IJICTR
International Journal of Information & Communication Technology Research

Volume 16- Number 1 – 2024 (42 -54)

An Optimized Deep Model for Bi-lingual Sentiment Analysis of Medical User Reviews

Shahla Nemati*
Department of Computer Engineering
Shahrekord University
Shahrekord, Iran
s.nemati@sku.ac.ir

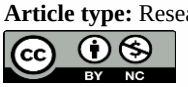
Reza Salehi Chegani
Department of Computer Engineering
Shahrekord University
Shahrekord, Iran
reza.sch@yahoo.com

Received: 26 April 2023 – Revised: 6 June 2023 - Accepted: 16 July 2023

Abstract—**Sentiment analysis of online doctor reviews helps patients to better evaluate and select the related doctors based on the previous patients' satisfaction. Although some studies are addressing this problem in the English language, only one preliminary study has been reported for the Persian language. In this study, we propose a new evolutionary deep model for sentiment analysis of Persian online doctor reviews. The proposed method utilizes both Persian reviews and their English translations as inputs of two separate deep models. Then, the outputs of the two models are combined in a single vector which is used for deciding the sentiment polarity of the review in the last layer of the proposed deep model. To improve the performance of the system, we propose an evolutionary approach to optimize the hyper-parameters of the proposed deep model. We also compared three evolutionary algorithms, namely, Ant Colony Optimization (ACO), Genetic Algorithm (GA), and Gray Wolf Optimization (GWO) algorithm, for this purpose. We evaluated the proposed model in two phases; In the first phase, we compared four deep models, namely, long short-term memory (LSTM), convolutional neural network (CNN), a hybrid of LSTM and CNN, and a bidirectional LSTM (BiLSTM) model with four traditional machine learning models including Naïve Bayes (NB), decision tree (DT), support vector machines (SVM), and random forest (RF). The results showed that the BiLSTM and CNN models outperform other methods, significantly. In the second phase, we compared the optimized version of two proposed bi-lingual models in which either two BiLSTM or two CNN models were used in parallel for processing Persian and English reviews. The results indicated that the optimization of the CNN using ACO and the optimization of BiLSTM using a genetic algorithm can achieve the best performance among other combinations of two deep models and three optimization algorithms. In the current study, we proposed two deep models for bi-lingual sentiment analysis of online Persian doctor reviews. Moreover, we optimized the proposed models using ACO, genetic algorithm, and gray wolf optimization methods. The results indicated that the proposed bi-lingual model outperforms a similar model using only Persian or English reviews. Also, optimizing the parameter of proposed deep models using ACO or genetic algorithms improved the performance of the models.**

Keywords: Medical Reviews, Online Doctor Reviews, Persian Sentiment Analysis, Bi-lingual Sentiment Analysis, Deep Learning, Evolutionary Optimization.

Article type: Research Article



Publisher: ICT Research Institute

* Corresponding Author

[DOI: 10.61186/itrc.16.1.42]

I. INTRODUCTION

The medical domain has attracted the attention of researchers in the natural language processing and data mining community [1]. In addition to the existing English websites, several online platforms have been provided for Persian-speaking users to interact with various medical fields. With the advent of social media, patients have the opportunity to express their opinions about doctors online. These platforms provide information such as the characteristics of specialists close to the user, the type of insurance they accept, the distance to the user, the percentage of patients who recommend the doctor, as well as patient opinions [2]. Extracting and analyzing the sentiment provided in user reviews can help other patients who are interested in finding out how much a physician cares about the treatment process of their patients. In addition, physician referral systems may use such information to assist patients in selecting the appropriate physician to treat their disease [3].

Due to its special medical features, opinions expressed in medical reviews have become an attractive source of information for sentiment analysis research [2]. However, textual information extracted from medical-related websites is more complex than in other fields because of the special medical terms and abbreviations used. Another problem in dealing with medical reviews is the indirect expression of sentiment [4]. In addition, the high lexical diversity and special terms make the analysis of medical sentiment more challenging [5].

Generally, there are two types of medical reviews on social media, including drug reviews and online doctor reviews. Drug reviews are written texts that describe various aspects of drug use such as side effects, price, and effectiveness [2]. Online doctor reviews are usually written by patients who want to express their opinion about the doctor or the treatment process they have experienced. These reviews may be used as the first step in finding a new doctor for patients. There are some similarities between drug reviews and online medical reviews. For example, both are typically written by non-professional users and also describe the user experience with the environment and the medical field. However, experience has shown that online medical reviews have fewer technical terms and are shorter [6].

Some studies addressed the problem of sentiment analysis for drug reviews and online doctor reviews written in English and other resource-rich languages in the literature [5], [7]. However, for the Persian language which is the 8th language of the Web concerning the content used in the Web and has the 6th market position in terms of popularity and traffic compared to the most popular content languages [8], there are few studies addressed the problem of medical review sentiment analysis. This motivated us in the current study to investigate the problem of the Persian language.

Because among Persian websites, there are more platforms for users to express their feelings and sentiment about doctors than drugs, in this study, we analyze online medical doctor reviews written in Persian. The results of this analysis may be used in doctor recommender systems as well as doctor ranking and retrieval systems. However, sentiment analysis in the Persian language has several problems such as a lack of resources, structural complexities, and informal language used in Persian social media [9]. These problems may be addressed using either traditional machine learning and natural language processing techniques, or deep neural models [10]. In the current study, due to their power of solving different NLP problems, we preferred to address these problems by proposing a deep bi-lingual model for sentiment analysis of online doctor reviews optimized by evolutionary algorithms.

To the best of our knowledge, only one study has analyzed the sentiment of Persian online doctor reviews [11]. They introduced the first online doctor review dataset for the Persian language, PODOR, and proposed a deep model for sentiment analysis of its reviews. PODOR contains 700 reviews written in Persian. To improve this study, we performed the following tasks:

- We added 300 more Persian reviews to the PODOR.
- We added the English translation of all 1000 reviews to the dataset to make the bi-lingual analysis of the reviews possible.
- We proposed two bi-lingual deep models in which both the English and Persian texts of reviews were used as the input of the system.
- We proposed an evolutionary method for optimizing the hyper-parameters of deep models and compare three evolutionary algorithms for this task.

II. RELATED WORK

Several studies have addressed the problem of sentiment analysis in medical domains. This section briefly reviews some of the most related studies in the literature. A more detailed review of medical sentiment analysis may be found in [12], [13].

In [6], a hybrid method called Sentihealth was presented to create a health-specific lexicon that used a bootstrapping-based method for effectively classifying and scoring health-related users' sentiments. They used N-gram features and compared their method with classical information retrieval methods including mutual information and information gain and showed that their method outperforms other methods in terms of both accuracy and F-measure.

In [14], the amount of patients' satisfaction was analyzed using feedback on physicians' performance. The authors proposed a method to predict the performance ranking in terms of knowledge and usefulness. To this aim, they tried out convolutional neural networks (CNN) with different loss functions

[DOI: 10.61186/itrc.16.1.42]

and optimization algorithms. They analyzed about 35000 reviews of more than 1000 physicians and reported an accuracy of 93% for binary classification of reviews.

In [3], a new healthcare recommender system called iDoctor was developed. This method was based on matrix factorization and used Sentiwordnet 3.0 [3], [15]. Also, it utilizes a topic model-based method to find the distributions of user preferences and doctor features. In the sentiment analysis part of their method, the authors proposed an emotion-aware method for specifying emotional offset [3].

In [16], researchers developed a new lexicon based on various areas, including diseases, symptoms, drugs, human anatomy, and medical terms. They applied various machine learning methods to the extracted lexicon. Specifically, they proposed a distributed framework in which long short-term memory (LSTM) neural network was utilized to predict the moods of cancer-affected patients from their tweets. They constructed a dataset and evaluated their method on this and two existing datasets.

