

OPEN ACCESS

\*Corresponding author

Soran Badawi

<mailto:soran.sedeeg@chu.edu.iq>

RECEIVED :07 /01 /2025

ACCEPTED : 16/06/ 2025

PUBLISHED : 31/ 10/ 2025

KEYWORDS:

text classification, Fake news detection, optimization, deep learning, machine learning.

# Deep Learning based Fake News Detection For The Kurdish Language

Soran Badawi

Language center, Charmo University, Chamchamal, Kurdistan Region, Iraq

## ABSTRACT

In today's world, technology has facilitated the publication of fake news on social media. Thus, it becomes very difficult to differentiate real news from fake ones particularly in low-resourced languages such as Kurdish. Therefore, in this study, the author bridges this gap by proposing a hybrid deep learning model named "ACO-CNN" that can effectively recognize false news. The hybrid approach includes the Ant Colony Optimization (ACO) that has been utilized to optimize the Convolutional Neural Network (CNN) hyperparameters to enhance its capabilities to successfully detect fake news. The proposed method was evaluated against the state-of-the-art methods that were incorporated for the same purpose including, Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest, (RF) and Decision Tree (DT) using two existing benchmark fake news dataset namely "KurdFake" and "KDFND". Our experimental results indicate that the ACO-CNN outperformed other approaches in terms of F1-score on both datasets. This work heavily contributes to the Kurdish Natural Language Processing (NLP) field and the development of effective deep learned based fake news detection tools for under-resourced languages.

## 1. Introduction

Technological revolution has transformed many areas of life. One of the areas that was impacted by it is news dissemination. Previously, people purchased magazines or tuned to news channels to acquire daily news. However, people currently obtain their news from social platforms such as Facebook and X (previously known as Twitter). This is because the news on these platforms is more timely and cheaper than traditional news media. Moreover, the platforms provide the opportunity to news seekers to voice their concern and share and comment their ideas with other news readers (Shu et al., 2017). Consequently, online users share fake news for political gains and financial success, which ultimately impacts the public opinion in a negative way by misleading others (Jardaneh et al., 2019).

The US election in 2016 serves as the defining moment for scholars to deem fake news as dangerous issue Fake news played major role causing major conflicts among political parties (Horne and Adali, 2017). Additionally, it easily misled and swayed voters by publishing untruthful news about political opponents. Sadly, the online facts checking systems are time consuming as they require human efforts to skillfully detect false claims and remove them from internet. Moreover, the systems place a great deal of emphasis on verifying the political news, neglecting the other political news and other false information that are rapidly shared on social media (Zhang and Ghorbani, 2020). There are several factors that cause fake news detection to be challenging. First, verifying the suspicious news and seeking evidence to the validate the news requires a lot of time. Second, the number of reliable sources to fact check news are limited. Finally, social media content contains noise and lack sufficient information to prove the news. These factors heavily influenced and declined the number of studies conducted in this area, particularly in low-resourced languages.

Kurdish language is an Indo-European language that are largely spoken in northern Iraq, Syria, eastern Turkey and western Iran. It has around 30-40 million users (Badawi et al., 2023). The

detection of fake news has been greatly studied in other languages such as English. They have designed effective and functional systems that can easily block misinformation. Sadly, the attempts on Kurdish language are scarcely limited and require immediate investigation.

In this study, a deep learning-based model called ACO-CNN was proposed to address this issue. Our model utilizes the ant colony optimizer algorithm to optimize the parameters of the convolutional neural network (CNN) and improve its accuracy. To evaluate the effectiveness of our proposed method, it is compared against existing state-of-the-art techniques for fake news detection using two different datasets: KDFND and KurdFake. The main contribution of this paper include;

- An optimized ACO; This study develops an enhanced optimization model for ACO that autonomously adjusts CNN parameters to maximize performance.
- Hybrid Model; The study proposes an ACO-CNN approach that combines CNN and optimized ACO for the optimal classification of Kurdish fake news.
- Comparison against the State-of-the-art models; The study not only focuses on the proposed deep learning approach but also includes a comparative analysis with traditional machine learning classifiers, such as K-Nearest Neighbors (KNN) and Naïve Bayes. The thorough experimental results demonstrate that the ACO-CNN model surpasses the other competing models in terms of recall, precision, accuracy, and F1 score.

The remainder of this paper is structured as follows: In section two, the author presents a literature review of studies focused on fake news detection. Section three provides a detailed explanation of the methodology employed in this study, including the processes of data collection and preprocessing, as well as the proposed method. In section four, the experimental results will be discussed in depth. Finally, section five concludes the study with a summary of our findings.

