تطبيقات الذكاء الاصطناعي	الهاتف المحمول أحد تطبيقات [mask] الاصطناعي	14023	الذكاء
لغة بايثون هي لغة برمجة	لغة بايثون هي [mask] برمجة	5042	لغة
الحاسوب مهم ويستخدم كثيرا في الحياة اليومية	الحاسوب مهم ويستخدم كثيرا في [mask]اليومية	12430	الحياة
السيارة وسيلة نقل من مكان الى مكان اخر	السيارة وسيلة نقل من مكان الى [mask] اخر	3108	مكان

Table 4. Results of Word Prediction Using CBOW Model

The Query	Prediction for Next Word
لغة بايثون من اللغات[mask]	المهمة
بايثون من اللغات المهمة[mask]	التي
من اللغات المهمة التي[mask]	تستخدم
اللغات المهمة التي تستخدم[mask]	في
المهمة التي تستخدم في[mask]	تطبيقات

Table 5. Results of Word Prediction Using CBOW Model

The Query	Prediction for Pervious Word
لغة بايثون [mask] اللغات المهمة	من
بايثون من[mask] المهمة التي	اللغات
من اللغات[mask] التي تستخدم	المهمة
اللغات المهمة [mask] تستخدم في	التي
المهمة التي [mask] في تطبيقات	تستخدم

The CBOW algorithm is also used to predict more than one next word, as in the following table(7) showing that the CBOW algorithm predicted three words and more, after being trained on a given database.

Table 6. Results of Words Prediction Using CBOW Model

The Query	Number of Word	Prediction for Next Word
الذكاء الاصطناعي هو مجال يدرس[mask][mask]	2	الذكاء الاصطناعي هو مجال يدرس كيف يمكن
الذكاء الاصطناعي هو مجال يدرس[mask][mask][mask]	3	الذكاء الاصطناعي هو مجال يدرس كيف يمكن للأنظمة
الذكاء الاصطناعي هو مجال يدرس[mask][mask][mask][mask]	4	الذكاء الاصطناعي هو مجال يدرس كيف يمكن للأنظمة الكمبيوترية
الذكاء الاصطناعي هو مجال يدرس[mask][mask][mask][mask][mask]	5	الذكاء الاصطناعي هو مجال يدرس كيف يمكن للأنظمة الكمبيوترية تنفيذ

We conclude from the results above that the BERT algorithm predicts the next and previous words without being trained on a given database because it was trained during its structure, while the algorithm of Continuous Bag of Words (CBOW) predicts the next and previous words but after it was trained on a given database.

To measure the accuracy of the natural language processing models for prediction, we use the following metrics:

precision: What is truly pertinent to the question is the percentage of recovered documents. The total number of pertinent recovered documents divided by the total number of recovered documents is the definition of accuracy. The true positive examples include accuracy, which can be expressed as the ratio of true positives to the sum of false positives and true positives[17,18]. Accuracy is calculated as follows:

$$\text{Precision} = \frac{\text{correctly identify}}{\text{All tested set words}} \quad \dots \dots (5)$$

$$\text{percision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \dots \dots (6)$$

Recall: the proportion of query-related documents that were truly retrieved. The total number of pertinent documents that were retrieved out of all the related documents in the database is the recall. Recall is the ratio of true positives to the sum of true positives and false negatives. It is also referred to as sensitivity[19]. The recall is calculated as follows:

$$\text{Recall} = \frac{\text{correctly identify}}{\text{All tested set words}} \quad \dots \dots (7)$$

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \dots \dots (8)$$

-F- Measure: The F-measure, which combines the model's precision and recall, is a measure of the model's accuracy on a dataset by classifying examples as "positive" or "negative." It is defined as the harmonic mean of the model's precision and recall [20]. The F- measure is calculated as follows:

$$\text{F\_measure} = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad \dots \dots (9)$$

Table 7. Accuracy Results

Algorithm	Next Word Prediction			Previous Word Prediction		
	Precision	**Recall**	**F-measure**	**Precision**	**Recall**	**F-measure**
BERT	0.90	0.90	0.90	0.80	0.80	0.80
CBOW	1.0	1.0	1.0	1.0	1.0	1.0

5. CONCLUSION

Users benefit from word prediction because it can speed up typing and help to avoid mistakes. Research on a personalized, predictive text input system is pertinent for all languages, but Arabic in particular. In this paper, we used the BERT algorithm and an CBOW algorithm where the BERT algorithm is an algorithm trained on a database in its structure and predicts words without training on a given database, while the algorithm CBOW does not make predictions unless it is trained on a given database. The introduction of a new pre-trained BERT text attachment model, which enables unprecedented precision results in many automated word processing tasks, was the main event in the field of natural language processing in 2019. In terms of prevalence, this model is probably going to overtake the well-known word2vec model and end up becoming the industry standard. In one way or another, the majority of scientific publications published in 2019 that addressed the issue of word processing in natural languages were a response to the introduction of this new model, whose creators are now among of the most cited researchers in the field of machine learning. In this paper, we focused on predicting the next word and the previous word of the sentence using the BERT algorithm and it was applied to an Arabic language database.

