An Intelligent Fusion Framework of Deep Learning with Secretary Bird Optimization Algorithm for Named Entity Recognition in Arabic Language Texts

Ebtesam Hussain Almansor[1,*]
[1]Computer Sciences Department, Applied College, Najran University, Najran 66462, Saudi Arabia
Email: ehmalmansoor@nu.edu.sa

Abstract

As increasingly Arabic textual data becomes accessible through the Intranet and Internet services, there is an important requirement for technologies and devices to handle the related data. Named Entity Recognition (NER) is an Information Extraction task that became a major part of several other Natural Language Processing (NLP) tasks. NER for Arabic has been obtaining improving attention, but possibilities for development in performance are even accessible. In recent decades, the Arabic NER (ANER) task has been confined to great effort to increase its performance. The ANER difficult task is to collect vast corpora or immense white gazetteers/lists that address probably the majority of Arabic language challenges like complexity, orthography, and ambiguity. Recently, deep learning (DL) has been the most typically applied NER model in the Arabic language and others. DL methods utilize the features of words and text to identify NEs. This paper presents a Secretary Bird Optimization Algorithm for Enhancing Fusion Deep Learning in Arabic Named Entity Recognition (SBOFDL-ANER) model. The main intention of the SBOFDL-ANER technique is to develop an effective method for NER in Arabic text. At first, the text pre-processing stage is applied to clean and transform the raw text into a structured format for analysis. Next, the word embedding method has been implemented by the Word2Vec method. Besides, the proposed SBOFDL-ANER technique designs ensemble models such as deep belief network (DBN), elman recurrent neural network (ERNN), and multi-graph convolutional networks (MGCN) for the process of classification. Eventually, the secretary bird optimization algorithm (SBOA) implements the hyperparameter choice of ensemble models. A wide-ranging simulation was applied to verify the performance of the SBOFDL-ANER method. The experimental outcomes demonstrated that the SBOFDL-ANER model highlighted improvement over other existing methods

Keywords: Fusin Deep Learning; Arabic Language Texts; Named Entity Recognition; Secretary Bird Optimization Algorithm

1. Introduction

Arabic is one of the most productive natural languages in the globe and the official language in the Arab world regarding derivation and morphological inflection [1]. Globally, the top ten most common languages with 420M native speakers and greater than twenty-five common languages. Arabic has multiple unique features that create knowledge of Arabic text either interesting or challenging [2]. Currently, the media platforms are rising, and so is the number of Arabic texts accessible on the Internet. For effectually processing these texts, Arabic Named Entity Recognition (ANER) captivates growing interest. During recent decades, the ANER challenge has gathered much effort to increase its performance [3]. The ANER task is to collect enormous corpora or white lists/gazetteers that handle probably most of the Arabic language namely ambiguity, complexity, and orthography.

NER is the task of recognizing mentions, which match particular kinds, like location, organization, and person name [4]. Moreover, common fields, in some particular areas like the drugs, medical field, and medical processes can be removed by NER. NER might be employed for diverse downstream challenges, namely event extraction,

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

relation extraction, machine translation, and entity linking [5]. Most NER challenges are advanced by leveraging either machine learning (ML)-or rule-based methodologies [6]. The process of NER is that named entities are classified and identified in an open-domain text [7]. NER is the most significant sub-tasks in Information Extraction. NER methods usually allow sub-tasks inside huge NLP methods. The quality of the NER method has a direct effect on the quality of the entire NLP method [8]. NER aims to classify and identify these names automatically in text into pre-defined classes. There has been significant growth in NER over the past decades, and the projected methods are modified by several NE techniques and methods that can be roughly categorized into deep learning (DL), rule-based, and ML methodologies. DL and ML techniques are more effective methods trained and simply enlarged to several language areas [9]. Recently, DL has captivated substantial attention owing to its accomplishments in multiple fields. DL-based NER methods with minimum feature engineering have been flourishing. In recent years, multiple analyses have implemented DL to NER and consecutively upgraded the sophisticated performance [10].

This paper presents a Secretary Bird Optimization Algorithm for Enhancing Fusion Deep Learning in Arabic NER (SBOFDL-ANER) model. At first, the text pre-processing stage is applied to clean and transform raw text into a structured format for analysis. Next, the word embedding method has been implemented by the Word2Vec method. Besides, the proposed SBOFDL-ANER technique designs ensemble models such as deep belief network (DBN), elman recurrent neural network (ERNN), and multi-graph convolutional networks (MGCN) for the process of classification. Eventually, the hyperparameter choice of ensemble models is implemented by secretary bird optimization algorithm (SBOA). The experimental results demonstrated that the SBOFDL-ANER model emphasized improvement over other existing methods.

2. Review of Literature

In [11], a structure is introduced for evaluating geopolitical news involved in public attention and local sentiment. This method concentrates on 4 major subjects employing particular keywords to retrieve appropriate data. The paper builds a Pre-trained Multilingual BERT Method, adjusted for challenges such as sentiment analysis and text classification, highlighting African lower-resource languages. In [12], an NN model that can effectually identify entities in unstandardized online clinical or health text has been projected. Furthermore, this model provides knowledge representations from an experience basis named multi-channel knowledge labels, and this model overwhelms the limitation from languages.