## 2. Related Works

In this section, the author extensively examines the research conducted on detecting fake news.

it is worth stating that researchers utilized machine learning techniques because of their superiority in the domain of classifications (Chiman Haydar Salh and Ali, 2022, Haddad, 2023). Moreover, they employed deep learning techniques due to efficiency and low-cost intelligent capabilities (Dana Mohammed Ali and A.Sadeq, 2022). First, Ruchansky et al. (2017) developed a powerful model from three models for recognizing fake news. The first model incorporates Recurrent Neural Network (RNN) to study the temporal patterns of user engagement with a given text. The second model examines the user behavior and activity patterns and the authors integrate these two models in third model to distinguish between authentic and fabricated news articles (Ruchansky, 2017). Researchers introduced a hybrid deep learning approach that merges CNN with Long Short-Term Memory (LSTM) to highlight false news in tweets. The hybrid method scored an accuracy of 82% (Ajao, 2018).

Similarly, another group of researchers designed a method that leverages machine learning methods to address issues faced fake news classification such as accuracy, time delays, and the extensive hours required to spend on preprocessing and handling high volume of tweets. They mined 400,000 tweets from the popular HSpam 14 dataset. The dataset contains 150,000 spam tweets and 250,000 tweets that were classified as non-spam. The authors extracted lightweight features and the top 30 words by using information gain from the bag of word model. The proposed method obtained an impressive accuracy of 91.65% outperforming other methods by nearly 18% (Sharma, 2020). The main shortcoming of this study is that it deals with tweets that are written in brief and users neglect the attentions to grammatical sides and punctuation, unlike news articles that are crafted professionally.

Likewise, The researchers developed a model that utilized the supervised machine learning method to identify fake news in Arabic tweets. The proposed method was a feature-based method for representing the data that employed user behavior, content analysis, and sentiment analysis. The proposed method was compared

other machine learning models and could correctly detect fake news in Arabic tweets with an accuracy of 76% (Jardaneh et al., 2019). Barrir and Djefal (2020) incorporated machine learning based methods for identifying fake news. The authors employed N-gram encoding and TF-IDF as the main feature extraction tool to extract critical features from their cleaned dataset which contained information about the author, article text, article source, article published data, and their sentiments. Moreover, SVM was employed and scored 100% in term of recognition rate (Fatima. and Djefal, 2020). The main limitation of this study is that it solely depends on feature selection on texts given the feature sets from one method to another vary that causes the results to be unstable.

Furthermore, the authors designed an AI tool to recognize fake news in the Arabic language. The tool combines multiple text vectorizations methods and machine learning algorithms. The core layer of the proposed model includes merging TF-IDF with XGBoost classifier to eliminate insignificant features and optimize the classifier, which evidently aided the model to achieve an accuracy score of 77.81% and an F-score 80.42% (Thafer et al., 2021).

Amjada et al. (2021) embarked on the task of automatically recognizing false news in the Urdu language. They designed a dataset that contained scraped articles from five different domains including health, sports, business, showbiz, and technology that each was classified either as "fake" or "real". The authors incorporated 60:40 split for training set and test set and experimented with machine learning algorithms and deep learning approaches such as LSTM and BERT. The outcomes indicate that logistic regression outperformed other methods in term of F1-macro score by scoring 90% (Amjada, 2021).

In the domain of Kurdish language, there only two studies conducted on identifying authentic news from false ones. The first study conducted by Azad et al. (2021). The authors initially developed a balanced dataset named "KurdFake" that includes 5000 text for each class that were sourced from Facebook pages and Sorani Kurdish websites. They experimented on

their dataset using five machine learning algorithm including Logistic Regression, Naive Bayes, Decision Tree, Support Vector Machine (SVM), and Random Forest. The experimental results showed that SVM achieved the highest accuracy of 88.71 (Azad et al., 2021).

The second conducted by Salh and Nabil (2023). Similar to the previous work, the authors designed a large dataset named "KDFND" that contains over 100k news articles. The authors extensively experimented on their dataset utilizing various techniques such word embeddings, TF-IDF, and count vectorization. Moreover, they applied RF, SVM and CNN to analyze the news articles. Their findings indicate that CNN was the superior model by achieving an notable accuracy of 95% for F1-score and an accuracy rate of 91% (Salh and Nabi, 2023).

The amount of works conducted on separating real news from fabricated news stresses the pressing need for research to study this issue from multiple perspectives for the Kurdish language. Kurdish scholars should perform extensive research and employ effective strategies to tackle this issue. Hence, in this study, the author proposes a hybrid novel deep learning approach with the purpose of motivating further research and developing innovative approaches to advance Kurdish NLP in this domain.

