Mapana – Journal of Sciences
2026, Vol. 25, No. 1, *137-156*
ISSN 0975-3303|https://doi.org/10.12723/mjs.76.9

QRMHF-DNK: A Hybrid Optimization and Deep Kernel approach for fake news detection

K. Saritha Devi* & M. Chidambaram*

Abstract

In this study, QRMHF-DNK (Quasi Reflection Metropolis Hasting Firefly- Deep Neural Kernel), a hybrid framework is proposed to enhance fake news detection on benchmark datasets. The framework integrates three major stages: data preprocessing to reduce sampling errors, feature selection using a swarm-based optimization strategy, and classification using a deep neural kernel model. This combination enables effective handling of high-dimensional textual data while accurately identifying informative features for classification. The proposed framework was evaluated on a publicly available Kaggle fake news dataset and compared with existing cooperative and multilingual deep learning methods. Experimental results show that QRMHF-DNK achieves a precision of 0.98 and a recall of 0.95, with a sampling error of 0.0671%, indicating that the sampled data closely represent the true class distribution. These results demonstrate the effectiveness of the proposed approach on the evaluated dataset and suggest its potential applicability to fake news detection tasks, while further validation on additional datasets is left for future work.

Keywords: Fake News Detection, Feature Selection, Swarm Intelligence, Deep Learning, Social Media Analysis.

1. Introduction

Fake news and false information have occurred long before the Internet, but the rise of digital media has made it easier for such content to spread quickly. To address this, a cooperative deep learning model was proposed in [1]. To address this growing problem, many fake news detection methods have been developed using NLP. However, most of these methods use very little human feedback in their systems. A Convolutional Neural Network

* Department of Computer Science, Rajah Serfoji Government College (Auto) (Affiliated to Bharathidasan University), Thanjavur, Tamil Nadu, India; pksdevi81@gmail.com

(CNN) converted user feedback into rankings, and the news items that were rated poorly were reused in training to make the model more accurate. Articles with low rankings were marked as fake, while higher ranked ones were considered reliable. A Capsule Neural Network was added to increase accuracy. To overcome this, a multilingual deep learning method was introduced in [2], using features like sentiment and named entities to analyze content more deeply. Even with these improvements, many systems could not fully capture the meaning and context of news written in different languages. Earlier models mainly focused only on text and ignored other types of data like images or videos. However, fake news often includes pictures and videos, which makes it more attractive and believable. Since fake news often includes both text and visuals, multimodal models have become very important. To tackle this, many studies have used Natural Language Processing (NLP), Machine Learning (ML), and Deep Learning (DL) approaches [6]. Some researchers have tried models with bidirectional encoders and multiple CNN layers [7], which improved classification accuracy. Although many of these models reach high accuracy, they often ignore precision or fail to reduce sampling errors. Other works, such as [9] and [10], reviewed and compared various deep learning and machine learning techniques for detecting fake new. The Multimodal Consistency Neural Network (MCNN), introduced in [8], is one example that considers both text and visual features to better identify fake content. Tests using several performance measures show that this model outperforms existing methods. It improves precision and reduces sampling errors through three stages: preprocessing with Stratum Variance Reduction, feature selection using the Quasi Reflection and Metropolis Hasting approach, and classification with a Deep Neural Kernel Perceptron that uses a Gudermannian Sigmoid function. To solve these problems, this paper presents a new method called the Quasi Reflection Metropolis Hasting Firefly and Deep Neural Kernel-based QRMHF-DNK model. Fake news can influence politics, society, and even public health. Many researchers have traveled deep learning approaches such as semi-supervised neural attention models [11], hybrid CNN RNN models [14], and multimodal compact bilinear pooling [15]. While these techniques have improved results, problems still remain in handling large, complex data and supporting multiple languages and content types. Therefore, there is a clear need for more flexible and efficient models that can accurately detect fake news across diverse datasets [12] [13] [17]. Others have studied sentiment-based classification [16], principal component analysis for dimensionality reduction [18], and ensemble deep learning [19] [20].

2. Methodology

Detecting fake news in real time is mostly difficult due to its dynamic nature. To address this, we propose a novel method called Quasi Reflection

Metropolis Hasting Firefly and Deep Neural Kernel-based QRMHF-DNK classification. With the rapid growth of the Internet, social media has become a major platform for the spread of fake news, posing major challenges to both academic circles and industry. The method aims to accurately predict the authenticity of news articles by employing a three-stage approach: preprocessing using Stratum Variance Reduction to minimize sampling error, optimal feature selection using a Quasi Reflection Metropolis Hastings Firefly algorithm, and final classification through a Deep Neural Kernel Perceptron. Further details of the QRMHF-DNK framework are presented in the following sections.

2.1 Dataset description

The data set employed for this analysis was obtained from the Kaggle competition entitled "Fake News Detection" and is referred to as the data set (train.csv). It comprises 20,800 articles of news that are classified based on their credibility that is, trusted or untrusted (label 0 and label 1 respectively). The dataset is almost balanced, with an equal measure of news set to 0 and 1, thus avoiding bias during model construction. For every news article, five attributes are provided: id, which is a distinctive identifier; title, which signifies the news title; author, which shows the author name; text, which comprises the complete news information; and finally, label that defines the authenticity of news. The existence of both textual and metadata attributes makes the collection appropriate for judging the fake news identification models. For testing purposes, this dataset is divided into three components: the training set containing labeled data for modeling, testing set with unseen labeled data, and sample submission file for judging the format of predictions. Due to its large size, well-balanced classes, and detailed textual data, this dataset is commonly used as a trusted benchmark for research in fake news classification.