In [17], [18], a deep learning approach was proposed to categorize the emotion of patients toward the hospital, before, during, and after leaving it. In this study, the authors used the SenticNet framework [18] emotion category and labeled their dataset using SenticNet. They compared different deep models with three traditional machine learning algorithms including support vector machines (SVM), logistic regression, and multinomial naïve Bayes, and showed that the combination of the gated recurrent unit (GRU) and CNN achieves the best performance among the others.

In [19], the authors proposed a text mining approach based on LDA to process online doctor reviews. Specifically, they tried to find different factors of patient satisfaction toward the United Kingdom's healthcare services. They focused on what patients like and what they dislike during their treatment process. The results showed that hospital-related factors including location, environment, and parking as well as doctor-related factors including knowledge and attitude are the most determinants of patient satisfaction while their treatment experience and staff bedside manners are the most concerning factors for the patients [19].

In [5], two different datasets were extracted from the social web to determine the best way to analyze the sentiment of Spanish users' comments in the medical domain. They applied supervised learning methods such as support vector machine (SVM) and compared the results with a lexicon-based method. In this study, both drug reviews and doctor reviews were analyzed. The results showed that drug reviews are more difficult to classify than doctor reviews and this is due to the linguistic differences between the two types of reviews.

In [20], a text mining method based on LDA is used to extract the topics of collected online doctor reviews from a Chinese online health platform. In this study, more than 500000 reviews of about 75000 Chinese doctors were analyzed. The results showed that the most frequent topics were the experience of finding doctors and doctors' technical skills and bedside manner.

In [21], the authors presented the first Russian-language textual corpus of drug reviews. They proposed an XLM-RoBERTa model for extracting entity relations and compared their method with four machine learning models including Logistic regression, SVM, multinomial naïve Bayes, and gradient boosting. The results showed that a state-of-the-art accuracy level for the task of pharmacological entity relation identification in the Russian language can be achieved using the proposed transformer-based model.

Finally, in [11], the authors introduced the first Persian dataset of online doctor reviews and proposed a deep learning model for sentiment analysis of its reviews. They showed that their deep model outperforms traditional machine learning methods. The main restrictions of this study are the use of only Persian reviews and their proposed deep model which has not been optimized. In the current study, we will address this research gap by adding the English translation of the reviews and by utilizing evolutionary algorithms for optimizing the hyper-parameters of the models. An overview of the above-mentioned studies is shown in Table 1.

III. CONTRIBUTION AND NOVELTY

As we pointed out in the introduction, only one study addressed the problem of sentiment analysis of Persian online doctor reviews [11]. In their study, the authors introduced the first online doctor review dataset for the Persian language, PODOR, containing only Persian reviews. They also proposed a deep model for sentiment analysis of the reviews. The differences between our current study and that mentioned study [11] are as follows.

First, we augment the PODOR dataset by adding 300 more Persian reviews which made the total number of reviews 1000. Then, we translated all 1000 reviews into English and then to the dataset which makes the bi-lingual study of online doctor reviews possible. Furthermore, we proposed bi-lingual deep fusion models for sentiment analysis of this dataset. Finally, we proposed a meta-heuristic optimization approach for optimizing the hyper-parameters of the proposed deep models. This step, to the best of our knowledge, is conducted for the first time in deep sentiment analysis of the Persian language. In this step, we compared two swarm intelligence methods namely ant colony optimization (ACO) and gray wolf optimization (GWO) methods with genetic algorithms (GA).

In the domain of bilingual Persian sentiment analysis, there are several studies in the literature. For

[DOI: 10.61186/itrc.16.1.42]

example, in [27], a transfer learning method based on bilingual embedding was proposed to compute the sentiment polarity of two English and Persian customer review datasets. They compared their model with CNN and LSTM and showed that their model achieves a better performance provided that a cross-lingual embedding is available.

TABLE I. AN OVERVIEW OF RELATED STUDIES FOR MEDICAL REVIEW SENTIMENT ANALYSIS.

Ref.	Year	Language	Method
[6]	2016	English	Lexicon-based
[22]	2016	English	Deep learning
[23]	2016	Chinese	ML
[3]	2017	English	ML
[16]	2019	English	Lexicon-based
[5]	2019	Spanish	ML
[11]	2020	Persian	Deep
[24]	2021	English	ML
[25]	2022	English	Deep
[26]	2022	Russian	Deep

In [28], a similar cross-lingual model was proposed for Persian sentiment analysis. They used the VecMap method to align English and Persian word embeddings in a joint space. They evaluated the method on a restaurant review dataset and showed that without using a Persian training dataset, their method achieved acceptable results.

There are two main differences between these two models and our proposed model. First, the focus of our method is on the optimization of the hyper-parameters of the proposed deep models using an evolutionary approach. Second, our method learns the word embeddings from the provided dataset while those two methods tried to find a joint space between English and Persian word embeddings. The reason for learning the word embeddings from the training dataset and not relying on the existing English word embedding is that our model is proposed for a special domain (i.e. medical reviews) and general word embeddings like those used in the two above-mentioned studies are not as accurate as a learned

IV. DATA AND PROCESSING

A. *Data Set*

In [11], a new dataset of Persian online doctor reviews, PODOR, was introduced for the first time. PODOR contains 700 manually gathered reviews from nobat.ir website. In the current study, we manually downloaded 300 more Persian reviews from the same website and added them to PODOR. Moreover, we added the English translations of all 1000 reviews to the dataset and named the new dataset, PODOR2. Each review in PODOR2 includes the unique ID of the doctor, the comment text in Persian, the comment text in English, the date of submission, and the comment label which is a binary number indicating the sentiment polarity of the review. Each review was manually labeled. The distribution of the positive and negative reviews in PODOR2 can be seen in Fig. 1.

Fig. 2 shows the distribution of review lengths in word count for the PODOR2 dataset. As shown in Fig. 2, the largest number of reviews belongs to the category of length 10 to 20 words with 329 reviews. The lowest number of reviews belongs to the category of 40 to 50 with 49 reviews. In addition, the average length of all reviews is 22.2 in terms of their number of words. A comparison between PODOR and PODOR2 datasets is shown in Table 2. It should be noted that, in the first version of the dataset, we used the rates provided on the website besides the reviews. Specifically, we considered reviews with a score of 5 as positive and those with a score of 2.5 or 0 as negative. In the second version of the dataset, we labeled the dataset manually to increase the accuracy of the labeling process. Therefore, the distribution of positive and negative reviews in the PODOR2 dataset has been changed in comparison to the PODOR dataset. Both versions of the dataset are available at https://github.com/mebasiri/ODR.

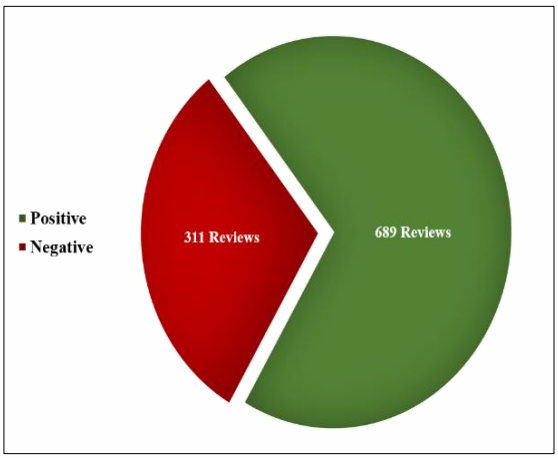


Figure 1. Distribution of positive and negative reviews in the PODOR2 dataset.