### 3. Methodology

Our methodology begins with the collection of data for experimentation. Next, the author will use the available tools to preprocess the data. Once the preprocessing is complete, the data will be trained using the proposed method. Finally, the results obtained will be compared from the proposed approach with those from the benchmark method, allowing us to evaluate the effectiveness of our methodology. Fig 1 illustrates a diagram of the proposed method.

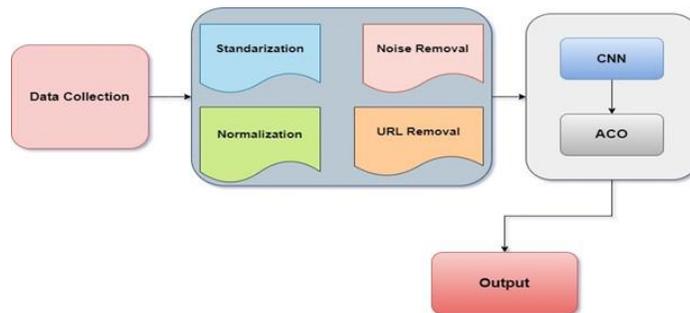


Figure 1: The diagram of the proposed method

#### 3.1 Data Collection

Identifying news articles as "false" can be a challenging and time-consuming task. It is possible for articles to present facts selectively, or use misleading framing while technically stating the truth. Politics, culture, and ideology can influence how the same information is interpreted. However, there is a pre-collected dataset of fake news known as the KDFND that can streamline this process. This dataset was compiled by researchers and is readily available on Mendeley Data (Salh and Nabi, 2023). It consists of news articles from various online resources, including well-known Kurdish news websites that are officially recognized by the Kurdistan Journalists Syndicate. Facebook pages were utilized as the primary source for gathering both fake and real news. The KDFND dataset covers news from three cities in the Kurdistan Region of Iraq: Erbil, Sulaimani, and Halabja. The articles are written in the Kurdish language, specifically in the Sorani dialect. The dataset includes 100,962 articles, categorized as either real or fake news across several topics, including economics, society, politics, entertainment, health, and sports, as provided in Table 1.

**Table 1:** The number of news items in each dataset

Dataset	No. of Fake news	No. of Real news	total
KDFND	50211	50751	100,962
KurdFake	5000	5000	10,000

The second dataset is named “KurdFake” which was developed by Azad et al. (2021). To create this corpus, two steps were followed. Initially, the authors scraped 5,000 authentic news articles from the official Kurdish news channels on Facebook. The articles covered various categories including, sport, technology, health, economy and etc. However, the lack of fact-checking websites and tools in the Kurdish language posed a major challenges while mining fake news contents in the second step. To optimize time, two techniques were employed to generate two sets of data. Set One contained 5,000 news articles from non-legitimate Facebook pages that met specific criteria, while Set Two included 5,000 news articles that were automatically manipulated using a Python script to create contradictory meanings. Both sets were then combined with the real news articles, resulting in a total of 10,000 annotated news articles.

### 3.2. Data Preprocessing

Pre-processing is a crucial phase in machine learning and deep learning. It helps classifiers produce better results. From a morphological perspective, Kurdish is a language rich in attachments, influenced by Arabic and Persian. Understanding these extensions can provide insights into pronouns, plurality, and prepositional location. However, pre-processing presents unique challenges due to the complexities involved in NLP for Kurdish. Fortunately, the author utilized the KLPT toolkit developed by Ahmedi (2020) as shown in Table 2. The libraries within KLPT assisted us with normalization, standardization, and tokenization (Ahmedi, 2020). It is important to note that there are no specific libraries available for identifying stop-words in the Kurdish language. Therefore, the author created the stop-word lists and

applied them to the datasets.