2.2. Stratum Variance Reduction based Preprocessing

After data collection, the input features from the raw dataset are uploaded to verify and validate their suitability for analysis. In this phase, news content encompassing both genuine and fake articles is collected. For the present study, the dataset was sourced from Kaggle. This approach divides the dataset into different strata based on four key features ID, title, author, and text. Prior to being fed into deep learning models, the dataset undergoes text preprocessing, which is essential for extracting the most relevant terms and enhancing model performance. A Stratum Variance Reduction SVR based preprocessing technique is employed to improve sampling precision by minimizing sampling error. Figure 1 shows the structure of the SVR based preprocessing model. Each stratum is randomly sampled to account for variations within the dataset, the error in detecting unpredictable news articles is reduced. Since strata with lower variance contribute to greater accuracy, this stratified method increases the whole consistency of the model.

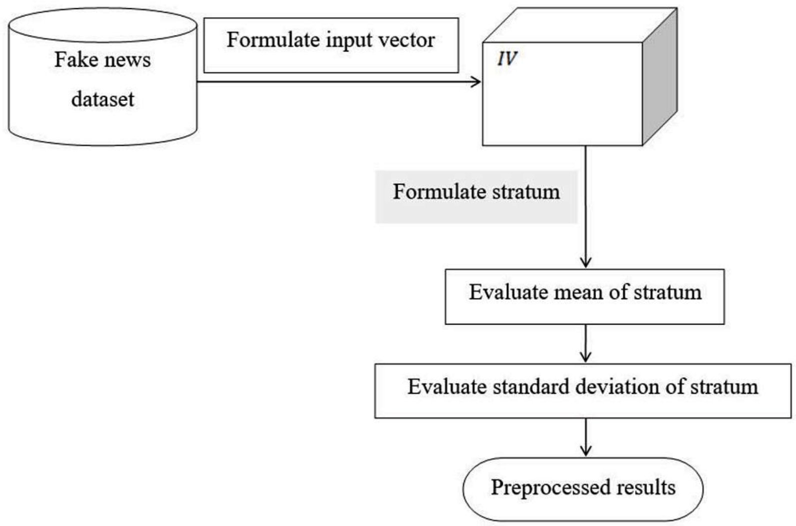

Figure 1: Structure of Stratum Variance Reduction based Preprocessing

As demonstrated in the above figure with the raw dataset acquired as input, the input vector matrix is formulated as given below.

$$IV = \begin{bmatrix} S_1F_1 & S_1F_2 & \dots & S_1F_n \\ S_2F_1 & S_2F_2 & \dots & S_2F_n \\ \dots & \dots & \dots & \dots \\ S_mF_1 & S_mF_2 & \dots & S_mF_n \end{bmatrix} \quad (1)$$

From equation (1), let 'n' represent the number of samples and 'm' denote the number of features extracted from the input dataset 'D' for raw data processing. Based on the formulated input vector, the mean of the Stratum Variance Reduction (SVR)–based sampling is expressed as follows:

$$S' = \frac{1}{N} \sum_{i=1}^{m} SSize_i S_i' \quad (2)$$

From the above equation (2) the mean of Stratum Variance Reduction based sampling 'S' for the given raw dataset 'DS' is obtained based on the sum of all stratum or sample sizes 'N', size of stratum or sample '$SSize_i$' and the sample mean of stratum or sample 'S_i'' respectively. Following which the variance of Stratum Variance Reduction based sampling is mathematically formulated as given below.

$$SD_{S'}^2 = \sum_{i=1}^{m} \left(\frac{SSize_i}{N}\right)^2 \left(\frac{SSize_i - Obs_i}{SSize_i - 1}\right) \frac{S_i^2}{Obs_i} \quad (3)$$

From the above equation (3) the standard deviation of Stratum Variance Reduction based sampling '$SD_{S'}^2$' for the given raw dataset 'DS' is obtained by taking into consideration the sum of all sample or stratum sizes 'N', size of sample or stratum 'i', number of observations '$Obsi$' in stratum 'i' and sample standard deviation of sample or stratum 'i' respectively.

$$PS = S'.SD_{S'}^2 \quad (4)$$

The above preprocessed sample 'PS' results (4) are obtained with less variability, therefore reducing sampling error considerably. The pseudo-code representation of Stratum Variance Reduction-based Preprocessing is given below.

Input: Dataset 'S', Samples '$S = \{S_1, S_2, \ldots, S_m\}$', Features '$F=\{F_1, F_2, \ldots, F_n\}$'
Output: Sampling error minimized preprocessed results 'PS'
1: **Initialize** 'm', 'n' 2: **Begin** 3: **For** each Dataset 'DS' with Samples 'S' and Features 'F' 4: Formulate the input vector matrix as given in equation (1) 5: Evaluate the mean of the Stratum Variance Reduction based sampling as given in equation (2) 6: Evaluate the standard deviation of the Stratum Variance Reduction based sampling as given in equation (3) 7: **Return** preprocessed sample results as given in equation (4) 8: **End for** 9: **End**

Algorithm 1: Stratum Variance Reduction based Preprocessing

This approach effectively produces sampled results with minimal sampling error. A stratum variation reduction function is applied to each stratum independently, enhancing sample precision and significantly reducing detection time. To minimize sampling errors in fake news detection, the input vector matrix derived from the raw dataset is subjected to stratified sampling.

2.3. Quasi Reflection Metropolis Hasting Firefly Optimal Feature selection model

A best feature subset improves classification performance, mostly in high-dimensional feature spaces. Feature selection plays a critical role in designing effective classification models for fake news detection using data-driven learning methods. Employing a well-organized feature selection model helps identify the most appropriate features while eliminating redundant ones. However, conventional approaches often suffer from premature convergence, leading to suboptimal feature selection. To overcome this limitation, the proposed model integrates a Quasi Reflection Function and Metropolis Hasting Attractiveness Mechanism, which together balance local and global exploration to identify the most discriminative features. Quasi Reflection is incorporated into the feature selection process to improve exploration efficiency in high-dimensional search spaces. Fake news datasets typically contain large numbers of redundant and noisy textual features, which can negatively affect classification performance. By generating quasi-reflected solutions alongside current candidates, the optimization process is better able to explore promising regions of the search spaces, leading to faster convergence and higher-quality feature selection. Figure 2 illustrates the structure of the proposed Quasi Reflection Metropolis Hasting Firefly Optimal Feature Selection Model.