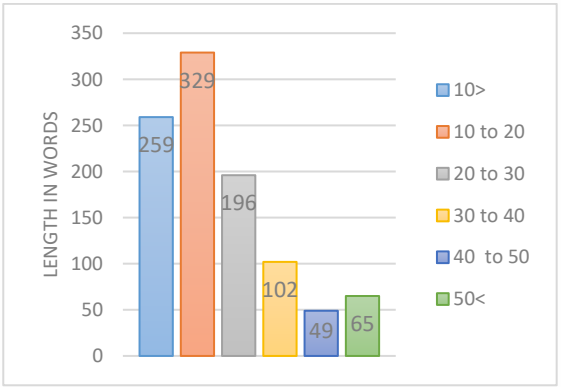


Figure 2. Distribution of comments in the PODOR2 dataset in terms of length by word count.

[DOI: 10.61186/itrc.16.1.42]

TABLE II. COMPARISON OF PODOR AND PODOR2 DATASETS

Parameter	PODOR	PODOR2	
		Per.	Eng.
Total number of reviews	700	1000	1000
Number of positive reviews	273	689	689
Number of negative reviews	427	311	311
avg word	21.67	22.22	24.25
avg sentence	2.3	1.62	2.56

V. DEEP BI-LINGUAL MODELS

Deep bilingual models, which are alternatively referred to as multilingual or multilingual neural machine translation (NMT) models, belong to the realm of artificial intelligence. These models possess the capability to translate text between two or more languages proficiently. Their purpose lies in acquiring a comprehensive understanding of the relationships and patterns within the data across multiple languages concurrently. A notable benefit of deep bilingual models is their aptitude for transferring knowledge across languages. Through the combined training on multiple language datasets, these models can acquire shared representations and enhance translation quality for various language combinations. A prominent characteristic of deep bilingual models involves the utilization of recurrent neural networks (RNNs) or transformer models. RNNs, including long short-term memory (LSTM) or Gated Recurrent Units (GRUs), have gained significant popularity in tasks involving sequences, such as machine translation. On the other hand, transformer models employ self-attention mechanisms that enable parallel processing and capture overall interdependencies [27].

In this study, to create bi-lingual models, the neural layers of each language are constructed independently of each other. Then, their output is merged and the final output is obtained. The performance of deep neural models is optimized using meta-heuristic algorithms. We divided our study into two phases; in the first phase, we evaluated four deep models, namely, long short-term memory (LSTM), three layered convolutional neural networks (3CNN), a hybrid of LSTM and 3CNN, and a bidirectional LSTM (BiLSTM). In the second phase, based on the results of the previous phase, we proposed two bi-lingual models. Specifically, the BiLSTM and 3CNN which outperformed LSTM and 3CNN-LSTM are selected to form the bi-lingual models as follows.

A. *Bi-LSTM-2 Model*

The first model we introduce is called Bi-LSTM-2, shown in Fig. 3.

This model has two inputs for Persian and English text, respectively. It should be noted that each of these inputs was preprocessed as we described in the data pre-processing section. The inputs are independently passed to word embedding layers. Then, the outputs of embedding layers are sent to the Bi-LSTM layers whose number of units is determined by meta-heuristic algorithms. In the BiLSTM layer, every LSTM unit consists of a memory cell c_t, which preserves its state over arbitrary time intervals, and three non-linear gates including an input gate i_t, a forget gate f_t, and an output gate o_t.

The output of the BiLSTM layer at time t, $h_{t_{BiLSTM}}$, is computed according to (1):

$$h_{t_{BiLSTM}} = [\overrightarrow{h_{t_{LSTM}}}, \overleftarrow{h_{t_{LSTM}}}] \quad (1)$$

Where $\overrightarrow{h_{t_{LSTM}}}$ and $\overleftarrow{h_{t_{LSTM}}}$ are forward and backward hidden layers used in the BiLSTM layer to capture the future and preceding context [24].

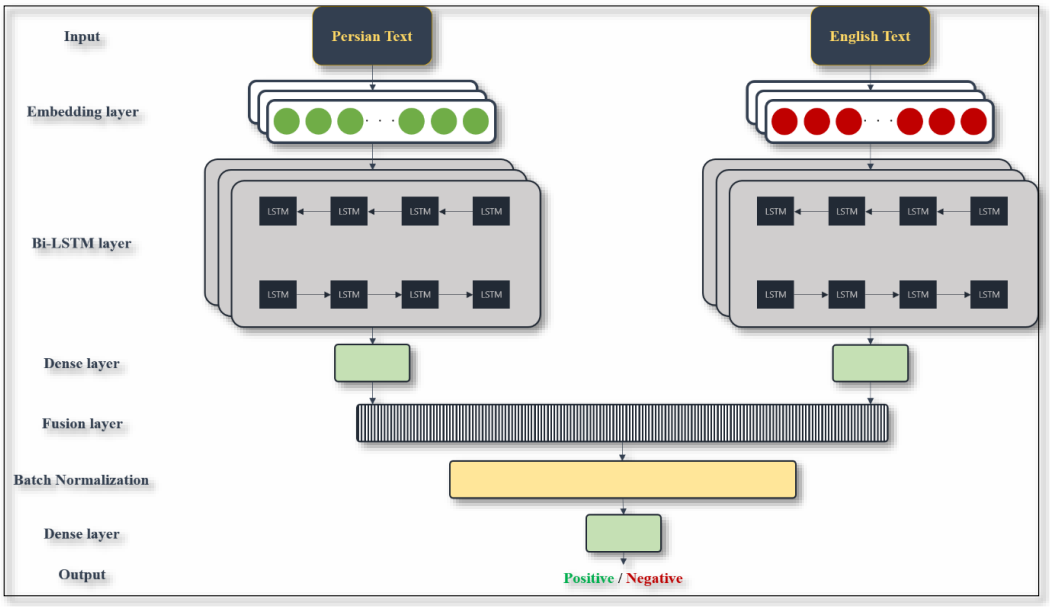


Figure 3. The proposed Bi-LSTM-2 model for sentiment analysis of Persian online doctor reviews.

[DOI: 10.61186/itrc.16.1.42]

Here, $h_{t_{LSTM}}$ which is called h_t hereafter, may be computed as:

$$h_t = o_t \odot \tanh c_t \quad (2)$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \quad (3)$$

where $\odot$, $\sigma(.)$, and tanh are the element-wise product, sigmoid function, and hyperbolic tangent function, respectively. U and W show the weight matrices of gates or cells for input x_t and b, denote the bias vectors. c_t is computed based on the forget and input gates as follows [29]:

$$f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right) \quad (4)$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (5)$$

$$\tilde{c}_t = \tanh(W_c h_{t-1} + U_c x_t + b_c) \quad (6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (7)$$

Before combining the Persian and English branches, a fully connected dense layer is placed for each of the languages. The outputs generated by both languages are then combined to form a common vector. To increase the speed and accuracy of the network in this model, the generated vector is passed through a batch normalization layer. Finally, to generate the output and determine the polarity, using a dense layer, a number in the range 0 and 1 is obtained that indicates the polarity of the input text as positive or negative; The closer to 1, the more the polarity tends to be positive, and vice versa for 0. In this model, activation functions are also selected by meta-heuristic algorithms.

B. 3CNN-2 Model

The Second model we introduce is called 3CNN-2, shown in Fig. 4. According to the structure in Fig. 4, the pre-processed input of this model, like the previous model, enters the embedding layer which allows processing by generating vector numbers. We then pass the generated bi-lingual vectors through three one-dimensional convolutional layers that have 3, 5, and 7 kernel sizes, respectively. The type of activation functions in convolutional layers is optimally specified by metaheuristic algorithms.

The 3CNN-2 model is designed in such a way that before combining the layers of the two languages, a dense layer is applied for each of them. Then, to produce a common vector of Persian and English languages, we combine the outputs of the two branches and by passing this common vector through a batch normalization layer, we send the result to a dense layer to detect the final polarity. The output of this numerical model is in the range of 0 and 1, which, like the previous model, expresses the tendency of the input polarity to each of the negative and positive classes. In this model, the activation functions used in dense layers are determined using the optimization power of meta-heuristic algorithms.