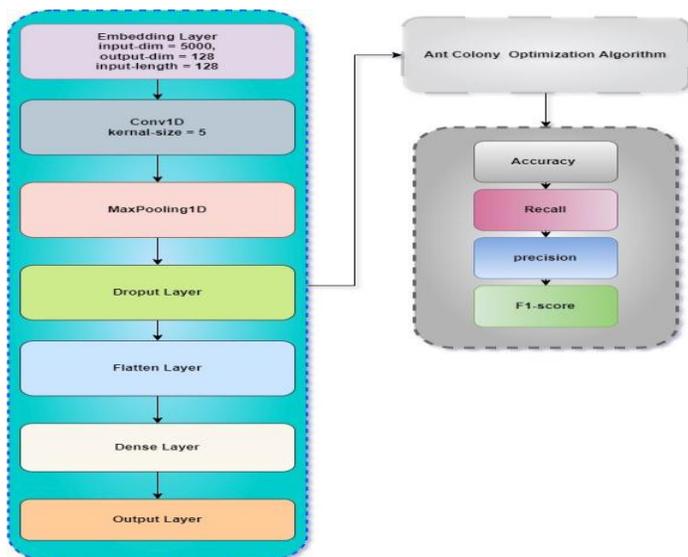
### 3.3. The Proposed Method

#### 3.3.1. CNN Architecture

The Convolutional Neural Network (CNN) architecture is designed to be flexible, allowing for adjustable hyperparameters as shown in Fig The network begins with an Embedding Layer that transforms tokenized text sequences into dense vectors. The output size of this layer is set to 128, meaning each word in the sequence is represented by a 128-dimensional vector. Following the embedding layer, the CNN consists of a series of Convolutional Layers. The number and configuration of these layers are not fixed but are determined dynamically using the ant colony optimization algorithm. The layers contains a Conv1D and a Rectified Linear Unit(ReLU) that serves at its activation function. It will be connected it with a MaxPooling1D to reduce the dimensionality of the training data (Badawi, 2024). To prevent the model from overfitting and yielding incorrect results, a dropout layer is added with a rate of (0.5). The dropout layer is followed by a flatten layer to reshape the output of convolutional layers into a single layer dimension, thus subsequently connecting it to full layers, which is then passed through a dense layer in which it's unit values were determined and optimized the ACO. This layer employs a ReLU activation function, introducing non-linearity to the network and enhancing its ability to learn from the data. Finally, the Output Layer of the CNN is designed for binary classification, as there are two classes to predict. This layer utilizes the softmax activation function to output probabilities for each of the two classes

In various classification tasks, CNNs have demonstrated impressive performance. There is, however, recent research that indicates that such methods are prone to overfitting, rely on dataset-specific artifacts, and tend to learn superficial features instead of deep semantic representations (McCoy, 2020, Lapuschkin et al., 2019). To address these issues, the proposed approach utilizes ACO in order to systematically tune the CNN's hyperparameters. Using this bio-inspired optimization technique, a wide range of





**Figure 2:** The diagram of the proposed method  
**3.3.2 Ant Colony**

Ants are found on every continent except Antarctica and thrive in organized colonies that rely heavily on pheromones for communication. These pheromones serve various purposes, including marking paths to food, signaling danger, and recognizing different types of ants. When an ant discovers a food source, it secretes a pheromone that leaves a trail back to its colony. Other ants that encounter this food source also contribute to the pheromone trail, creating a network that leads back to the food. Ants leverage the pheromone that carries the strongest signal that leads to starting an autocatalytic process that normally guide the ants to detect the shortest path to the food source. Eventually, the pheromone route vanishes into thin air that ultimately affects its signal and reduces its strength. Luckily, this process assists the ants not select the old trails that may misled them to bad food. These behaviors by the ants follows the natural rules that have inspired the development optimization algorithm to exploit its capacity to find solutions for the problems in fields of mathematics, computer science and engineering (Dorigo et al., 2006).

Needless to say, ACO is regarded as the effective method that successfully and efficiently addresses complex problems by being implemented in multiple industrial domains. Conversely, the traditional version of ACO has its

own pitfalls that can hinder its capacity to yield the optimal scores. In this paper, the author aims to exploit the power of ACO to uncover the optimal parameters and reduce training time for CNN. Our works begins with optimizing ACCO and fuses it with CNN.

The main weakness of traditional ACO is merely dependent on the ants pheromone trails and heuristic information guide other ants. Thus, additional parameters is introduced that is linked to CNN hyperparameters. The additional parameters helps the algorithm to cover more areas in the search space of CNN configurations more efficiently. The edge attractiveness function is equated as follows:

$$T_{ij} = (T_{ij}^{\theta}) \cdot (T_{ij}^{\beta}) \cdot (T_{ij}^{\gamma})$$

In the equation,  $T_{ij}$  denotes the level of the pheromone on the edge of (i, j) the main heuristic information. The pheromone level and heuristic information is controlled by the  $\theta$ ,  $\beta$ , and  $\gamma$ .