Cordeiro et al. [13] progress a procedural model, which integrates statistical, and artificial intelligence (AI) models to discover and study text data from digital media. The structure, titled Data Analysis Framework for Information and Media (DAFIM), comprises approaches for collecting data done by web scraping and APIs, data enrichment, and textual data processing employing AI solutions, comprising NER and the recognition of clickbait in news. Text clustering and SA models are combined to aid content analysis. Dash et al. [14] handle the difficult challenge of Bio-NER (Bio-NER) over a deep neural network (DNN) model for integrating refined hyperparameters and biological corpus statistics. Pre-trained word embedding namely GloVe or word2vec aid for embedding creation depends upon input. Yang et al. [15] advance a DL-based Landslide NER (LNER) model, which employs BERT for word embedding and combines the Conditional Random Field (CRF) model and projected gradient descent (PGD) adversarial NNs.

Qiu et al. [16] developed an innovative DL technique that incorporates either the GeoBERT or various characteristics to create representations from byte sequences. This representation is consequently passed and fused over a BiLSTM-A method for training. Eventually, entity classification outcomes are attained by utilizing CRF. Zhang et al. [17] introduces the SecureBERT method, surpassing the premier open-source solution for NER in the system security field, utilizes a pre-trained encoder adapted for cybersecurity text for identifying different entities and related data embedded inside natural language sections.

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

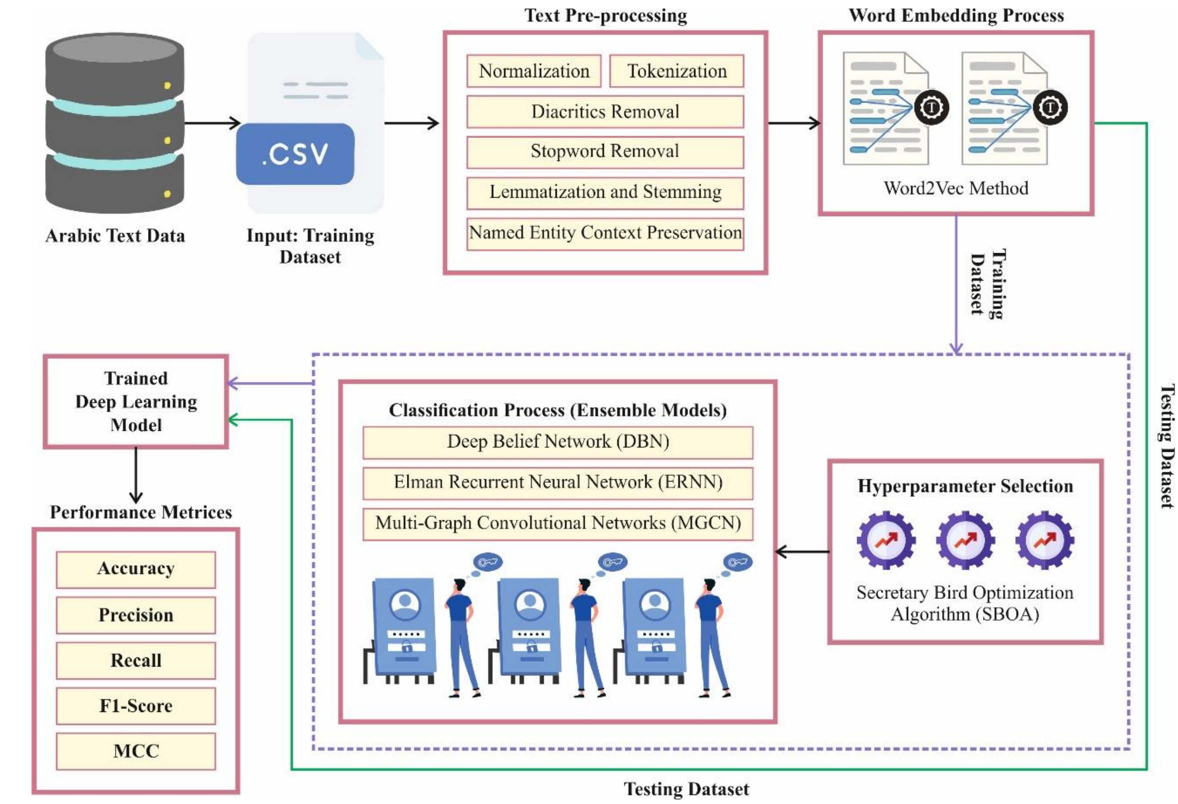

Figure 1. Workflow of SBOFDL-ANER model

3. Methodology

In this paper, we have projected a new SBOFDL-ANER model. The main intention of the SBOFDL-ANER technique is to develop an effective method for NER in Arabic text. It contains various kinds of stages involved as text preprocessing, word embedding, ensemble classification, and parameter fine-tuning. Fig. 1 illustrates the workflow of the SBOFDL-ANER model.

A. Text Preprocessing

At first, the text pre-processing stage is applied to clean and transform the raw text into a structured format for analysis. Arabic NER is a difficult task owing to the composite diacritics, variations, and morphology in the word framework [18]. Efficient text pre-processing is important to enhance the performance of the model in identifying entities such as locations, organizations, and names.

- Normalization

Arabic text has various types of characters that are required to be united. For instance, Alef (أ, إ, آ) is substituted with (ا), and Ta Marbuta (ة) is transformed into (ه). Moreover, lengthened characters (for example ســـلام) are standardized to their standardized forms (سلام).

- Diacritics Elimination

Diacritics (Harakat) such as Kasra (ِ), Fatha (َ), and Damma (ُ) are normally eliminated to decrease complexities and normalize the text model.

- Tokenization

Arabic words were intensely modified, making tokenization demanding. Statistical, DL-based, or Rule-based tokenization techniques are applied to divide text into meaningful components whereas processing clitics (for example: "بالمدرسة" → "بـ" + "المدرسة").

- Stop-word Removal

General Arabic words (e.g., "من", "على", "في") don't benefit NER and were eliminated to improve efficacy.