Bi-LSTM is applied on the output of the embedding layer to process the sequences of arbitrary length and extract long dependencies in both forward and backward directions. Moreover, we used Bi-LSTM because it is designed to handle the vanishing/exploding problem of traditional RNNs. In the current study, we also used CNN because it is used in NLP applications for local feature extraction.

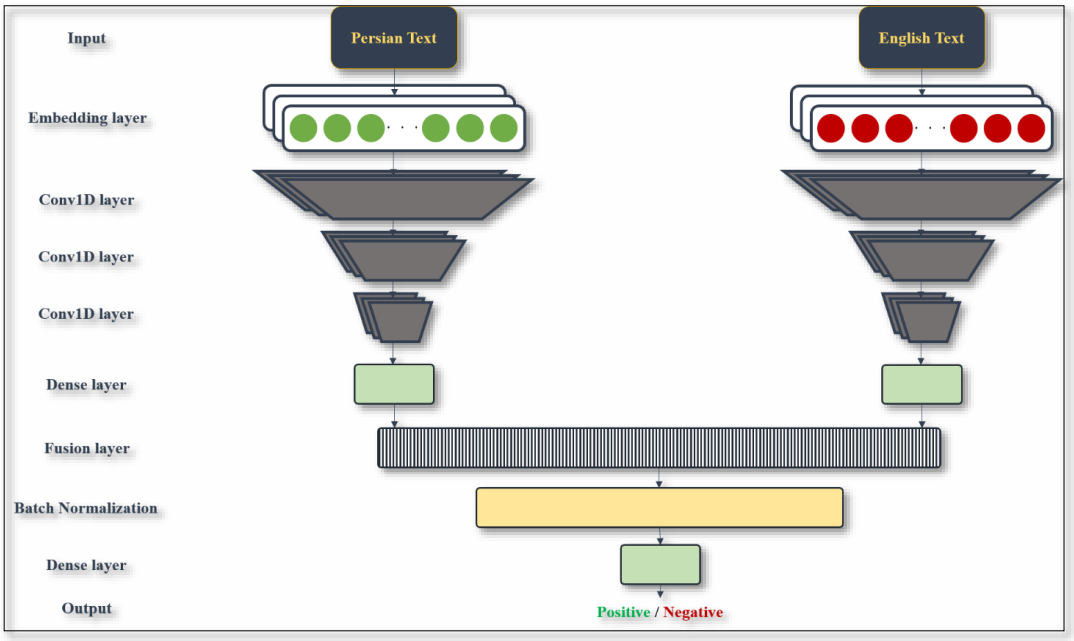


Figure 4. The proposed 3CNN-2 model for sentiment analysis of Persian online doctor reviews.

[DOI: 10.61186/itrc.16.1.42]

This is achieved by convolution operation on the input features via linear filters. Both deep networks have shown promising results in different sentiment analysis research [10] and we expect to achieve comparable results.

VI. META-HEURISTIC ALGORITHMS

Meta-heuristic algorithms belong to a category of optimization algorithms designed to tackle intricate problems by systematically navigating the search space in order to discover solutions that are close to being optimal. In contrast to conventional optimization methods that depend on particular problem structures or mathematical characteristics, meta-heuristic algorithms are versatile approaches that can be employed across diverse problem domains. Meta-heuristic algorithms often derive inspiration from natural or social phenomena, such as evolution, swarm behavior, or physical processes. By emulating these phenomena, the algorithms engage in an intelligent exploration of the solution space, aiming to discover favorable solutions. One advantage of meta-heuristic algorithms is their ability to find near-optimal solutions in reasonable time frames, even for highly complex problems. However, they do not guarantee to find the global optimum and can sometimes get stuck in local optima [30].

One of the major challenges in deep neural models is the selection of appropriate hyper-parameters in the training phase. On the other hand, finding the optimal combination of these hyper-parameters is a difficult task that requires a lot of experience; In addition, this combination may vary in each project. After designing deep bi-lingual neural models, to detect the polarity of online medical opinions in Persian, it is time to optimize the performance of these models. In the last step of the second phase, we optimize the bi-lingual models using genetic algorithms (GA), ant colony optimization (ACO), and gray wolf optimization (GWO) methods. The reason for selecting these meta-heuristic algorithms in the current study is as follows.

GA is a widely-used metaheuristic optimization algorithm that is used as a baseline optimization algorithm in several domains [31]. It has been also used in sentiment analysis-related tasks and has shown great performance for different tasks in this domain [31]. ACO has been also successfully used for different text mining tasks including feature selection [32] and is used recently for several sentiment analysis tasks [33]. Finally, GWO is a relatively newer optimization algorithm that has shown its potential for solving different problems in sentiment analysis literature [34]. We selected this newer algorithm to compare its performance for hyperparameter optimization with GA and ACO algorithms.

In the following, the list of allowed values for each hyperparameter is introduced, and the implementation of each metaheuristic algorithm is described.

For the Bi-LSTM-2 model, several units, activation function, optimizer, loss function, batch size, and the number of epochs are hyper-parameters of the model and we try to find their optimal values using evolutionary algorithms. Hyper-parameters of the 3CNN-2 model are the same except for the number of units which is not applicable here. Table 3 shows the acceptable values we tested for the hyper-parameters.

A. *Genetic Algorithm*

The genetic algorithm is one of the most widely used meta-heuristic algorithms that is inspired by the genetics of living organisms [31][35]. Fig. 5 shows the different stages of the genetic algorithm in this study.

In the first step, using the list of allowed values, the initial population is made up of 10 chromosomes. Each gene on a chromosome contains an allowable value of hyper-parameters, which are randomly selected. Each of the generated chromosomes represents a combination of hyper-parameters needed to teach a bi-lingual model. In the second step, these chromosomes are evaluated using the fitness function; for each chromosome, a bi-lingual neural model is run once and after training, it is evaluated with the test data. The training and testing process is the same for both bi-lingual models in this study.

Once the population has been assessed, it is time to select the parent to implement inheritance in the genetic algorithm. In this method, using a roulette wheel, valuable chromosomes are selected with a higher probability than other chromosomes. By selecting two parents, the input of the fourth stage is prepared. In this step, we combine the parent chromosomes with the multipoint method. The result is two new chromosomes that inherit the parent hyper-parameters. The generated chromosomes are then mutated and added to the current population.

TABLE III. ACCEPTABLE VALUES FOR THE HYPERPARAMETERS.

Hyperparameter	Allowed values
number of units	7, 10, 16
activation function	"sigmoid", "relu", "selu", "elu", "softmax", "softplus", "tanh"
optimizer	"adamax", "adadelta", "adam", "adagrad", "rmsprop", "nadam"
loss function	"binary_crossentropy", "mse", "mae"
batch size	2, 4, 8, 16, 32, 64
number of epochs	range (5, 50)

[DOI: 10.61186/itrc.16.1.42]

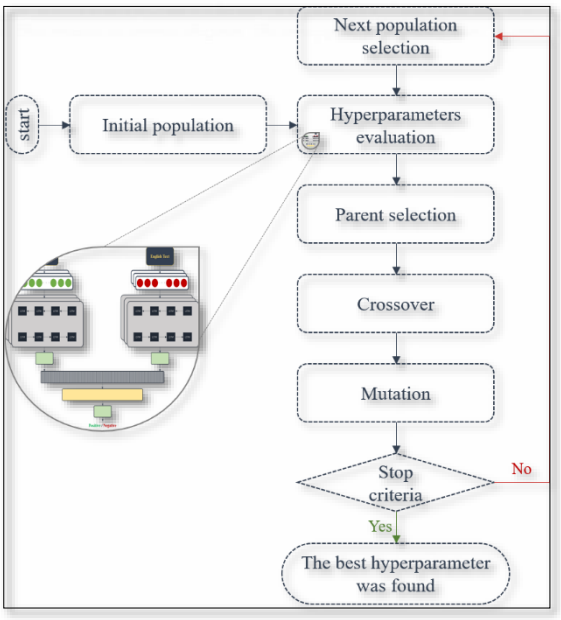


Figure 5. Steps of the genetic algorithm for optimizing proposed bi-lingual deep models.