ACKNOWLEDGEMENT

The lab and required documents were provided by the Department of Computer Science, Faculty of Education for Girls, University of Kufa, Najaf, Iraq, for which the authors are grateful.

REFERENCE

[1]. R.M, Duwairi, Marji, N., Sha'ban & S. Rushaidat (2014, April). Sentiment analysis in arabic tweets. In 2014 5th international conference on information and communication systems (ICICS) (pp.1-6). IEEE. DOI: 10.1109/GlobConPT57482.2022.9938153

[2]. C. ,Aliprandi, N. Carmignani, N. Deha, P Mancarella, & M. Rubino (2008). Advances in nlp applied to word prediction. University of Pisa, Italy February.

[3]. R. Sharma, N. Goel, N. Aggarwal, P. Kaur & C. Prakash (2019, September). Next word prediction in hindi using deep learning techniques. In 2019 International conference on data science and engineering (ICDSE) (pp. 55-60). IEEE. DOI: 10.1109/ICDSE47409.2019.8971796

[4]. A. Atçili, O. Özkaraca, G. Sariman, & B. Patrut, (2021, October). Next Word Prediction with Deep Learning Models. In The International Conference on Artificial Intelligence and Applied Mathematics in Engineering (pp. 523-531). Cham: Springer International Publishing.

[5]. K. Shakhovska, I. Dumyn, N. Kryvinska, & M.K. Kagita, (2021). An Approach for a Next-Word Prediction for Ukrainian Language. Wireless Communications and Mobile Computing, 2021, 1-9.DOI.org/10.1155/2021/5886119

[6]. A.F. Ganai, & F. Khursheed, (2019, November). Predicting next word using RNN and LSTM cells: Stastical language modeling. In 2019 Fifth International Conference on Image Information Processing (ICIIP) (pp. 469-474). IEEE. DOI: 10.1109/ICIIP47207.2019.8985885

[7]. S. Agarwal, A. Sukritin, Sharma, & A. Mishra, (2022). Next Word Prediction Using Hindi Language. In Ambient Communications and Computer Systems: Proceedings of RACCCS 2021 (pp. 99-108). Singapore: Springer Nature Singapore.

[8]. U.Anil, &M. Akcayol, M. (2020). Deep learning based prediction model for the next purchase. Advances in Electrical and Computer Engineering, 20(2).

[9]. S. González-Carvajal, & E.C. Garrido-Merchán, (2020). Comparing BERT against traditional machine learning text classification. arXiv preprint arXiv:2005.13012.

[10]. M.V. Koroteev, (2021). BERT: a review of applications in natural language processing and understanding. arXiv preprint arXiv:2103.11943.

[11]. S., Chakraborty, M. Borhan Uddin Talukdar, P., Sikdar, & J., Uddin, (2024). An Efficient Sentiment Analysis Model for Crime Articles' Comments using a Fine-tuned BERT Deep Architecture and Pre-Processing Techniques. Journal of Information Systems and Telecommunication (JIST), 1(45), 1.

[12]. S., Sina, R., Ramin (2024). An Aspect-Level Sentiment Analysis Based on LDA Topic Modeling. Journal of Information Systems and Telecommunication (JIST), (46), 117.

[13]. M. Milicevic, M. Baranovic, & K. Zubrinic, (2015). Application of machine learning algorithms for the query performance prediction. Advances in Electrical and Computer Engineering, 15(3), 33-44.

[14]. Karani, D. (2018). Introduction to word embedding and word2vec. Towards Data Science, 1.

[15]. M., Jaderyan, & H., Khotanlou(2020). SGF (Semantic Graphs Fusion): A Knowledge-based Representation of Textual Resources for Text Mining Applications. Journal of Information Systems and Telecommunication (JIST), 2(26), 120.

[16]. S. Sivakumar, L.S. Videla,T.R. Kumar, J. Nagaraj, S. Itnal, & D. Haritha, (2020, September). Review on word2vec word embedding neural net. In 2020 international conference on smart electronics and communication (ICOSEC) (pp. 282-290). IEEE. DOI: 10.1109/ICOSEC49089.2020.9215319

[17]. R. Wang, & J. Li, (2019, July). Bayes test of precision, recall, and F1 measure for comparison of two natural language processing models. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 4135-4145).

[18]. D. Jatnika, M.A. Bijaksana, & A.A.Suryani, (2019). Word2vec model analysis for semantic similarities in english words. Procedia Computer Science, 157, 160-167.

[19]. B. Sidaoui, &K. Sadouni, (2023). Epilepsy Seizure Prediction from EEG Signal Using Machine Learning Techniques. Advances in Electrical & Computer Engineering, 23(2).D. Karani, (2018). Introduction to word embedding and word2vec. Towards Data Science, 1.

[20]. Taher, H. A., Abdulameer, M. H., & Mahdi, B. (2022). Information Retrieval Scheme Via Similarity Technique. International Journal on Technical and Physical Problems of Engineering (IJTPE), (51), 375-379.