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

- Lemmatization and Stemming

Stemming decreases words to their base form (for example: "مدارس" → "مدرس"), whereas lemmatization gives the basis dictionary form.

- Named Entity Context Preservation

Unlike general NLP tasks, extreme pre-processing (for example: stop-word removal) should be balanced to maintain context, safeguarding that named entities remain constant.

B. Word2Vec Method

Next, the word embedding method has been implemented by the Word2Vec method. Word2Vec is an NN prediction approach, thus accomplishing it computationally efficiently [19]. The method reflects all words in a large corpus that can be specified as input. The principle underlying this model is that words follow similar conditions and have related meanings. This term has better relationships, and the vectors will overlap. It comprises the Skip-Gram (SG) and the CBOW method. Utilizing its context ("يلبي احتياجات المستهلك بشكل كبير") the CBOW method forecasts the specific word (for example: "احتياجات") SG method performs the opposite, predicting the contextual words assumed in the specific word. Based on consideration of the entire context as the particular remark, the CBOW method accurately smooths over the considerable sum of distributed data. It performs well using small quantities of data. At tth time-step, the CBOW's aim is the word w_t. The method has specified a window of n words nearby $wt,$ and the function of loss J was stated as among Eq. (1):

$$J = \frac{1}{T}\sum_{t=1}^{T} \log\, P(w_t|w_{t-n}, \dots, w_{t-1}, w_{t+1}, \dots, w_{t+n}) \tag{1}$$

In contrast, the SG method forecasts the adjacent words w_{t+j} utilizing the center-word w_t.

$$J = \frac{1}{T}\sum_{t=1}^{T} \sum_{-n\le j\le n, n\ne 0}^{t} ,\ \log P\left(w_{t+j}\middle|w_t\right) \tag{2}$$

C. Ensemble Learning-Based Classification

Besides, the proposed SBOFDL-ANER technique designs ensemble models such as DBN, ERNN, and MGCN for the process of classification.

i) DBN Model

A restricted Boltzmann machine (RBM) can remove features and recover the input database however is not now able to contend with gradient blur [20]. Still, with a clever compound of particular RBMs and a coordinator, the neural network (NN) was made that can resolve this issue. The visible layer (L) contains numerous binary stochastic components that are visible. These components were identified utilizing the visible vector v; furthermore, the hidden layer (L) contains binary random components that are hidden and were identified utilizing the hidden vector characterized by h. Basically, a DBN has been presented to two layers together and has been considered as the RBM. Throughout the system, the hidden layer (HL) of RBM acts as the input layer of RBM. Then, the initial RBM was trained and the outputs were applied as the following input RBM. This procedure is repetitive whereas the output was obtained. Following these training procedures, it can recognize patterns associated with the data. In another type of deeper network, layers generally learn complex designs progressively, a DBN, in contrast, trains hidden designs at the overall level. There were undirected weights and biases amongst the hidden and the visible layers. In explaining the joint layers' distribution function, energy function became mathematical representation utilizing the succeeding equations:

$$\rho\left(L_y, L_h\right) = \frac{e^{-E(L_v, L_h)}}{F_p} \tag{3}$$

$$F_p = \sum L_y\,, L_h e^{-E(L_y, L_h)} \tag{4}$$

Here, F_p validates the partition function that is obtained over the sum-up of the possible pairs for the visible and hidden layers, and L_{h_j} and L_{y_i} establish the binary condition of i^{th} visible and j^{th} hidden layer, correspondingly.

Whereas, $E(L_y, L_h)$ validates the jointed function energy of hidden and visible layers, and its grades were attained utilizing the equation described under:

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

$$E(L_v, L_h) = -\sum i = 1\alpha_i L_{v_i} - \sum j = 1\beta_j L_{h_j} - \sum i\,, jL_{h_j} w_{ij} \quad (5)$$

Whereas, w_{ij} presents the weight between visible and hidden components, α_i and β_j establishes the biases existing in visible and hidden layers. To update the RBM weights, the succeeding equation has been utilized:

$$\Delta w_{ij} = E_t\left(L_{y_i} L_{h_j}\right) - E_m\left(L_{y_i} L_{h_j}\right) \quad (6)$$

Whereas, $E_t(L_{y_i} L_{h_j})$ and $E_m(L_{y_i} L_{h_j})$ describe assurance in the training dataset and the method, arranged.

ii) ERNN Method

The ERNN is one of the ML approaches that can mimic dynamic methods owing to the existence of feedback connections between nodes [21]. ERNN is a particular case of recurrent network method in which the individual nodes are organized in a certain fashion. To make responses, hidden layer (HL) of ERNN feeds data from layer of context. The feedback connection can report timing data for output and input designs. Every layer of ERNN has neurons utilizing weights depending on the overall number of input items to create a non-linear function and transfer that data to the following layer. The mathematical expression of the input layer is given:

$$X_{it}(k) = \sum_{i=1}^{n} X_{it}\,(k-1) \quad (7)$$

Now, X_{it} denotes an input at time t with n neuron counts in the layer of input.

$$ner_{jt}(k) = \Sigma_{i=1}^{n} W_{ij} X_{it}(k-1) + \Sigma_{j=1}^{p} C_j r_{jt}(k) \quad (8)$$

W_{ij} specifies the weight of connection among HL and input layer, C_j is the weight of connection among the recurrent layer and HL. The HLs output is given:

$$Z_{jt}(k) = f(net_{jk}(k) = \Sigma_{i=1}^{n} W_j x_{it}(k-1) + \Sigma_{j=1}^{p} C_j R_{jt}(k) \quad (9)$$