If the accuracy of the bi-lingual model is reached 99% by each of the chromosomes, the algorithm is terminated and the genes on the best chromosome are introduced as the optimal hyper-parameters; otherwise, two members of the weakest chromosomes of the previous generation are removed and the remaining 10 chromosomes are selected as the next generation. Then the algorithm goes to the second stage to evaluate hyper-parameters. This cycle is repeated until the final condition is met or the number of generations of the algorithm is completed. Finally, the best member of the population is selected as the most valuable hyper-parameters to determine the polarity of online medical reviews.

B. Ant Colony Optimization Algorithm

The second meta-heuristic algorithm in this research for optimizing the proposed deep neural networks is ant colony optimization. Ant Colony Optimization (ACO) is a meta-heuristic algorithm that takes inspiration from how ants forage. It has proven to be valuable in addressing optimization problems characterized by graph-like structures, including challenges like the traveling salesman problem and vehicle routing problem. The ACO algorithm emulates the behavior of real ant colonies, where individual ants leave pheromone trails while searching for food. These pheromone trails act as channels of communication, allowing ants to exchange information about favorable paths to food sources. ACO leverages this natural mechanism with the aim of discovering optimal or near-optimal solutions to complex problems. In the ACO algorithm, a group of artificial ants is utilized to explore the solution space. Each ant represents a potential solution and traverses the problem graph, systematically constructing paths based on predefined rules. At each step, an ant determines its next move by considering two factors: the quantity of pheromone present on the edges and a heuristic estimation of the desirability of the particular move [36]. Fig. 6 shows the steps of the ant colony algorithm.

In the first step of this algorithm, all possible paths are created. The second step is to select a certain number of ants. In this study, 10 ants are determined to move in the space of the paths. In the third step of the ant colony algorithm, a path is selected for each ant. The route chosen is a combination of the ant's personal choice and the suitability of the route according to the number of pheromones in the route. In this study, the effect of the pheromone and the ant's random choice are considered the same. Ants' random selection is used for escaping from local optimums [36], [37].

In the fourth step, the selected paths are evaluated by simulating the passage of ants. For each path, a bi-lingual model is implemented once, the result of which will affect the pheromone level of the path; the higher the accuracy of the model, the higher the amount of pheromone produced. In the next step, the pheromone in all paths is updated. At this stage, we have set the evaporation coefficient to 0.5. In this algorithm, we have set the threshold value of 99% as the stop condition. If the stop condition did not meet, the algorithm jumps to the third step and this cycle continues until the stop condition is met or the number of iterations reaches 50. Finally, the hyper-parameters in the best path are used to optimize the answers.

C. Gray Wolf Optimization Algorithm

The gray wolf algorithm can be used for optimizing various problems by simulating the group life of wolves. In the current study, the gray wolf algorithm is implemented for the first time for optimizing deep neural networks. Fig. 7 shows the implementation steps of the gray wolf algorithm. The first step is to design its state space [38][39]

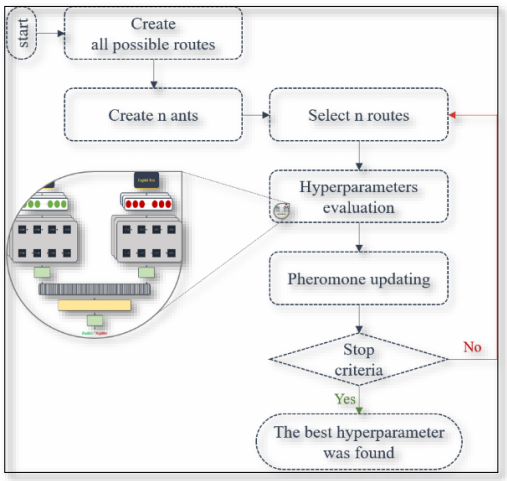


Figure 6. Steps of the ant colony optimization algorithm.

[DOI: 10.61186/itrc.16.1.42]

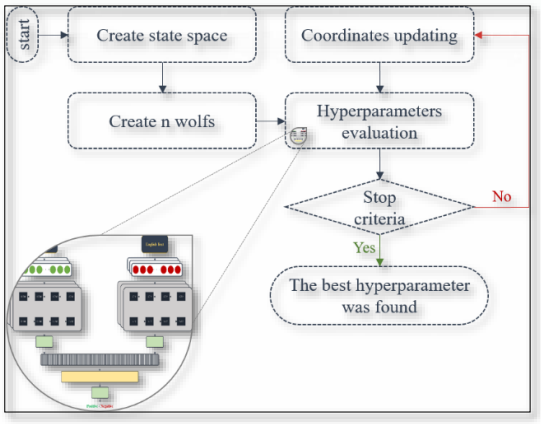


Figure 7. Steps of the gray wolf algorithm for optimizing proposed bi-lingual deep models.

In the first step, a multidimensional space is designed, each axis of which corresponds to a hyperparameter. This creates the coordinates needed to release the wolves into space. Then, 10 wolves are generated in space, and the coordinates of each define a combination of hyper-parameters needed to run proposed bi-lingual models. In the third step, we evaluate the coordinates of each of the wolves. In this process, a bi-lingual neural model is performed based on the hyper-parameters in the coordinates of each wolf. The top three wolves in the evaluation phase are considered the alpha, beta, and delta wolves, respectively, and the rest are called omega.

The fourth step of the algorithm examines the final condition. If the top wolves of the population reach the threshold accuracy of 99%, the algorithm ends and the alpha wolf coordinates are returned as the optimal hyper-parameters. Otherwise, omega wolves update their coordinate hyper-parameters with the help of the coordinates of the top three wolves to get closer to the algorithm target. Then the population goes to the third stage and this cycle continues until the goal is reached or the number of iterations reaches 50. Finally, the hyper-parameters in the position of the most valuable wolf are used to optimize results.

VII. Experimental Results

A. *Experimental Settings*

All steps in this research have been implemented using Python programming language in the Google Collaboratory environment. This Linux environment eliminates the need for powerful hardware to train machine learning models by providing a cloud-based processing space. The specifications of the hardware provided by this environment are as follows:

- GPU: 1x Tesla K80, 12GB GDDR5 VRAM
- CPU: 1x Single Core Hyper Threaded Xeon Processors @2.3Ghz, 45MB Cache
- RAM: 12.6 GB
- Disk: 68.3 GB

B. *The First Phase of Research*

As discussed in the previous section, we conducted our study in two phases; The first phase includes implementing four deep models for sentiment analysis of Persian online doctor reviews and comparing them with four traditional machine learning models such as Naïve Bayes (NB), decision tree (DT), support vector machines (SVM), and random forest (RF) as described in previous sections.

In Fig. 8, the accuracy of all the algorithms and models used in the first phase of the research are compared. According to the results obtained in this phase, deep learning models showed better performance than traditional machine learning models. Among the traditional machine learning algorithms, the best result was achieved by a random forest model with an accuracy of 74%, but still, its accuracy was lower than all deep learning models. Among the deep learning models, Bi-LSTM was able to rank at the top with 81% accuracy.Table 4 shows the performance of models in negative, positive, and average classes. The results in Table 4 show the superiority of models based on deep neural networks over traditional machine learning algorithms.

C. *Results of bi-lingual mode*

To show the performance of models in classifying the English translation of the reviews, we showed the results obtained using both traditional and deep models on the Persian and English reviews in Table 5.

The results in the table show that all deep models outperform traditional models, significantly. Therefore, for the next phase, we only select the deep models for further improvement by applying metaheuristic algorithms to their hyper-parameters.