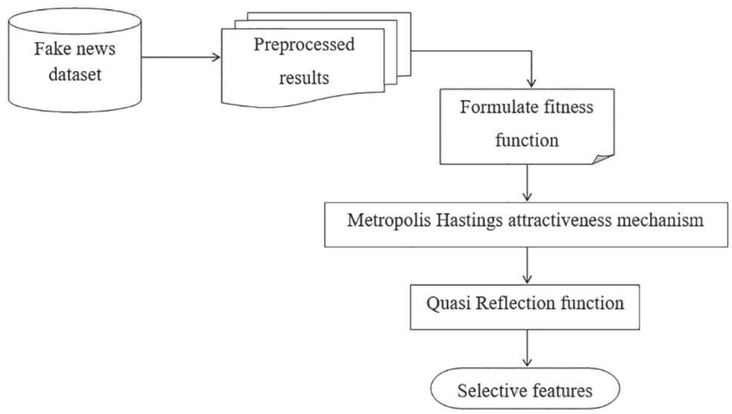

Figure 2: Structure of Quasi Reflection Metropolis Hasting Firefly Optimal Feature selection model

As shown in the figure, the Quasi Reflection Function and Metropolis Hasting Attractiveness Mechanism operate cooperatively to overcome premature convergence and achieve global optimality, thereby ensuring the selection of the most discriminative features. The objective function used to evaluate the fitness of each feature subset is mathematically formulated as follows.

$$fitness(S) = W_1 * Acc_i + W_2 * (F_i)^{-1} \tag{5}$$

From the above equation (5), 'W_1' represents the weight of fake news detection accuracy 'Acc_i', whereas 'W_2' represents the weight of the feature 'F_i' subset respectively with '$W_1 + W_2 = 1$'. Owing to the reason that fake news detection accuracy is more significant than the features being selected, 'W_1' is assigned as '0.9' and 'W_2' is assigned as '0.1'. The proposed Quasi Reflection Metropolis Hasting Firefly Optimal Feature selection model employs the Metropolis Hasting for population initialization to increase diversification in the swarm population (i.e., overall sample space).

After ranking all fireflies, i.e., features based on their fitness values, the global best solution is identified using the Metropolis Hasting mechanism. This secondary leader exhibits a comparable fitness value with minimal correlation to the global best. To enhance search diversity and prevent entrapment in local optima, a second swarm leader, i.e., a secondary feature is determined using the Quasi Reflection Function. Furthermore, the optimal offspring, derived from the mean position of the two leaders and their neighboring features, guides the proposed attractiveness search, enabling fireflies with lower light intensities to move toward optimal solutions. The Metropolis Hasting (MH) Attractiveness Parameter, governed by the Acceptance Ratio, is mathematically defined as follows. Since both swarm leaders explore distinct search regions, the likelihood of premature convergence is significantly reduced.

$$S_i = S_i + AR(S_j' - S_i) + AR(g_{best}' - S_i) + \frac{S}{Max\_Iter} Sign\left[AN - \frac{1}{2}\right] \tag{6}$$

$$AR = f(S') \big/ f(S_{iter}) \tag{7}$$

$$S_j' = S_j + \lambda_1 \tag{8}$$

$$g_{best}' = \mu(g_{best} - S_{best}) + \lambda_2 \tag{9}$$

From the above equation (6), 'S_j'' represents the brighter neighboring firefly or samples identified by MH as given in equation defined in Equation (8), whereas 'g'_{best}' denotes mean of two swarm leaders (i.e., two consecutive features) occasioned by MH as defined in equation (9). Moreover '*AR*' denotes the acceptance ratio of desired probability function '$f(S')$' for the corresponding sample from a candidate sample '$f(S')$' with respect to iteration '$f(Siter)$'. The acceptance ratio aids in deciding whether to accept or reject the candidate feature.

In addition '$\frac{S}{Max\_Iter}$' is employed as an adaptive parameter with '*Max_Iter*' representing the maximum number of iterations and 'λ_1, λ_2' the random search strategy. Due to its ability to accelerate convergence, the Metropolis Hasting (MH) mechanism is employed to fine-tune the attractiveness parameters. In particular, it adaptively updates both the local and global parameter coefficients, as represented in Equation (6). Each newly generated solution from the MH process is compared with the existing solutions, and the best solution is selected to replace the current firefly, thereby improving overall search efficiency and convergence stability. Finally, the Quasi Reflected feature 'j' of sample 'S' (S_j^{QR}) to overcome local optimal functionality is mathematically stated as given on the next page.

$$FS = S_i^{QR} = RND\left(\left[\frac{LB_j + UB_j}{2}\right], S_i\right) \tag{10}$$

From the above equation (10) the most selective features selected 'FS' are obtained based on the lower 'LB_j' and upper 'UB_j' bound of the '$j-th_j$' sample, mean '$\frac{LB_j+UB_j}{2}$' of the interval '$[LB_j, UB_j]$' whereas '$RND\left(\left[\frac{LB_j+UB_j}{2}\right], S_i\right)$' creates the pseudo-random . The pseudo-code representation of Quasi Reflection Metropolis Hasting Firefly Optimal feature selection is given below.