The recurrent layer is expressed as:

$$R_{j\tau}(k) = Z_{jt}(k-1) \quad (10)$$

The output layer is calculated:

$$Y_t(k) = f(\Sigma_{j=1}^{p} V_j Z_{jt}(k) \quad (11)$$

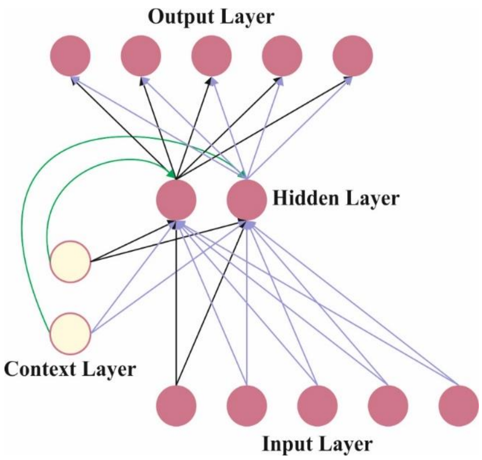

Figure 2. Architecture of ERNN

iii) MGCN System

Multi-graph convolutional networks apply graph convolution operations to the graph representations [22]. To facilitate the expression, utilize A to uniformly signify the normalized proximity matrix, the regularized k-nearest neighbor matrices, and the correlation matrices, for example: $\overline{A}_P, \overline{A}_{N,Q}, \overline{A}_{N,V}$, $A_{R,Q}$ and $A_{R,V}$. Formerly, the initial layer graph convolution operation on speed (flow) observations is typically expressed as

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

$$H_{Q,1} = \sigma\left(A\tilde{Q}W_{Q,0}\right) \tag{12}$$

$$H_{V,1} = \sigma\left(A\tilde{V}W_{V,0}\right) \tag{13}$$

Whereas $W_{Q,0}$ and $W_{V,0}$ represents weighted matrices in the initial layer. $H_{Q,1}$ and $H_{V,1}$ represents hidden matrices gained after the initial layer of graph convolution operations, and σ means activation function. Implementing sequential layers of graph convolution operations. The general term to transition layer z to layer z+1 is as shown.

$$H_{Q,z+1} = \sigma\left(AH_{Q,z}W_{Q,z}\right) \tag{14}$$

$$H_{V,z+1} = \sigma\left(AH_{Q,z}W_{V,z}\right) \tag{15}$$

The output matrices imitating the spatial features of traffic flow and speed observations were removed by performing complete Z layers of graph convolution operations. After the sequence of graph convolutional operations, the output matrices are gained by joining the proximity, k-nearest neighbor, and correlation matrices utilizing a fully connected neural network layer. The output matrices can be compressed into spatial feature vectors as outputs.

D. Hyperparameter Fine-Tuning through SBOA

Eventually, the hyperparameter selection of ensemble models is implemented by SBOA. SBOA is a new population-based metaheuristic model designed by mimicking the secretary bird's survival behaviors in nature [23]. In detail, SBOA generally considered the escape and hunting behaviors of the SB in its natural environment. Generally, the initial stage of the model to implement the problem of FS is to make a collection of candidate solutions.

$$S_i = Lb + r \odot (LIb - Lb), i = 1,2, \dots N \tag{16}$$

Whereas S_i represents the ith individual location. Ub and Lb epitomizer the upper and lower boundary restrictions of the problem is resolved, correspondingly. r symbolizes a random vector in the interval of [0,1]. S_i, Lb, Ub, and r are dimension vectors $1 \times dim$, with dim representing the decision variable counts of the problem is resolved. N signifies the candidate solution counts inside the population. Then the produced N individuals build the population initialization in the iterative procedure of the model stated as the succeeding Eq. (17).

$$S = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_i \\ \vdots \\ S_N \end{bmatrix} = \begin{bmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,j} & \dots & s_{1,dim} \\ s_{2,1} & s_{2,2} & \dots & s_{2,j} & \dots & s_{2,dim} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_{i,1} & s_{i,2} & \dots & s_{i,j} & \dots & s_{i,dim} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_{N,1} & s_{N,2} & \dots & s_{N,j} & \dots & s_{N,dim} \end{bmatrix}_{N \times dim} \tag{17}$$

Whereas S represents the initialized population containing$N$ individuals, $S_j = \left(s_{1,1}s_{1,2} \dots s_{1,j} \dots s_{1,dim}\right)$. After generating the population initialization, the SBOA carries out feature sub-set system tuning by mimicking the escape and hunting behaviors of the SB that are mathematically represented in the following sub-sections. Especially assess the individual's quality utilizing the fitness function (FF) values, as stated by Eq. (18).

$$F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N \times 1} = \begin{bmatrix} F(S_1) \\ \vdots \\ F(S_i) \\ \vdots \\ F(S_N) \end{bmatrix}_{N \times 1} \tag{18}$$

F characterizes the FF value vector, and F_i characterizes the FF value of the ith SB individual.

Mathematical Modelling of Hunting Behavior

In such cases, the hunting behavior of the SB contains 3 key stages namely pursuing one's prey, attacking prey and depleting prey's energy. The mathematical model of the 3 steps that constitute the hunting behavior.

1. Seeking after One's Prey

After searching for prey, the SB is initially required to search for the prey's position and it can rapidly locate the unseen victim among the grass by its benefit in height along with vision, and this procedure is demonstrated in Eq. (19).