D. *The second phase of research*

In the second phase of the research, we implement bi-lingual models as previously described in the section "Deep Bi-lingual Models". Each bi-lingual model has two separate inputs; online medical reviews in Persian and their translation in English. To show the effect of using the bi-lingual mode, we summarize the results obtained using four deep models in Table 6.

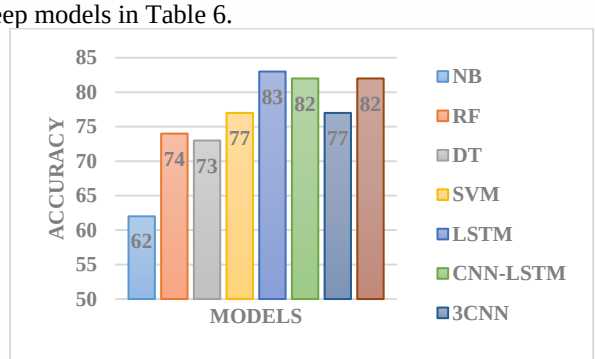


Figure 8. Comparison of models in the first phase of research according to their accuracy. Naïve Bayes (NB), decision tree (DT), support vector machines (SVM), and random forest (RF) are traditional ML classifiers and LSTM, 3CNN-LSTM, 3CNN, and Bi-LSTM are deep models.

[DOI: 10.61186/itrc.16.1.42]

TABLE IV. PERFORMANCE OF MODELS IN THE FIRST PHASE OF RESEARCH. AVG IS THE WEIGHTED AVERAGE OF RESULTS OBTAINED FOR THE NEGATIVE (NEG) AND POSITIVE (POS) REVIEWS.

Model	Label (class)	Precision	Recall	F1
NB [40]	Neg	44%	82%	58%
	Pos	87%	54%	66%
	Avg	74%	62%	64%
DT [40]	Neg	56%	53%	55%
	Pos	79%	81%	80%
	Avg	72%	72%	72%
SVM [40]	Neg	74%	37%	49%
	Pos	77%	94%	85%
	Avg	76%	77%	74%
RF [40]	Neg	60%	52%	56%
	Pos	80%	85%	82%
	Avg	74%	74%	74%
LSTM [41]	Neg	77%	65%	**70%**
	Pos	85%	91%	88%
	Avg	83%	83%	**83%**
CNN-LSTM [42]	Neg	71%	66%	68%
	Pos	85%	88%	86%
	Avg	81%	81%	81%
3CNN [43]	Neg	62%	63%	62%
	Pos	83%	83%	73%
	Avg	77%	77%	77%
Bi-LSTM [11]	Neg	83%	55%	66%
	Pos	82%	95%	**88%**
	Avg	83%	82%	81%

TABLE V. COMPARISON BETWEEN THE PERFORMANCE OF MODELS ON PERSIAN (PER) AND ENGLISH (EN) REVIEWS.

Model	Lang.	Precision	Recall	F1
NB [40]	Per	74%	62%	64%
	En	68%	60%	61%
DT [40]	Per	72%	72%	72%
	En	65%	66%	65%
SVM [40]	Per	76%	77%	74%
	En	73%	74%	70%
RF [40]	Per	74%	74%	74%
	En	73%	72%	73%
LSTM [41]	Per	83%	83%	**83%**
	En	83%	83%	83%
CNN-LSTM [42]	Per	81%	81%	81%
	En	79%	80%	81%
3CNN [43]	Per	77%	77%	77%
	En	82%	82%	82%
Bi-LSTM [11]	Per	83%	82%	81%
	En	84%	84%	**84%**

As shown in Table 6, both LSTM-based models perform similarly and outperform CNN-based models. The 3CNN-2 model performs slightly better than the CNN-LSTM model. Therefore, we selected Bi-LSTM2 from the first type and 3CNN-2 from the second category for further improvement through optimization using metaheuristic algorithms. As discussed in the section "Meta-heuristic algorithms", in this research, GA, ACO, and GWO algorithms are used to optimize and increase the accuracy of the deep models.

TABLE VI. COMPARISON BETWEEN THE PERFORMANCES OF DEEP BILINGUAL MODELS.

Model	Accuracy	F1
LSTM [41]	84%	**84%**
CNN-LSTM [42]	82%	82%
3CNN-2 [43]	83%	83%
Bi-LSTM-2 [11]	84%	84%

To this aim, a total of 1000 bi-lingual neural networks were implemented for each of the genetic algorithms, ant colony, and gray wolf optimization methods. In other words, in the second phase, 3000 bi-lingual neural networks were used for training and testing. Fig. 9 shows the final results of this step for Bi-LSTM-2 and 3CNN-2 methods concerning their accuracy.

According to Fig. 9, all three algorithms based on the two bi-lingual models have presented similar results. In the 3CNN-2 model's optimization process, the ant colony algorithm performed better than the others with 85% accuracy. In optimizing the Bi-LSTM-2 model, the genetic algorithm with 89% accuracy is superior to other algorithms.

Several studies compare the ACO and genetic algorithm for different optimization problems [37]. The main difference between ACO and GA is their underlying principles. It can be said that both ACO and GA have their own strengths and weaknesses when it comes to optimizing hyperparameters of deep neural networks. The choice between the two depends largely on the specific characteristics of the problem being solved and the available computational resources. However, ACO converges faster than GA in most cases due to its ability to quickly exploit promising search space regions.

In general, the Bi-LSTM-2 bi-lingual model performed better than the other models in this study. Using the power of optimizing the genetic algorithm, this model was able to improve the highest accuracy obtained in the first phase of the research by 5%. The results obtained in the second phase show that by using the properties of meta-heuristic algorithms, the third goal of the research has been achieved.

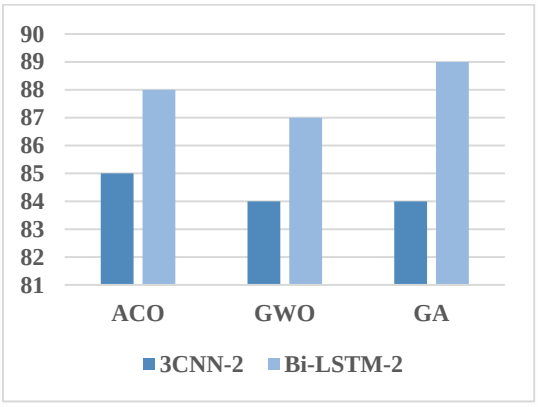


Figure 9. Comparison of models in the second phase of research.

Fig. 10 summarizes the best results obtained in the implementation of all the models discussed in this research. According to this figure, the best performance among all models was achieved by the deep bi-lingual model Bi-LSTM-2.