Input: Dataset '*DS*', Samples '$S = \{S_1, S_2, ..., S_m\}$', Features '$F = \{F_1, F_2, ..., F_n\}$'
Output: accurate and precise selective features '*FS*'
1: **Initialize** preprocessed results '*PS*', 'm', 'n', '$W_1 = 0.9$', '$W_2 = 0.1$' 2: **Begin** 3: **For** each Dataset '*DS*' with preprocessed sample results '*PS*' and Features '*F*' 4: Formulate fitness function as given in equation (5) 5: Measure Metropolis Hasting (MH) Attractiveness Parameters guided by Acceptance Ratio as given in equations (6), (7), (8), and (9) 6: Obtain Quasi Reflected feature 'j' of sample '*S*' as given in equation (10) 7: **Return** most selective features '*F S*' 8: **End for** 9: **End**

Algorithm 2: Quasi Reflection Metropolis Hasting Firefly Optimal feature selection

As described in the algorithm, the primary objective is to enhance the true positive rate, that is, the probability of correctly identifying fake news as fake. To mitigate premature convergence, population diversity is maintained so that, even with a limited number of features in the raw dataset, diversification ensures optimal feature selection. Feature selection plays a vital role in achieving this goal. Subsequently, the QuasiReflection function is employed to overcome local optima, prevent premature convergence, and further effectively enhance the true positive rate. Initially, the Metropolis Hasting function is applied to generate a sequence of sample values whose distribution progressively approximates the desired probability distribution, thereby improving the detection of actual positive cases.

2.4. Deep Neural Kernel Perceptron-based Classifier for fake news detection

However, these existing approaches often look limited in terms of precision and recall, highlighting the need for a more robust detection mechanism. In recent years, widespread research hasled to the development of well-organized structures for improving detection accuracy using machine learning and deep learning methods. Several deep learning approaches, such as Cooperative Deep Learning 1 and Capsule Neural Networks 2, have travelled for fake news detection. Feature classification extends conventional classification methods by leveraging information from grouped samples referred to as fields or features. To address this, the present study assumes that each feature represents a homogeneous group containing consistent news samples. Under the assumption of independent and identically distributed (i.i.d.) samples, conventional models tend to overlook the inherent uniformity within these feature groups, which restricts their detection capability. Accordingly, this work proposes a Deep Neural Kernel Perceptron based Classifier to enhance both precision and recall. The model integrates News Kernel Perceptron and Classifier Kernel Perceptron modules where the former learns from specific feature groups and the latter performs classification based on the learned representations. This architecture ensures improved accuracy and robustness in fake news detection. Figure 3 illustrates the structure of the proposed Deep Neural Kernel Perceptron based Classification Model.

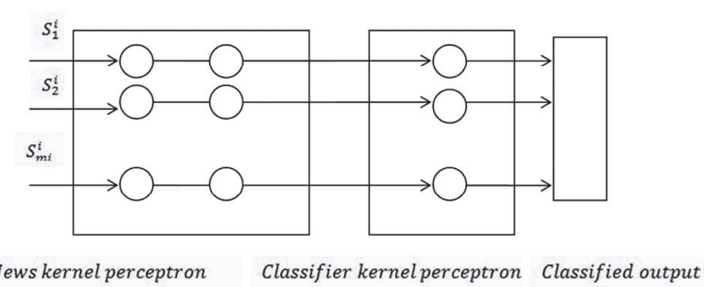

Figure 3: Structure of Deep Neural Kernel Perceptron-based Classification model

As shown in the figure, the sample news articles are provided as input to the model. Given a preprocessed dataset containing the most significant selected features, a set of n news perceptrons is established, where each perceptron matches to a specific feature sample. Following the News Kernel Perceptron concept, a separate perceptron is used to represent each feature of the news data accurately. A sigmoid activation function is working to activate each news perceptron. After applying the News Kernel transformation, a News Invariant Classifier Kernel Perceptron is trained on the transformed samples and subsequently activated using the sigmoid function, ensuring nonlinear mapping and improved classification accuracy. Let us represent a group of patterns or sample news articles as '$f_i = \{S_1^i, S_2^i, \dots, S_{mi}^i\}$' and the corresponding labels as '$C_i = \{O_1^i, O_2^i, \dots, O_{mi}^i\}$'.

If all patterns or sample news articles in 'f_i' share the same news article with different content, we define 'f_i' as a feature-sample with feature sample length 'mi' and 'C_i' is the feature class. Then 'L' field sample news articles '$\{f_i, C_i\}_{i=1}^{L}$' are given to train using News Kernel Perceptron for the corresponding labels. Then a perceptron for learning a binary classifier (i.e., fake news or real news) that maps input 'x' to an output value 'y' is formulated as given on the next page.

$$f(x \in S) = H(W.x + b) \tag{11}$$

From the above equation (11) the perceptron is modeled using the unit step function 'H' based on the weights 'W' and bias 'b' subjected to input samples 'x'. Then, the output of News Kernel Perceptron for the analogous labels is then mathematically formulated as given below.

$$y^n = f^{CK}(f_i^{NK}(S; W_i^{NK}); W^{CK}) \tag{12}$$

From the above equation (12), 'f_i^{NK}' denotes the news kernel perceptron for feature 'i' and 'f^{CK}' represents the classifier kernel perceptron for the corresponding samples 'S' and '$X = \{S_i\}_{i=1}^{L}$', '$Y = \{O_i\}_{i=1}^{L}$'. Also 'W_i^{NK}' and 'W^{CK}' denotes the weight matrix for news kernel and classifier kernel perceptron respectively. We further assume that each distinct feature is derived from the corresponding sample news article distribution, as data within the same feature group share common characteristics. Accordingly, the News Kernel Perceptron is formulated based on a Bayesian framework, which captures the probabilistic relationship between features and news article distributions. The mathematical model of the News Kernel Perceptron based on this Bayesian structure is expressed as follows.