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

$$s_{i,j}^{new,P1} = s_{i,j} + \left(s_{random1,j} - s_{random2,j}\right) \cdot R_1 \tag{19}$$

Whereas $s_{i,j}^{new,P1}$ symbolizes the novel state of the jth size of the ith SB individual after location update, $s_{i,j}$ symbolizes the jth dimension value of the ith individual, $s_{random1,j}$ and $s_{random2,j}$ represent the jth dimension values of dual mutually exclusive individuals, correspondingly and R_1 signifies a randomly generated number within the range [0,1]. Afterward, utilize the subsequent formulation to maintain the novel state of the individual.

$$S_i = \begin{cases} S_i^{new,P1}, & if\ F_i^{new,P1} < F_i \\ S_i, & else \end{cases} \tag{20}$$

Here, S_i characterizes the location of the ith individual, $S_i^{new,P1}$ symbolizes the new location of the ith individual update, F_i embodies the FF value of the individual S_i, and $F_i^{new,P1}$ signifies the FF value of individual $S_i^{new,P1}$

2. Depleting Prey's Energy

After looking for prey owing to the combativeness of the victim, the SB is even required to consume the energy of the prey before attack. The depleting prey's energy behavior is demonstrated in Eq. (21).

$$s_{i,j}^{new,P1} = s_{best,j} + e^{(\frac{t}{T})^4} \cdot (R_2 - 0.5) \cdot \left(s_{best,j} - s_{i,j}\right) \tag{21}$$

Now, $s_{besi,j}$ epitomizes the jth dimension value of the high standard individual, e epitomizes exponential process, t symbolizes the present iteration number, T exemplifies the maximal iteration number, and R_2 signifies a randomly generated number that follows a normal standard distribution.

3. Attacking Prey

Once the prey's energy is exhausted and the prey doesn't have the capability for attacking, the SB will initiate an assault on the prey to attain the target of hunting. This procedure is represented by Eq. (20) to maintain the novel state of the individual SB.

$$s_{i,j}^{new,P1} = s_{best,j} + \left(1 - \frac{t}{T}\right)^{2\frac{t}{T}} .. s_{i,j} \cdot RL \tag{22}$$

Here, RL characterizes the fight state of the SB after initiating attacking behavior, stated using the next Eq. (23).

$$RL = 0.5 \cdot Levy(dim) \tag{23}$$

Whereas, Levy(dim) signifies the function of Levy distribution and is resolved by utilizing Eq. (24).

$$Levy\ (dim) = s \cdot \frac{\mu \cdot \sigma}{|\nu|^{\frac{1}{\eta}}} \tag{24}$$

Now, s and η denote constants with values of 0.01 and 1.5 respectively, μ and ν denote randomly generated numbers inside the range [0,1], and σ are described in Eq. (25).

$$\sigma = \left(\frac{\Gamma(1+\eta) \times \sin\left(\frac{\pi\eta}{2}\right)}{\Gamma\left(\frac{1+\eta}{2}\right) \times \eta \times 2\left(\frac{\eta-1}{2}\right)} \right)^{\frac{1}{\eta}} \tag{25}$$

Here, Γ characterizes the gamma function.

Mathematical Modelling of Escape Behaviour

The escape behavior of SB can be primarily separated into two types such as escaping by running or flying away and Hiding with the Environment. The mathematical representation of these dual types of escape behaviors is mentioned in the following sections.

1. Escaping by Run or Fly Away

Once the SB is attack by prey and possesses good flying condition, and the nearby surroundings are not hidden, it will rapidly fly to run away from the prey's attack.

$$s_{i,j}^{new,P2} = s_{i,j} + R_2 \cdot \left(s_{random,j} - K \cdot s_{i,j}\right) \tag{26}$$

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

Whereas, $s_{i,j}^{new,P2}$ epitomizes the novel state of the jth size of the ith SB individual after location update, $s_{i,j}$ and $s_{random,j}$ characterizes the jth dimension value of the ith individual and a random individual, R_2 characterizes a randomly generated number following normal standard distributions, K symbolizes a continuous arbitrarily chosen among the sets $\{1, 2\}$, stated as the succeeding Eq. (27).

$$K = round(1 + rand) \tag{27}$$

Here, $round$ characterizes the integer-up function, and $rand$ embodies randomly generated numbers in the interval of [0,1]. After utilizing the succeeding equation to maintain the novel state of the individual SB.

$$S_i = \begin{cases} S_i^{new,P2}, & if\ P_i^{new,P2} < F_i \\ S_j, & else \end{cases} \tag{28}$$

Now, S_j characterizes the location of the ith individual, $S_i^{new,P2}$ characterizes the upgrade new location of the ith individual, F_i signifies the FF value of the individual S_j, and $F_i^{new,P2}$ embodies the FF value of individual $S_i^{new,P2}$

2. Hiding with the Environment

After being attacked by normal enemies, when the nearby environment is extremely equivalent to itself, the SB will incorporate into the environment to hide and avoid the attack.

$$s_{i,j}^{new,P2} = s_{best,j} + (2 \cdot RB - 1) \cdot \left(1 - \frac{t}{T}\right)^2 \cdot s_{i,j} \tag{29}$$

Here, RB characterizes randomly generated numbers, which emulates a normal standard distribution, and then Eq. (13) is applied to maintain the novel state of the individual.

The SBOA advances a fitness function (FF) to achieve better performance of the classification. It establishes a progressive integer to signify the enhanced performance of the candidate solutions. The reduction of the classification error ratio is measured as the FF, as provided in Eq. (30).

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{no\ of\ misclassified\ samples}{Total\ no\ of\ samples} * 100 \tag{30}$$

4. Performance Validation

Here, the experimental validation of the SBOFDL-ANER technique is examined under the dataset [24]. This dataset holds 400 samples under two classes namely Propaganda and Transparent as depicted in Table 1.