Each ants evaluates the CNN performance on the set of validation. The power of CNN that determines the quality of the solution is incorporated to finetune the trails of the pheromone based on the standard ACO. Moreover, the pheromone trail path is updated for each ant to empower the CNN model by selecting the optimal parameters. The pheromone update rule is calculated utilizing the below equation;

$$T_{ij} = (1 - p) \cdot T_{ij} + \frac{Q}{S}$$

In the equation (2),  $p$  is the pheromone vanishing rate, the amount of pheromone accumulated is  $Q$ , and the quality of solution is indicated as  $S$ . The modified ACO has developed to take into account the validation performance of diverse configurations of CNN that guarantee that pheromone trails are continuously updated depending on the performance of CNN on the validation set which is measured using the below equation;

$$T_{ij} = (1 - p) \cdot T_{ij} + \frac{Q}{V - S}$$

By updating the pheromone trails based on the performance of CNN model configurations on the validation set, the search is directed towards configurations that provide better classification accuracy. Consequently, the ACO serves as the powerful method for tuning the CNN hyperparameters that enable effective and functional way for text classification tasks. The chief comparison between the standard ACO and modified ACO presented in this study lies in the representation of candidate solutions and the pheromone matrix. In traditional ACO, the candidate solutions are characterized as a series of distinct decisions. Whilst, in the adapted ACO, the candidate solutions represent continuous configurations of the CNN model hyperparameters. This requires a different approach to constructing and updating the pheromone matrix, which guides the exploration of the solution space. The Pseudo code for the algorithm is provided below.

---



---

#### Algorithm 1 ACO-CNN Algorithm

---



---

##### Begin

**Initialize** ACO-CNN parameters:

m = number of ants  
k\_max = maximum iterations  
alpha = pheromone importance  
beta = distance importance  
rho = evaporation rate  
Q = pheromone deposit  
tau = initial pheromone matrix with value 0.1  
eta = 1/distance between parameter choices  
best\_score = infinity  
best\_solution = null

**Load** the training data, including features and labels.

**while** num\_iterations < k\_max do

**Calculate** edge attractiveness:

A = (tau<sup>alpha</sup>) \* (eta<sup>beta</sup>)

**Apply** pheromone evaporation:

tau = tau \* (1-rho)

**for** each ant in population do

**Generate** CNN parameters:

conv\_layers = random(1,

max\_conv\_layers)

filters = random\_choice(filter\_options)

kernel\_size =  
random\_choice(kernel\_size\_options)  
dense\_units =  
random\_choice(dense\_unit\_options)  
**Build** CNN architecture:  
**Add** Embedding layer  
**for** i = 1 to conv\_layers do  
**Add** Conv1D layer with filters,  
kernel\_size  
**Add** MaxPooling1D layer  
**end for**  
**Add** Flatten layer  
**Add** Dense layer with dense\_units  
**Add** Output layer  
**Train** CNN model:  
**Apply** early stopping  
**Train** on training data  
**Calculate** fitness:  
score = 1 -  
accuracy\_on\_validation\_set  
**if** score > best\_score then  
best\_solution = current\_parameters  
best\_score = score  
Update pheromone trails:  
**for** each parameter pair (i,j) do  
tau[i,j] = (1-rho)\*tau[i,j] + Q/score  
**end for**  
**end if**  
**if** time\_limit\_reached then  
**Extract** optimized parameters  
**Return** best\_solution, best\_score  
**end if**  
**end for**  
**end while**  
**Extract** the optimized parameters from  
best\_solution  
**Build** final CNN model with optimized  
parameters  
**Compile** CNN model with loss function,  
optimizer, metrics  
**Train** final model using complete training  
dataset  
**Evaluate** model's performance on test set  
**Return** optimized parameters and trained  
CNN model  
**End.**

### 3.3.3 Classifier Evaluation

The F1 measure and accuracy are key criteria for evaluating the outcomes of various methods. These algorithms operate based on two possible results for each test case, as shown in Table 3. This table presents the label "actual," which signifies the data prior to the application of the classifier, and "predicting," which reflects the data after the classifier has processed the test instances.

In this context, TP, FP, FN, and TN represent true positive, false positive, false negative, and true negative classifications, respectively. "True" indicates cases that have been correctly classified, while "false" refers to misclassified cases. Additionally, "positive" denotes terms identified within a specific class, whereas "negative" pertains to terms not associated with that class (Saeed et al., 2024).

Table 2: Confusion Matrix

Prediction/Actual	Positive	Negative
Positive	TP	FP
Negative	FN	TN

$$f1 - measure = \frac{2.precision \cdot recall}{precision+recall}$$

$$precision = \frac{TP}{TP+FP}$$

$$recall = \frac{TP}{TP+FN}$$

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

### 4. Experiments and Results

our experiments initiates with cleaning the news articles. The author performs the standard preprocess including normalization, standardization, stop words removal, punctuation elimination and etc. The author accomplished these by exploiting the KLPT library designed by Ahmedi 2020. Following this step, The author selected the popular base models that were employed for the same purpose to compare our proposed models against them. The base models included Support Vector Machine (SVM), Multinomial Naive Bayes (MNB), K-nearest

neighbors (KNN), and Decision Tree. The author replicated the same parameters and configuration as outlined in their studies. Moreover, F1-score, accuracy, precision and recall were selected as the main evaluation metrics.