$$Prob(W^{NK}, W^{CK}|X, Y) = Prob(W^{NK}, W^{CK}) * Prob(Y|X, W^{CK}, W^{NK}) \tag{13}$$

$$Prob(W^{NK}, W^{CK}) * \prod_{i=1}^{L} \prod_{m=1}^{mi} Prob\left(y_m^i | x_m^i, W^{CK}, W^{NK}\right) \tag{14}$$

From the above equations (13) and (14), 'W^{CK}', 'W^{NK}' represents the weight for the classifier kernel and news kernel of corresponding sample 'i' and '$\left\{x_m^i, y_m^i\right\}$' denotes the sample to be simulated for fake news detection. Then, the weight matrix of the news kernel of corresponding sample 'i' is formulated as given below.

$$W^{NK} = \begin{bmatrix} w_{11}^{NK} & w_{12}^{NK} & .. & w_{1n}^{NK} \\ w_{21}^{NK} & w_{22}^{NK} & .. & w_{2n}^{NK} \\ .. & .. & .. & .. \\ w_{L1}^{NK} & w_{L2}^{NK} & .. & w_{Ln}^{NK} \end{bmatrix} \tag{15}$$

Finally, the classifier output, referred to as the Class Kernel CK, is obtained using the Gudermannian Sigmoid Activation Function. This activation function is used for fake news detection because it establishes a smooth and continuous mapping between the title represented as a point on a hyperbola and the text content represented as a point on a semicircle. Through this nonlinear mapping, meaningful patterns specifically, unreliable or fake news articles are effectively recognized, thereby enhancing both the accuracy and robustness of the detection model. The Gudermannian Sigmoid Activation Function used for fake news detection is mathematically expressed as follows.

$$f(y^n) = gd(y) = \int_0^y \frac{dt}{\cosh t} = 2\arctan\left(\tanh\left(\frac{y}{2}\right)\right) \tag{16}$$

The results from Equation (16) help identify fake news articles in the preprocessed dataset using the most important extracted features. The overall process is explained using the pseudo-code of the Deep Neural Kernel Perceptron–based classification model shown below.

Input: Dataset '*DS*', Samples '*S* = {S_1, S_2,..., S_m}', Features '*F*={F_1, F_2,..., F_n}'
Output: precise and accurate classified results
1: **Initialize** preprocessed results '*PS*', '*m*', '*n*', selective features '*FS*' 2: **Begin** 3: **For** each Dataset '*DS*' with preprocessed sample results '*PS*' and selective features '*FS*' 4: Formulate perceptron function as given in equation (11) 5: Formulate output of News Kernel Perceptron for the analogous labels as given in equation (12) 6: Model News Kernel Perceptron based on Bayesian structure as given in equations (13) and (14) 7: Formulate weight matrix as given in equation (15) 8: Formulate output using Gudermannian Sigmoid Activation function as given in equation (16) 9: **If** '$f(y^n) = \frac{1}{2}\Pi$' 10: **Then** the samples are reliable news articles 11**: Return** results as reliable news articles 12: Return to Step 4 13: **End if** 14: **If** '$f(y^n) = -\frac{1}{2}\Pi$' 15: **Then** the samples are fake news articles 16: **Return** results as fake news articles 17: **Go to** step 4 18: **End if** 19: **End for** 20: **End**

Algorithm 3: Deep Neural Kernel Perceptron-based Classification

As described in the algorithm, to enhance precision and recall, a Deep Neural Kernel Perceptron based Classifier is utilized. Using the preprocessed samples with the most significant features as input, the model is divided into two components the News Kernel and the Classifier Kernel, each processed through perceptrons for effective classification. In the News Kernel Perceptron, a Bayesian framework is employed to model a set of text variables with conditional dependencies linked to the news title, generating probabilistic outcomes that contribute to improved precision. Subsequently, the Classifier Kernel is activated using the Gudermannian Sigmoid Function which maps the feature variables from a hyperbolic to a circular domain-based on the title text relationship. The Gudermannian Sigmoid activation function is employed in the classifier due to its smooth transition between linear and non-linear behavior without sharp saturation. This property helps maintain stable gradients during training and improves sensitivity to subtle linguistic patterns and semantic inconsistencies often present in fake news content. As a result, the classifier achieves better discrimination between reliable and unreliable news articles. This transformation enables accurate classification of fake and reliable news articles, thereby significantly improving the overall recall performance.

3. Experimental setup

The performance of the suggested method was compared with existing models, namely Cooperative Deep Learning [1] and Multilingual Deep Learning [2]. Evaluation was conducted using key performance metrics such as precision, recall, sampling error, and true positive rate.

The proposed fake news detection approach, termed the Quasi Reflection Metropolis Hasting Firefly and Deep Neural Kernel-based QRMHF-DNK classification model, was implemented and evaluated using Python on a Microsoft Windows platform furnished with an Intel i3 2350 processor and 2 GB RAM. The comparative analysis demonstrates the performance efficiency of the QRMHF-DNK model over the existing methods. For experimentation, a maximum of 20,000 samples were utilized, with 500 files and a maximum file size of 800 KB considered.

4. Implementation details

The method is evaluated using a fake news dataset and compared against two existing approaches, Cooperative Deep Learning [1] and Multilingual Deep Learning [2], to validate its performance. In this study, a novel fake news detection model named Quasi Reflection Metropolis Hasting Firefly and Deep Neural Kernel-based QRMHF-DNK classification is proposed to achieve higher precision, improved recall, and minimal sampling error. The proposed QRMHF-DNK framework consists of three main modules: preprocessing, feature selection, and classification. Initially, raw data comprising four key features is extracted from the input dataset. Using this approach, the mean and standard deviation of each sample stratum are computed, resulting in preprocessed samples with minimized sampling error. In the preprocessing phase, the Stratum Variance Reduction based Preprocessing Algorithm is employed to achieve balanced and optimal sampling. Finally, the Deep Neural Kernel Perceptron based Classifier is applied to the selected features to significantly improve both precision and recall in fake news detection. This process integrates the Metropolis Hasting (MH) Attractiveness Parameters, guided by the Acceptance Ratio, along with the Quasi Reflection Function, to improve the true positive rate and precision. In the feature selection phase, the Quasi Reflection Metropolis Hasting Firefly Optimal Feature Selection Algorithm is applied to identify the most discriminative features. Based on these implementation steps, the performance evaluation of the proposed method is conducted using four key metrics, which are discussed in the following section.