Table 1: Details of dataset

Class	No. of Samples
Propaganda	200
Transparent	200
Total Samples	400

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

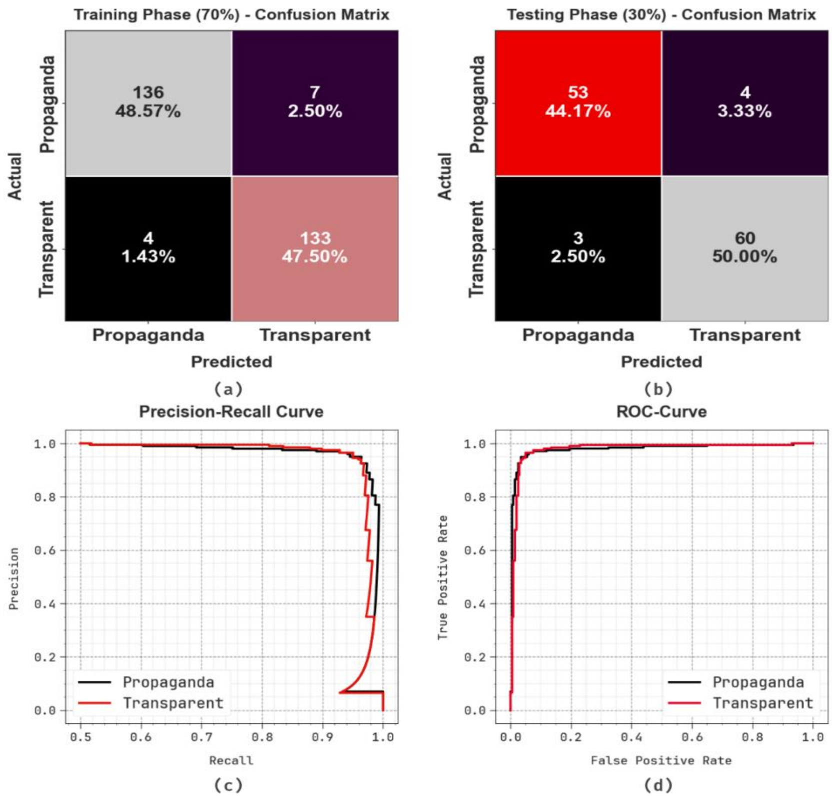

Figure 3. 70%TRAPHA and 30%TESPHA of (a-b) of confusion matrices, (c) curve of PR, and (d) curve of ROC

Fig. 3 illustrates the classifier solutions of SBOFDL-ANER methodology. Figs. 3a-3b depicts the confusion matrix with precise recognition of 2 classes under 70%TRAPHA and 20%TESPHA. Fig. 3c represents the PR analysis, specifying maximal solution within 2 class labels. Eventually, Fig. 3d demonstrates the ROC analysis, specifying capable performance with higher value of ROC for 2 classes.

Table 2 and Fig. 4 depict the classifier performance of SBOFDL-ANER method under 70%TRAPHA and 30%TESPHA. The outcomes suggest that the SBOFDL-ANER methodology appropriately recognized the instances. With 70%TRAPHA, the SBOFDL-ANER model delivers average $accu_y$, $prec_n$, $reca_l, F1_{score}$, and MCC of 96.09%, 96.07%, 96.09%, 96.07%, and 92.16%, respectively. Furthermore, depending on 20%TESPHA, the SBOFDL-ANER model provides average $accu_y$, $prec_n$, $reca_l, F1_{score}$, and MCC of 94.11%, 94.20%, 94.11%, 94.15%, and 88.31%, respectively.

Table 2: Classifier outcome of SBOFDL-ANER model under 70%TRAPHA and 30%TESPHA

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$	MCC
TRAPHA (70%)					
Propaganda	95.10	97.14	95.10	96.11	92.16
Transparent	97.08	95.00	97.08	96.03	92.16
Average	96.09	96.07	96.09	96.07	92.16
TESPHA (30%)					
Propaganda	92.98	94.64	92.98	93.81	88.31
Transparent	95.24	93.75	95.24	94.49	88.31
Average	94.11	94.20	94.11	94.15	88.31