The author conducted the experiment by leveraging Google Colab notebook that was designed by google for programmers to explore their programming codes. The notebook has both subscription and free based model. In this study, The author used the free based version. In the study, The author employed libraris such as TensorFlow, NumPy, Scikit-learn, and Pandas. The computer which was employed to perform evaluations equipped with an Intel Core i5 CPU, Windows 11, and 16 Gigabytes of RAM. Moreover, The author applied holdout techniques that employes 80:20 split for splitting data into train and test. For deep learning model, The author specified the number of epochs to two to avoid overfitting. The 5-fold cross-validation was implemented to calculate the average value for each evaluation metric. The results of our experimented are shown in Tables 3 and 4.

**Table 3:** The evaluation metrics of each model for the KurdFake dataset

Classifiers	F1-score	Accuracy	Recall	Precision
SVM	72.05	72.09	72.11	72.24
Random Forest	83.26	83.26	83.26	83.27
Naïve Bayes	66.49	66.72	66.66	66.09
Decision Tree	82.18	82.19	82.18	82.23
ACO-CNN	92.10	92.07	90.98	92.02

**Table 4.** The evaluation metrics of each model for the KDFND dataset

Classifiers	F1-score	Accuracy	Recall	Precision
SVM	80.09	79.96	79.94	80.43
Random Forest	81.17	80.36	80.36	81.6
Naïve Bayes	79.85	79.85	79.50	80.00
Decision Tree	73.82	73.82	73.52	74.00
ACO-CNN	83.47	83.24	83.13	82.97

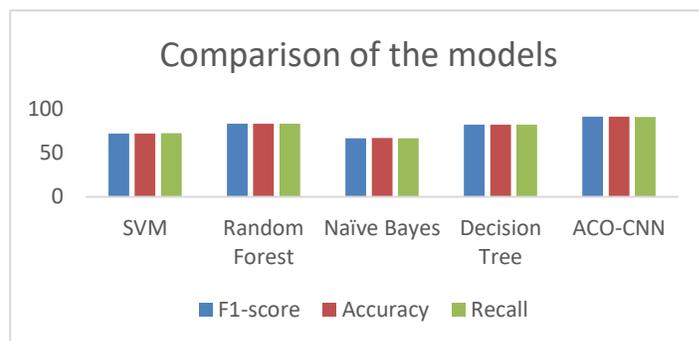
## 5. Discussion

In Table 2, numerous machine learning algorithms were experimented on the KurdFake dataset and evaluated against one another. It provides information about F1-score, accuracy, recall, and precision for each model.

Experimental results indicate that ACO-CNN method shows the highest performance across all of the metrics as displayed in Fig 3. The proposed method scored an F1-score of 92.10%, an accuracy of 92.07%, a recall of 90.08%, and a precision of 92.02%. These results denote that ACO-CNN triumphantly detects the fake news in the KurdFake dataset, that evidently proves that fusing ACO was successful in optimizing the CNN parameters to generate higher accuracy. In addition to its excellent performance, ACO-CNN can also be attributed to its hybrid architecture, which combines the features extraction capabilities of CNN with the optimization capabilities of ACO. As a result of the ACO algorithm, hyperparameters of the CNN, such as the learning rate, filter sizes, and network depth, are fine-tuned. Accordingly, an optimal representation of features for Kurdish text classification is achieved as a result of this optimization.

In contrast, the RF and DT that are popular for their success in managing non-linear relationships perform relatively well by scoring F1-scores of 83.26% and 82.18%, respectively. The SVM and Naïve Bayes method scored lower than the tree-based methods by achieving F1-scores of 72.05% and 66.49%, respectively. Considering the linguistic complexity of Kurdish text, Random Forest offers excellent performance. This effectiveness is likely due to the fact that it is an ensemble decision tree, which combines multiple decision trees in order to minimize overfitting and enhance generalization. Based on the similar performance of Random Forest and Decision Trees, it appears that tree-based methods are particularly well suited for capturing hierarchical patterns in the detection of Kurdish fake news. The reason for their suitability may be that they are capable of handling both categorical and numerical features derived from the preprocessing of text. This underperformance may be due to the simplistic assumptions made

by these models mainly the feature independent assumptions that causes to be less effective on the complex nature of the KurdFake dataset. SVMs are not as effective as they should be due to the curse of dimensionality in high-dimensional text feature spaces, as SVM struggles to identify optimal hyperplanes when dealing with sparse text representations. Additionally, the linear kernel used in SVM may not be able to effectively capture non-linear relationships found in Kurdish linguistic patterns. Moreover, the underperformance of Naive Bayes, despite there being a strong assumption of independence among features, emphasizes the interconnected nature of words and phrases in Kurdish fake news. As a result, contextual dependencies have a significant impact on classification accuracy in this context.