[DOI: 10.61186/itrc.16.1.42]

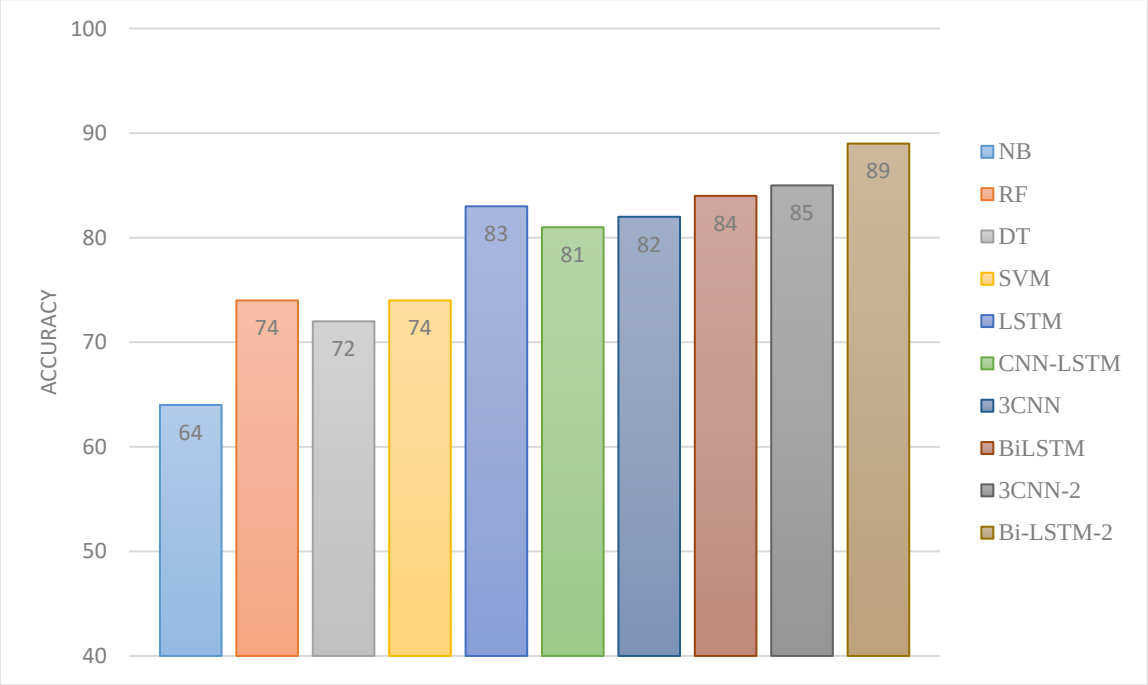


Figure 10. Comparison of the best results obtained by all research models.

VIII. Conclusion

With the ever-expanding Internet and online medical platforms around the world, large volumes of medical data are being generated at an ascending rate.

One type of medical data is online user feedback on the physician experience. The information contained in these user reviews is of great importance. Choosing the right doctor, on the other hand, is a relatively difficult and time-consuming task that can sometimes confuse people. Analyzing the sentiment in online medical reviews can largely solve this problem by extracting valuable sentiment information.

In this study, we proposed a bi-lingual deep model for analyzing online medical reviews in the Persian language. Also, using the optimization power of meta-heuristic algorithms, we optimized the hyper-parameters of the proposed deep model. To show the power of the proposed model we compared it against four traditional machine learning models including support vector machine, decision tree, naïve Bayes, and random forest. Also, to compare these algorithms with deep neural models, LSTM, 3CNN, 3CNN-LSTM, and Bi-LSTM models were selected from the literature and compared with the proposed model. All of these deep models performed better than traditional machine learning algorithms. We proposed the Bi-LSTM-2 model which has two parallel input layers for the Persian and English versions of reviews. The proposed model was optimized using three meta-heuristic algorithms including genetic algorithm, ant colony, and gray wolf optimization. The implementation results showed the superiority of the proposed optimized model in comparison with all mentioned traditional and deep models.

The obtained results show that without optimizing hyper-parameters, the accuracy of LSTM-based methods is slightly higher than CNN-based methods while using meta-heuristic methods to optimize hyper-parameters makes this difference disappear. Also, the obtained results emphasize the use of optimization methods to increase the accuracy of deep learning methods in sentiment analysis. In addition, the findings of this research show that although the use of meta-heuristic methods to optimize the hyper-parameters of learning methods improves the accuracy of the sentiment analysis system, there is not much difference between these meta-heuristic methods. Based on these findings, it can be suggested that other meta-heuristic and population-based methods to optimize the hyper-parameters of deep learning methods should be investigated in future research. Also, other deep learning methods such as methods based on transformers can be used to increase the accuracy of the sentiment analysis system.

Given that the present study is the first bi-lingual meta-heuristic deep model for the sentiment analysis of Persian online doctor reviews, we naturally faced a variety of challenges. The first challenge we encountered in this study was the lack of enough data in the field of online medical opinions in Persian.In this study, by collecting the required data, we proceeded to complement the first collection of online medical reviews in Persian. Another challenge related to the data set was the lack of a proper tag for the reviews, which we also solved by using a questionnaire and a survey of 6 people. One of the most important limitations we encountered in this study was the hardware required to run deep neural models. This prevented the acceleration of the

[DOI: 10.61186/itrc.16.1.42]

research process in the implementation of some models. In each of the metaheuristic algorithms, 50 deep bi-lingual neural networks were implemented, which produced a heavy processing load for the problem. Implementing this number of deep neural networks in a sequential (serial) form is time-consuming. In this study, we tried to solve this problem by parallelizing these models. Unfortunately, due to the lack of access to powerful hardware, it was not possible to use parallel execution.

For future work, the proposed data sets can be expanded to enhance the quality of the learning process in deep neural networks. Another important part of this field is the architecture of deep neural networks. More advanced architectures can help improve the sentiment analysis process. Moreover, several meta-heuristic algorithms can be used for architecture search. For this purpose, open-source toolkits such as OpenNAS may be used for comparing swarm intelligence meta-heuristic algorithms. Another suggestion for future work is to use ensemble methods such as stack ensemble algorithms for improving the performance of optimized deep models for online doctor review sentiment analysis. Given the bi-lingual nature of this research, another suggestion for future work could be to combine other languages. In this way, the system's overall quality can be improved by combining different structures specific to each language.

ACKNOWLEDGMENT

This work has been financially supported by the research deputy of Shahrekord University (grant number 0GRD34M44264).

REFERENCES

[1] M. E. Basiri, S. Nemati, M. Abdar, S. Asadi, and U. R. Acharrya, “A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets,” *Knowl Based Syst*, vol. 228, 2021, doi: 10.1016/j.knosys.2021.107242.

[2] K. Denecke and Y. Deng, “Sentiment analysis in medical settings: New opportunities and challenges,” *Artif Intell Med*, vol. 64, no. 1, pp. 17–27, 2015.

[3] Y. Zhang, M. Chen, D. Huang, D. Wu, and Y. Li, “iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization,” *Future Generation Computer Systems*, vol. 66, pp. 30–35, 2017.

[4] S. Liu and I. Lee, “Extracting features with medical sentiment lexicon and position encoding for drug reviews,” *Health Inf Sci Syst*, vol. 7, no. 1, pp. 1–10, 2019.

[5] S. M. Jiménez-Zafra, M. T. Mart\’\in-Valdivia, M. D. Molina-González, and L. A. Ureña-López, “How do we talk about doctors and drugs? Sentiment analysis in forums expressing opinions for medical domain,” *Artif Intell Med*, vol. 93, pp. 50–57, 2019.

[6] M. Z. Asghar, S. Ahmad, M. Qasim, S. R. Zahra, and F. M. Kundi, “SentiHealth: creating health-related sentiment lexicon using hybrid approach,” *Springerplus*, vol. 5, no. 1, pp. 1–23, 2016.

[7] A. Shukla, W. Wang, G. G. Gao, and R. Agarwal, “Catch me if you can—detecting fraudulent online reviews of doctors using deep learning,” *Ritu, Catch Me If You Can—Detecting Fraudulent Online Reviews of Doctors Using Deep Learning (January 14, 2019)*, 2019.

[8] “W3Techs - World Wide Web Technology Surveys.” Accessed: Jan. 01, 2017. [Online]. Available: https://w3techs.com/

[9] M. E. Basiri and A. Kabiri, “Words Are Important: Improving Sentiment Analysis in the Persian Language by Lexicon Refining,” *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 17, no. 4, p. 26, 2018.

[10] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, “ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis,” *Future Generation Computer Systems*, vol. 115, 2021, doi: 10.1016/j.future.2020.08.005.

[11] M. E. Basiri, R. S. Chegeni, A. N. Karimvand, and S. Nemati, “Bidirectional LSTM Deep Model for Online Doctor Reviews Polarity Detection,” in *2020 6th International Conference on Web Research, ICWR 2020*, 2020. doi: 10.1109/ICWR49608.2020.9122289.