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

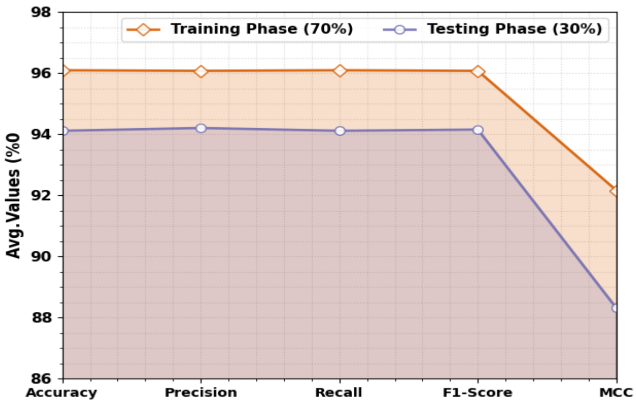

Figure 4. Average of SBOFDL-ANER model under 70%TRAPHA and 30%TESPHA

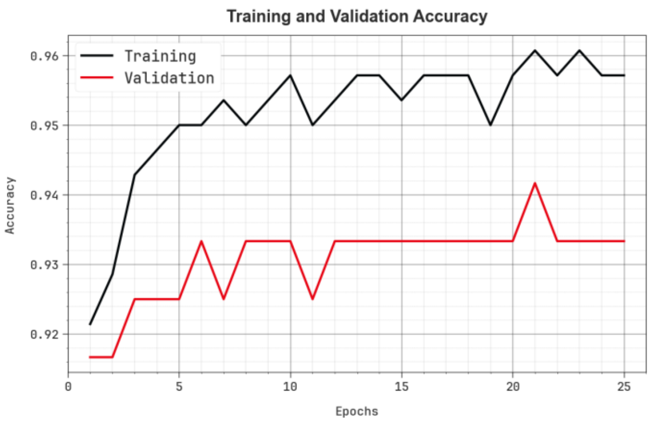

Figure 5. $Accu_y$ Curve of SBOFDL-ANER model

In Fig. 5, the training (TRAN) $accu_y$ and validation (VALN) $accu_y$ performance of the SBOFDL-ANER method is represented. The $accu_y$ values are calculated around an interval of 0-25 epochs. The figure emphasized that the either $accu_y$ values displays an increasing trend that reported the capability of the SBOFDL-ANER method with enhanced outcomes through multiple iteration counts. Furthermore, the either $accu_y$endures closer over the epoch counts, which signifies minimum over-fitting and displays an improved solution of the SBOFDL-ANER technique, ensuring dependable prediction on unseen instances.

In Fig. 6, the TRAN loss (TRANLOS) and VALN loss (VALNLOS) curve of the SBOFDL-ANER technique is shown. The loss values are calculated around an interval of 0-25 epochs. It is depicted that either value shows a minimal trend, reporting the ability of the SBOFDL-ANER method to balance a trade-off between generalization and data fitting. The recurrent decrease in loss values moreover ensures the greater outcome of the SBOFDL-ANER methodology and adjusts the prediction outcomes over time.

Figure 6. Loss curve of SBOFDL-ANER model

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

Table 3 and Fig. 7 demonstrates the comparison solution of SBOFDL-ANER model with current methodologies under several measures [25, 26]. The solution highlighted that the K-Nearest Neighbour, SVC, Neural network, Adaboost, and Att-RNN techniques are attained poor performance. Eventually, Bag of Words and CNN methods attained closer outcomes. Additionally, the presented SBOFDL-ANER technique stated enhanced solution with maximal $accu_y, prec_n, reca_l$, and $F1_{score}$ of 96.09%, 96.07%, 96.09%, and 96.07%, respectively.

Table 3: Comparative analysis of SBOFDL-ANER method with existing models

Technique	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$
SBOFDL-ANER	96.09	96.07	96.09	96.07
CNN Classifier	93.05	95.1	79.09	81.72
K-Nearest Neighbour	82.23	95.87	85.63	84.52
SVC Model	87.75	94.49	86.58	80.65
Neural Network	82.04	80.84	92.61	93.3
Bag of Words	94.35	80.97	95.55	79.68
Adaboost	83.67	87.35	83.89	84.45
Att-RNN	86.39	79.48	94.78	91.7

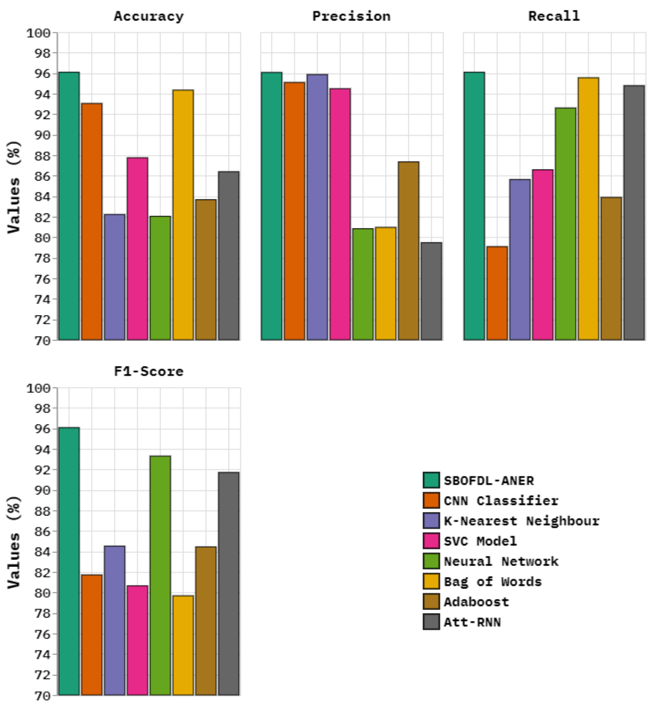

Figure 7. Comparative outcome of SBOFDL-ANER method with existing models

5. Conclusion

In this paper, we have presented a novel SBOFDL-ANER model. The main intention of the SBOFDL-ANER technique is to develop an effective method for NER in Arabic text. At first, the text pre-processing stage is applied to clean and transform raw text into a structured format for analysis. Next, the word embedding method has been implemented by the Word2Vec method. Besides, the proposed SBOFDL-ANER technique designs ensemble

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

models such as DBN, ERNN, and MGCN for the process of classification. Eventually, the hyperparameter choice of ensemble models is implemented by SBOA. A wide-ranging simulation was applied to verify the solution of the SBOFDL-ANER. The experimental outcomes demonstrated that the SBOFDL-ANER model highlighted improvement over other existing methods.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] X. Qu, Y. Gu, Q. Xia, Z. Li, Z. Wang, and B. Huai, "A survey on Arabic named entity recognition: Past, recent advances, and future trends," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 3, pp. 943–959, 2023.

[2] M. Khalifa and K. Shaalan, "Character convolutions for Arabic named entity recognition with long short-term memory networks," *Computer Speech & Language*, vol. 58, pp. 335–346, 2019.

[3] C. Helwe and S. Elbassuoni, "Arabic named entity recognition via deep co-learning," *Artificial Intelligence Review*, vol. 52, pp. 197–215, 2019.