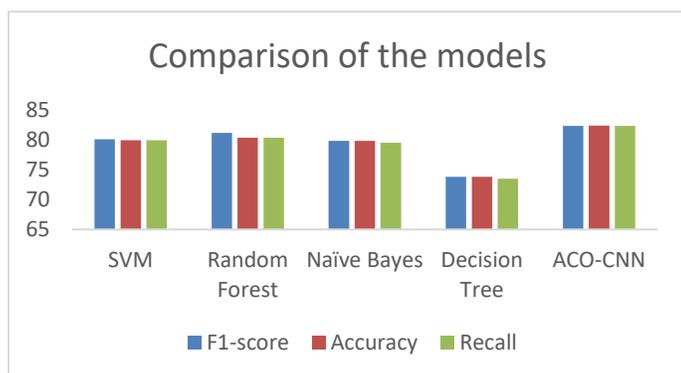
**Figure 3:** Comparison of the models on KurdFake dataset

Regarding the KDFND dataset, the ACO-CNN bested the other methods by achieving a highest F1-score of 83.47, a highest accuracy of 83.24, and a highest precision of 82.97, as shown in Fig 4. These results highlight the efficiency of ACO-CNN model in precisely classifying the news contents in the dataset. In comparison with KurdFake, KDFND achieved slightly lower F1-scores (82.37% vs. 91.07%), suggesting that the dataset presents more challenging cases or has linguistic features that complicate classification. This dataset contains more subtle forms of misinformation or feature a more balanced distribution of fake and real news, making the classification task inherently more complex.

It is worth noting that DT scored the lowest score in terms of F1-score, accuracy, recall, and precision against other methods. This shows that DT faced challenges in accurately identifying

authentic news from fabricated news in the KDFND dataset. On the KDFND dataset, the Decision Tree algorithm performed poorly (73.82% F1-score), whereas on the KurdFake dataset, it performed relatively well (82.18% F1-score). This indicates that single decision trees have a tendency to overfit, particularly when working with complex or varied datasets. Considering this performance gap, ensemble methods, such as Random Forest, can provide more consistent results between the two datasets.

Accordingly, the robustness of the proposed approach can be demonstrated by the consistent superiority of the ACO-CNN model across both datasets. Compared to the other datasets, F1-score differences of 8.7 percent (91.07% vs. 82.37%) indicate that the method is effective, but its performance depends on the characteristics of the datasets, including linguistic complexity, diversity of news domains, and degree of sophistication of the fake news patterns. Furthermore, ACO-CNN exhibits balanced performance across precision, recall, and F1-score metrics, indicating that it does not suffer from a high rate of false positives nor does it fail to detect instances of genuine fake news. It is essential to achieve a balance between recall and precision when deploying the technology, as failing to detect fake news (resulting in a low recall) and incorrectly flagging legitimate news (resulting in a low precision) could prove very detrimental to society.



**Figure 4:** Comparison of the models on KDFND dataset

In short, the outcomes state that the proposed ACO-CNN model can accurately detect fake news on both dataset, demonstrating its success in spotting instances of fake or genuine news in

the Kurdish language.

## Conclusions

The alarming increase of the dissemination of false news stands as the major threat that threatens the balance of well-educated community mainly in languages where NLP resources are scarcely short. Therefore, in this study, The study proposed ACO-CNN a hybrid deep learning models that can be incorporated for detecting and preventing the spread of the fake news in these under-resourced languages. The proposed model includes an optimized version of ACO fused with CNN for finetuning the CNN parameters and reduce the training time. The proposed method was evaluated against the state-of-the-art methods using two publicly available datasets in the language. The experimental results showed that our proposed method outperforms other models. Future research can study the functionality of ACO-CNN model on other low-resourced languages and suggest other improvements to the model.