[12] L. Abualigah, H. E. Alfar, M. Shehab, and A. M. A. Hussein, “Sentiment analysis in healthcare: a brief review,” *Recent advances in NLP: the case of Arabic language*, pp. 129–141, 2020.

[13] A. Zunic, P. Corcoran, and I. Spasic, “Sentiment analysis in health and well-being: systematic review,” *JMIR Med Inform*, vol. 8, no. 1, p. e16023, 2020.

[14] R. D. Sharma, S. Tripathi, S. K. Sahu, S. Mittal, and A. Anand, “Predicting online doctor ratings from user reviews using convolutional neural networks,” *Int J Mach Learn Comput*, vol. 6, no. 2, p. 149, 2016.

[15] S. Baccianella, A. Esuli, and F. Sebastiani, “Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining,” in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*, 2010.

[16] D. C. Edara, L. P. Vanukuri, V. Sistla, and V. K. K. Kolli, “Sentiment analysis and text categorization of cancer medical records with LSTM,” *J Ambient Intell Humaniz Comput*, pp. 1–17, 2019.

[17] J. Serrano-Guerrero, M. Bani-Doumi, F. P. Romero, and J. A. Olivas, “Understanding what patients think about hospitals: A deep learning approach for detecting emotions in patient opinions,” *Artif Intell Med*, vol. 128, p. 102298, 2022.

[18] E. Cambria, R. Speer, C. Havasi, and A. Hussain, “Senticnet: A publicly available semantic resource for opinion mining.,” in *AAAI fall symposium: commonsense knowledge*, 2010.

[19] A. M. Shah, X. Yan, S. Tariq, and M. Ali, “What patients like or dislike in physicians: Analyzing drivers of patient satisfaction and dissatisfaction using a digital topic modeling approach,” *Inf Process Manag*, vol. 58, no. 3, p. 102516, 2021.

[20] H. Hao and K. Zhang, “The voice of Chinese health consumers: a text mining approach to web-based physician reviews,” *J Med Internet Res*, vol. 18, no. 5, p. e108, 2016.

[21] A. Sboev *et al.*, “Extraction of the relations among significant pharmacological entities in Russian-language reviews of internet users on medications,” *Big Data and Cognitive Computing*, vol. 6, no. 1, p. 10, 2022.

[22] R. D. Sharma, S. Tripathi, S. K. Sahu, S. Mittal, and A. Anand, “Predicting online doctor ratings from user reviews using convolutional neural networks,” *Int J Mach Learn Comput*, vol. 6, no. 2, p. 149, 2016.

[23] H. Hao, K. Zhang, and others, “The voice of Chinese health consumers: a text mining approach to web-based physician reviews,” *J Med Internet Res*, vol. 18, no. 5, p. e4430, 2016.

[24] A. M. Shah, X. Yan, S. Tariq, and M. Ali, “What patients like or dislike in physicians: Analyzing drivers of patient satisfaction and dissatisfaction using a digital topic modeling approach,” *Inf Process Manag*, vol. 58, no. 3, p. 102516, 2021.

[25] J. Serrano-Guerrero, M. Bani-Doumi, F. P. Romero, and J. A. Olivas, “Understanding what patients think about hospitals: A deep learning approach for detecting

[DOI: 10.61186/itrc.16.1.42]

emotions in patient opinions," *Artif Intell Med*, vol. 128, p. 102298, 2022.
[26] A. Sboev *et al.*, "Extraction of the Relations among Significant Pharmacological Entities in Russian-Language Reviews of Internet Users on Medications," *Big Data and Cognitive Computing*, vol. 6, no. 1, 2022, doi: 10.3390/bdcc6010010.
[27] R. Ghasemi, S. A. Ashrafi Asli, and S. Momtazi, "Deep Persian sentiment analysis: Cross-lingual training for low-resource languages," *J Inf Sci*, vol. 48, no. 4, pp. 449–462, 2022.
[28] M. Aliramezani, E. Doostmohammadi, M. H. Bokaei, and H. Sameti, "Persian sentiment analysis without training data using cross-lingual word embeddings," in *2020 10th International Symposium onTelecommunications (IST)*, 2020, pp. 78–82.
[29] B. Anbalagan, "Detecting Depressive Online User Behavior During Global Pandemic by Fusing LSTM and CNN Models," in *Proceedings of 2nd International Conference on Artificial Intelligence: Advances and Applications*, 2022, pp. 1–10.
[30] K. Rajwar, K. Deep, and S. Das, "An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges," *Artif Intell Rev*, pp. 1–71, 2023.
[31] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimed Tools Appl*, vol. 80, no. 5, pp. 8091–8126, 2021.
[32] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text feature selection using ant colony optimization," *Expert Syst Appl*, vol. 36, no. 3 PART 2, pp. 6843–6853, Apr. 2009.
[33] S. L. Vemula and N. Rathee, "Bio-Inspired Feature Selection Techniques for Sentiment Analysis–Review," in *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, 2023, pp. 808–816.
[34] M. A. Salam and M. Ali, "Optimizing extreme learning machine using GWO algorithm for sentiment analysis," *Int J Comput Appl*, vol. 176, no. 38, pp. 22–28, 2020.
[35] Y. Li, M. Jia, X. Han, and X.-S. Bai, "Towards a comprehensive optimization of engine efficiency and emissions by coupling artificial neural network (ANN) with genetic algorithm (GA)," *Energy*, vol. 225, p. 120331, 2021.
[36] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Comput Intell Mag*, vol. 1, no. 4, pp. 28–39, 2006.
[37] M. E. Basiri and S. Nemati, "A novel hybrid ACO-GA algorithm for text feature selection," in *2009 IEEE Congress on Evolutionary Computation, CEC 2009*, 2009. doi: 10.1109/CEC.2009.4983263.
[38] X. Huang *et al.*, "A Gray Wolf optimization-based improved probabilistic neural network algorithm for surrounding rock squeezing classification in tunnel engineering," *Front Earth Sci (Lausanne)*, vol. 10, p. 857463, 2022.
[39] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in engineering software*, vol. 69, pp. 46–61, 2014.
[40] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.
[41] J. Nowak, A. Taspinar, and R. Scherer, "LSTM recurrent neural networks for short text and sentiment classification," in *International Conference on Artificial Intelligence and Soft Computing*, 2017, pp. 553–562.
[42] X. Wang, W. Jiang, and Z. Luo, "Combination of convolutional and recurrent neural network for sentiment analysis of short texts," in *Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers*, 2016, pp. 2428–2437.
[43] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.

Shahla Nemati was born in Shiraz, Iran, in 1982. She received the B.Sc. degree in Hardware Engineering from Shiraz University, Shiraz, in 2005, the M.Sc. degree from the Isfahan University of Technology, Isfahan, Iran, in 2008, and the Ph.D. degree in Computer Engineering from Isfahan University, Isfahan, in 2016. Since 2017, she has been an Assistant Professor with the Computer Engineering Department, Shahrekord University, Shahrekord, Iran. She has written several articles in the fields of Data Fusion, Emotion Recognition, Affective Computing, and Audio Processing. Her research interests include Deep Learning, Data Fusion, Affective Computing, and Data Mining.
- Email: s.nemati@sku.ac.ir

Reza Salehi Chegeni received the B.Sc. degree in Software Engineering from Lorestan University, Khorramabad, Iran, in 2018, the M.Sc. degree from Shahrekord University, Shahrekord, Iran, in 2021. Currently, he is a Ph.D. candidate in Software Engineering at Amirkabir University of Technology in Tehran, Iran. His research interests encompass Natural Language Processing, Deep Learning, Language Modeling, Evolutionary Algorithms and Graph Data Science.
- Email: reza.sch@yahoo.com

[DOI: 10.61186/itrc.16.1.42]