5. Discussion

In this section, the performance of the proposed Quasi Reflection Metropolis Hasting Firefly and Deep Neural Kernel based QRMHF-DNK classification method is validated and analyzed through a comparative study with two state-of-the-art techniques Cooperative Deep Learning [1] and Multilingual Deep Learning [2]. To ensure a fair and unbiased comparison, the same dataset with identical sample records is employed for all models, allowing for a consistent and detailed evaluation of performance across the different approaches.

5.1. Performance analysis of precision

In fake news detection, precision and recall serve as key performance metrics for evaluating the accuracy of classification based on the sampled news articles. The precision rate is mathematically expressed as follows. Precision measures the fraction of correctly identified fake news articles out of all articles predicted as fake, indicating the reliability of positive predictions made by the model.

$$Precision = \frac{TP}{TP+FP} \tag{17}$$

Table 1, given below, lists the precision rate obtained using the three methods: QRMHF-DNK, cooperative deep learning [1] and multilingual deep learning [2].

Table 1: Tabulation of precision

Samples	Precision		
2000	QRMHF-DNK	cooperative deep learning	multilingual deep learning
4000	0.98	0.97	0.96
6000	0.96	0.95	0.93
8000	0.95	0.93	0.9
10000	0.94	0.92	0.89
12000	0.95	0.93	0.9
14000	0.96	0.94	0.91
16000	0.95	0.92	0.87
18000	0.96	0.93	0.89
20000	0.97	0.94	0.9

The performance evaluation of precision with respect to the number of distinct samples is presented in Table 1. For instance, with 2000 samples considered in the simulation of which 1980 were reliable news articles the precision rate achieved by the QRMHF-DNK method was 0.98, while the Cooperative Deep Learning [1] and Multilingual Deep Learning [2] methods achieved precision rates of 0.97 and 0.96, respectively. When compared with the existing methods, the proposed QRMHF-DNK model demonstrates a notable improvement in precision by accurately identifying unreliable news articles. Overall, the QRMHF DNK method attained an improvement in precision of approximately 2% and 6% compared to the existing approaches. Across multiple model runs, similar patterns were observed, confirming the greater performance of the proposed model.

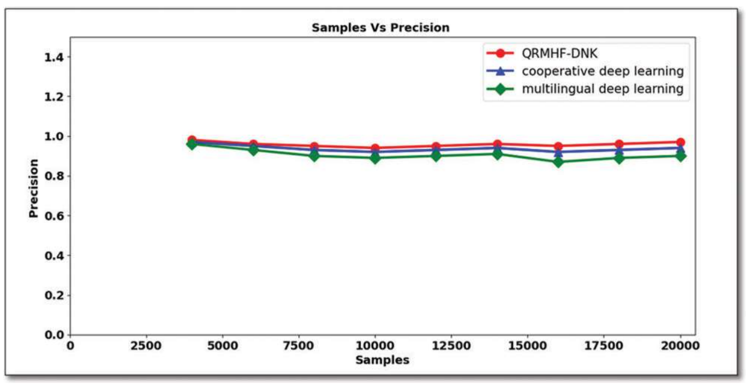

Figure 4: Comparison of precision using QRMHF-DNK, cooperative deep learning [1] and multilingual deep learning [2]

Figure 4 illustrates the graphical comparison of precision rates obtained using the proposed QRMHFDNK method and the existing Cooperative Deep Learning [1] and Multilingual Deep Learning [2] techniques. However, the QRMHF-DNK method consistently achieves a higher precision rate compared to the other two methods. From the figure, it is observed that the precision rate remains relatively stable with an increase in the number of samples, showing neither a consistent rise nor decline.

This improvement is attributed to the design of the QRMHF-DNK framework, where preprocessed samples and the most important features are used as inputs. The deep neural model divides the data into News Kernel and Classifier Kernel, with the News Kernel further processed through a perceptron for classification. Within this perceptron, a Bayesian structure is employed to establish conditional dependencies between textual features and titles, generating probabilistic outcomes that enhance precision. Consequently, the proposed method demonstrates superior precision performance, and the graph further confirms that increasing the number of samples does not compromise precision stability.

5.2. Performance analysis of recall

In this section, the performance metric recall is evaluated. While precision measures the quality of fake news detection by assessing how many identified fake news articles are actually fake, recall measures the quantity, indicating how effectively the model detects all actual fake news instances within the dataset. The recall rate is mathematically expressed as follows

$$Recall = \frac{TP}{TP+FN} \tag{18}$$

Table 2, given below, lists the recall rate evaluated with the help of the three methods, QRMHF-DNK, cooperative deep learning [1], and multilingual deep learning [2]. The results of recall rate by applying equation (18) are given in table 2.