[4] A. Oueslati, F. Bouchahoua, and M. Sellami, "Enhancing Arabic named entity recognition using deep learning and pre-trained contextual embeddings," *International Journal of Information Retrieval Research*, vol. 13, no. 1, pp. 1–16, 2023.

[5] B. A. B. Ali, S. Mihi, I. El Bazi, and N. Laachfoubi, "A recent survey of Arabic named entity recognition on social media," *Revue d'Intelligence Artificielle*, vol. 34, no. 2, pp. 125–135, 2020.

[6] M. Al-Smadi, S. Al-Zboon, Y. Jararweh, and P. Juola, "Transfer learning for Arabic named entity recognition with deep neural networks," *IEEE Access*, vol. 8, pp. 37736–37745, 2020.

[7] N. Alsaaran and M. Alrabiah, "Arabic named entity recognition: A BERT-BGRU approach," *Computers, Materials & Continua*, vol. 68, no. 1, pp. 471–485, 2021.

[8] C. Helwe, G. Dib, M. Shamas, and S. Elbassuoni, "A semi-supervised BERT approach for Arabic named entity recognition," in *Proc. Fifth Arabic Natural Language Processing Workshop*, Dec. 2020, pp. 49–57.

[9] L. Liu, J. Shang, and J. Han, "Arabic named entity recognition: What works and what's next," in *Proc. Fourth Arabic Natural Language Processing Workshop*, Aug. 2019, pp. 60–67.

[10] A. El Mahdaouy, H. Touir, and K. Boukhalfa, "Sentiment analysis and named entity recognition in Arabic social media texts," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 3, pp. 374–386, 2023.

[11] N. Abdou Mohamed, I. Benelallam, and Y. Rahmani, "Monitoring African geopolitics: A multilingual sentiment and public attention framework," *Applied Intelligence*, vol. 55, no. 2, pp. 1–19, 2025.

[12] L. Cao, C. Wu, G. Luo, C. Guo, and A. Zheng, "Online biomedical named entities recognition by data and knowledge-driven model," *Artificial Intelligence in Medicine*, vol. 150, p. 102813, 2024.

[13] D. Cordeiro, C. Lopezosa, and J. Guallar, "A methodological framework for AI-driven textual data analysis in digital media," *Future Internet*, vol. 17, no. 2, p. 59, 2025.

[14] A. Dash, S. Darshana, D. K. Yadav, and V. Gupta, "A clinical named entity recognition model using pretrained word embedding and deep neural networks," *Decision Analytics Journal*, vol. 10, p. 100426, 2024.

[15] B. Yang, C. Zhou, S. Li, and Y. Wang, "A Chinese named entity recognition method for landslide geological disasters based on deep learning," *Engineering Applications of Artificial Intelligence*, vol. 139, p. 109537, 2025.

[16] Q. Qiu et al., "Chinese engineering geological named entity recognition by fusing multi-features and data enhancement using deep learning," *Expert Systems with Applications*, vol. 238, p. 121925, 2024.

[17] Y. Zhang, J. Liu, X. Zhong, and L. Wu, "SecLMNER: A framework for enhanced named entity recognition in multi-source cybersecurity data using large language models," *Expert Systems with Applications*, p. 126651, 2025.

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025

[18] R. Salah, M. Mukred, L. Q. binti Zakaria, and F. A. Al-Yarimi, "A machine learning approach for named entity recognition in Classical Arabic natural language processing," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 18, no. 10, pp. 2895–2919, 2024.

[19] D. Mustafa, S. M. Khabour, M. Al-kfairy, and A. Shatnawi, "Leveraging sentiment analysis of food delivery services reviews using deep learning and word embedding," *PeerJ Computer Science*, vol. 11, p. e2669, 2025.

[20] F. Liu, Z. Dong, and B. Gorbani, "Utilizing deep belief network optimized by balanced Manta ray foraging optimization algorithm for estimating the shear Wall's shear strength," *Scientific Reports*, vol. 15, no. 1, p. 7443, 2025.

[21] B. Sivasakthi, K. Preetha, and D. Selvanayagi, "Osteoporosis disease detection using optimized Elman recurrent neural network based on hybrid bacterial colony optimization and tabu search algorithm," *International Research Journal of Multidisciplinary Technovation*, vol. 7, no. 1, pp. 1–16, 2025.

[22] F. Shao, H. Shao, X. Wu, Q. Cheng, and W. Lam, "A physics-informed machine learning framework for speed-flow prediction: Integrating an S-shaped traffic stream model with deep learning models," *SSRN*, 2025. [Online]. Available: SSRN 5164944.

[23] F. Chen, S. Ye, J. Wang, and J. Luo, "Multi-strategy improved binary secretarial bird optimization algorithm for feature selection," *Mathematics*, vol. 13, no. 4, p. 668, 2025.

[24] B. M. Almotairy, M. Abdullah, and D. H. Alahmadi, "Dataset for detecting and characterizing Arab computation propaganda on X," *Data in Brief*, vol. 53, p. 110089, 2024.

[25] M. Gaikwad, S. Ahirrao, K. Kotecha, and A. Abraham, "Multi-ideology multi-class extremism classification using deep learning techniques," *IEEE Access*, vol. 10, pp. 104829–104843, 2022.

[26] A. N. Maseri, A. A. Norman, C. I. Eke, A. Ahmad, and N. N. A. Molok, "Socio-technical mitigation effort to combat cyber propaganda: A systematic literature mapping," *IEEE Access*, vol. 8, pp. 92929–92944, 2020.

DOI: https://doi.org/10.54216/FPA.190227
Received: December 15, 2024 Revised: February 07, 2025 Accepted: March 03, 2025