**Acknowledgment:** Not applicable.

**Financial support:** No financial support.

**Potential conflicts of interest.** All authors report no conflicts of interest relevant to this article.

## References

- Ahmedi, S. KLPT – Kurdish Language Processing Toolkit. Proceedings of Second Workshop for NLP Open Source Software (NLP-OSS), 2020. Association for Computational Linguistics, 72-84.
- Ajao, O., Deepayan Bhowmik, Shahrzad Zargari. Fake news identification on twitter with hybrid cnn and rnn models. Proceedings of the 9th international conference on social media and society, 2018. 226-230.
- Amjada, M., Butta, S., Amjad, H. I., Zhilab, A., Sidorova, G. & Gelbukha. 2021. A. Overview of the Shared Task on Fake News Detection in Urdu at FIRE 2021. *Forum for Information Retrieval Evaluation*.
- Azad, R., Bilal Mohammed, Rawaz Mahmud, Lanya Zrar & Sdiq, S. 2021. Fake News Detection in Low-Resourced Languages 'Kurdish Language' Using Machine Learning Algorithms. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 4219-4225.
- Badawi, S., Saeed, A. M., Ahmed, S. A., Abdalla, P. A. & Hassan, D. A. 2023. Kurdish News Dataset Headlines (KNDH) through multiclass classification. *Data Brief*, 48, 109120.
- Badawi, S. S. 2024. Bridging the Gap: Enhancing Kurdish News Classification with RFO-CNN Hybrid Model. *Aro-the Scientific Journal of Koya University*, 12, 100-107.

- Chiman Haydar Salh & Ali, A. M. 2022. Comprehensive Study for Breast Cancer Using Deep Learning and Traditional Machine Learning. *Zanco Journal of Pure and Applied Sciences*, 34.
- Dana Mohammed Ali & A.Sadeq, H. 2022. Road Pothole Detection Using Unmanned Aerial Vehicle Imagery and Deep Learning Technique. *Zanco Journal of Pure and Applied Sciences*, 34, 107-115
- Dorigo, M., Birattari, M. & Stutzle, T. 2006. Ant colony optimization. *IEEE Computational Intelligence Magazine*, 1, 28-39.
- Fatima., B. N. & Djeflal, A. Fake news detection using machine learning. 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHS), 2020 Algeria. IEEE, 125-130.
- Haddad, H. A. 2023. Online Student Feedback System Using Machine Learning. *Zanco Journal of Pure and Applied Sciences*, 35, 78-85.
- Horne, B. & Adali, S. This Just In: Fake News Packs A Lot In Title, Uses Simpler, Repetitive Content in Text Body, More Similar To Satire Than Real News. Proceedings of the International AAAI Conference on Web and Social Media, 2017. 759-766.
- Ilyas, A., Shibani Santurkar, Dimitris Tsipras, Logan Engstrom, Brandon Tran, Aleksander Madry. Adversarial examples are not bugs, they are features. *Advances in neural information processing systems*, 2019.
- Jardaneh, G., Abdelhaq, H., Buzz, M. & Johnson, D. Classifying Arabic Tweets Based on Credibility Using Content and User Features. In 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 2019. IEEE, 596-601.
- Lapuschkin, S., Waldchen, S., Binder, A., Montavon, G., Samek, W. & Muller, K. R. 2019. Unmasking Clever Hans predictors and assessing what machines really learn. *Nat Commun*, 10, 1096.
- Mccoy, R. T., Ellie Pavlick, Tal Linzen. Right for the wrong reasons: Diagnosing syntactic heuristics in natural language inference. ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, 2020 Florence, Italy. 3428-3448.
- Ruchansky, N., Sungyong Seo, Yan Liu. Csi: A hybrid deep model for fake news detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2017.
- Saeed, A. M., Badawi, S., Ahmed, S. A. & A, D. 2024. Comparison of feature selection methods in Kurdish text classification. *Iran Journal of Computer Science*, 7, 55-64.
- Salh, D. & Nabi, R. 2023. Kurdish Fake News Detection Based on Machine Learning Approaches. *Passer Journal of Basic and Applied Sciences*, 5, 262-271.
- Sharma, U., Sidarth Saran, Shankar M. Patil. 2020. Fake news detection using machine learning algorithms. *International Journal of creative research thoughts (IJCRT)*, 8, 509-518.
- Shu, K., Sliva, A., Wang, S., Tang, J. & Liu, H. 2017. Fake News Detection on Social Media. *ACM SIGKDD Explorations Newsletter*, 19, 22-36.
- Thaher, T., Saheb, M., Turabieh, H. & Chantar, H. 2021. Intelligent Detection of False Information in Arabic Tweets Utilizing Hybrid Harris Hawks Based Feature Selection and Machine Learning Models. *Symmetry*, 13.
- Zhang, X. & Ghorbani, A. A. 2020. An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, 57.