Table 2: Tabulation of recall

Samples	Recall		
2000	QRMHF-DNK	cooperative deep learning	multilingual deep learning
4000	0.95	0.9	0.8
6000	0.94	0.88	0.78
8000	0.92	0.87	0.75
10000	0.9	0.85	0.73
12000	0.88	0.83	0.72
14000	0.85	0.81	0.7
16000	0.86	0.82	0.75
18000	0.88	0.83	0.78
20000	0.93	0.85	0.81

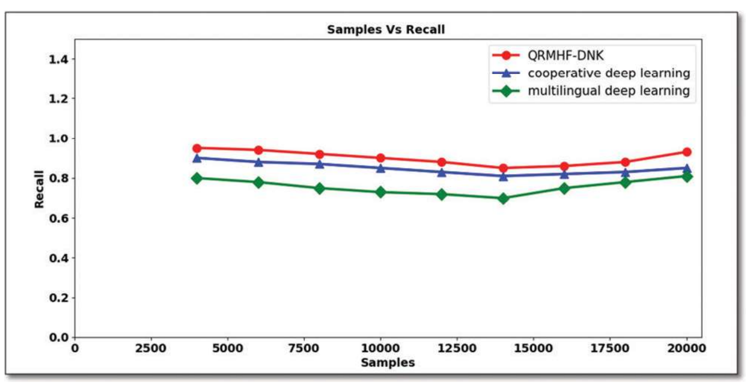

Figure 5: Comparison of recall using QRMHF-DNK, cooperative deep learning [1] and multilingual deep learning [2]

Figure 5 shows the graphical representation of the recall rate with respect to 20,000 different samples collected across 10 iterations. As shown, the recall rate remains steady regardless of the increase in the number of models, approving that a larger sample size does not compromise the overall recall performance. This reliability also gives an improvement in the true positive rate for fake news detection. For example, in a simulation involving 2000 samples, where 20 were fake news articles, the QRMHF-DNK method correctly detected 19, while Cooperative Deep Learning [1] and Multilingual Deep Learning [2] identified 18 and 16, respectively. Similarly, the recall rates were 0.95 for QRMHF-DNK, 0.90 for [1], and 0.80 for [2]. These results clearly show that the offered method achieves a higher recall rate compared to existing approaches. The improvement in recall using QRMHF-DNK is attributed to the application of the Deep Neural Kernel Perceptron-based Classification Algorithm. In this model, the Classifier Kernel is activated using the Gudermannian Sigmoid Function, which maps text representations from a hyperbolic angle (news content) to a circular angle (news title), allowing for more accurate discrimination between fake and consistent articles. As a result, the overall recall rate achieved by the QRMHF-DNK method improves by approximately 5% compared to [1] and 15% compared to [2].

5.3. Performance analysis of sampling error

In fake news detection, sampling errors arise when statistical characteristics of a population are estimated based on only a subset of that population. Meanwhile the selected sample may not fully represent all the structures or occurrences present in the overall dataset, the calculated means, variances, or quartiles may differ from the true population values. This deviation between the sample statistic and the population factor is defined as the sampling error. The sampling error can be mathematically expressed using the following formula.

$$SE = Z * \sqrt{(p * (1-p)/n)} * (1 - \sqrt{(n/N)}) \tag{19}$$

where N = population size (total number of articles in the dataset), n = sample size used for training or evaluation, p = proportion of the positive class (e.g., unreliable news) in the sample, Z = Z-score corresponding to the 95% confidence level (1.96).

Table 3, given below, lists the sampling error measured using the three methods, QRMHF-DNK, cooperative deep learning [1] and multilingual deep learning [2].

Table 3: Tabulation for sampling error

Particulars	QRMHF-DNK	cooperative deep learning [1]	multilingual deep learning [2]
Population size (N)	25000	25000	25000
Sample size (n)	20000	20000	20000

Proportion of sample who said yes (p)	70%	65%	61%
Z-score for 95% confidence level (Z)	1.96	1.96	1.96
Sampling error	0.000671	0.000698	0.000714
Sampling error (%)	0.0671	0.0698	0.0714

As shown in the above table, a total of 20,000 news articles were sampled from a fake news dataset representing an overall population of 25,000 individuals surveyed to determine whether the news items were fake or real. Among these, 70% of the sample responded positively indicating real news using the proposed QRMHF-DNK method, while 65 and 61 responded positively using Cooperative Deep Learning [1] and Multilingual Deep Learning[2], respectively. The sampling error was computed at a 95% confidence level for all three methods. For the dataset used in this study, this computation resulted in a **sampling error of 0.0671%**, indicating that the sampled training data closely represent the true class distribution of the full dataset. Similar calculations for baseline models yielded 0.0698% and 0.0714%, confirming that the proposed preprocessing and feature selection strategies maintain high representativeness. This process ensures balanced sampling and significantly reduces the overall sampling error.

5.4. Performance analysis of true positive rate

In this section, the True Positive Rate TPR is evaluated. It measures the model's ability to accurately recognize genuine news articles and is a key indicator of its effectiveness in distinguishing between real and fake news. The true positive rate represents the probability that an actual positive case in this context, a reliable news article is correctly identified as positive by the detection model. The true positive rate is mathematically expressed as follows.

$$TPR = \frac{TP}{TP+FN} \tag{20}$$

Table 4, given below, provides the true positive rate results obtained using the three methods.

Table 4: Tabulation for true positive rate

Samples	True positive rate		
2000	QRMHF-DNK	cooperative deep learning	multilingual deep learning
4000	0.98	0.97	0.96
6000	0.96	0.94	0.92
8000	0.95	0.92	0.89
10000	0.93	0.9	0.86

12000	0.91	0.88	0.85
14000	0.9	0.86	0.83
16000	0.9	0.86	0.83
18000	0.93	0.88	0.84
20000	0.94	0.89	0.85

Table 4 gives us the data how the True Positive Rate TPR changes for different sample sizes, extending from 2,000 to 20,000. The results clearly indicate that the proposed QRMHF-DNK method achieves a superior true positive rate compared to the existing models. For instance, with 2,000 samples, the true positive rate obtained using QRMHF-DNK was 0.98, while Cooperative Deep Learning [1] and Multilingual Deep Learning [2] achieved 0.97 and 0.96, respectively. Based on the average results from these simulations, it is observed that the proposed QRMHF DNK method improves the true positive rate by approximately 4% compared to [1] and 7% compared to [2]. This enhancement demonstrates the robustness and reliability of the QRMHF DNK model in accurately identifying genuine news articles. To ensure a fair comparison, an overall simulation consisting of 10 iterations was conducted using distinct sets of samples for all three methods.

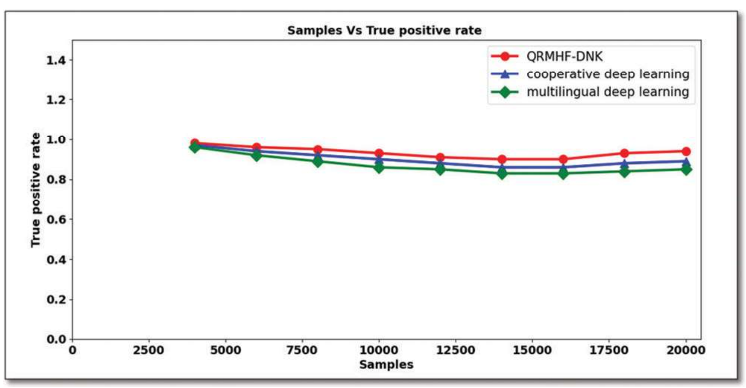

Figure 6: Comparison of true positive rate using QRMHF-DNK, cooperative deep learning [1] and multilingual deep learning [2]

Figure 6 illustrates the graphical representation of the True Positive Rate TPR for 20,000 distinct samples, averaged over 10 simulation runs. The graphical analysis shows that the proposed QRMHF-DNK method consistently outperforms the existing Cooperative Deep Learning 1 and Multilingual Deep Learning 2 approaches. In the figure, the x-axis represents the number of samples acquired at different time intervals, while the y-axis indicates the corresponding true positive rate achieved using the three methods under comparison. In this algorithm, the Metropolis Hasting function first generates a sequence of sample values that progressively approximate the desired distribution, effectively identifying actual positive cases reliable news articles with high accuracy. To further address local optimality, the Quasi Reflection function is employed, which not only prevents premature convergence but also significantly improves the true positive rate. The superior performance of the QRMHF DNK method is primarily attributed to the integration of the Quasi Reflection Metropolis Hasting Firefly Optimal Feature Selection Algorithm. This combined mechanism ensures both diversity and precision in feature selection, thereby improving the overall detection reliability of the QRMHF-DNK model.

6. Conclusion

In this study, a comprehensive fake news detection framework Quasi Reflection Metropolis Hasting Firefly and Deep Neural Kernel-based QRMHF-DNK classification model was developed to enhance detection accuracy, precision, and recall while minimizing sampling error. Experimental analyses confirm that QRMHF-DNK steadily outperforms state-of-the-art Cooperative and Multilingual Deep Learning models, achieving higher true positive rates and more stable detection outcomes. Finally, the QRMHF DNK structure gives strong, adaptive and interpretable approach good for real-world fake news detection across multilingual and multimodal datasets. The combination of a Gudermannian Sigmoid activation further improves classification accuracy by nonlinearly mapping textual and contextual features. Future work may focus on extending this model to handle real-time data streams and adding cross-platform verification mechanisms to increase scalability and trustworthiness in large-scale information ecosystems.

References

[1]. Mallick, C., Mishra, S., Senapati, M.R., 2023. *A cooperative deep learning model for fake news detection in online social networks*. *Journal of Ambient Intelligence and Humanized Computing*, Springer, —. https://doi.org/10.1007/s12652-023-04680-3

[2]. Mohawesh, R., Maqsood, S., Althebyan, Q., 2023. *Multilingual deep learning framework for fake news detection using capsule neural network*. *Journal of Intelligent Information Systems*, Springer, —. https://doi.org/10.1007/s10844-023-00850-4

[3]. Zeng, J., Zhang, Y., Ma, X., 2021. *Fake news detection for epidemic emergencies via deep correlations between text and images*. *Sustainable Cities and Society*, Elsevier, 66, 102676. https://doi.org/10.1016/j.scs.2020.102676

[4]. Fang, Y., Gao, J., Huang, C., Peng, H., Wu, R., 2019. *Self Multi-Head Attention-based Convolutional Neural Networks for fake news detection*. *PLOS ONE*, 14(9), e0222713. https://doi.org/10.1371/journal.pone.0222713

[5]. Nasir, J.A., Khan, O.S., Varlamis, I., 2021. *Fake news detection: A hybrid CNN-RNN based deep learning approach*. *International Journal of Information Management Data Insights*, 1(1), 100007. https://doi.org/10.1016/j.jjimei.2020.100007

[6]. Madani, Y., Erritali, M., Bouikhalene, B., 2021. *Using artificial intelligence techniques for detecting Covid-19 epidemic fake news in Moroccan tweets*. *Results in Physics*, 24, 104186. https://doi.org/10.1016/j.rinp.2021.104186

[7]. Kaliyar, R.K., Goswami, A., Narang, P., 2020. *FakeBERT: Fake news detection in social media with a BERT-based deep learning approach*. *Multimedia Tools and Applications*, 80, 11765–11788. https://doi.org/10.1007/s11042-020-10183-2

[8]. Xue, J., Wang, Y., Tian, Y., Li, Y., Shi, L., Wei, L., 2021. *Detecting fake news by exploring the consistency of multimodal data*. *Information Processing and Management*, 58(2), 102437. https://doi.org/10.1016/j.ipm.2020.102437

[9]. Hu, L., Wei, S., Zhao, Z., Wu, B., 2022. *Deep learning for fake news detection: A comprehensive survey*. *AI Open*, 3, 1–20. https://doi.org/10.1016/j.aiopen.2022.09.003

[10]. Ahmad, I., Yousaf, M., Yousaf, S., Ahmad, M.O., 2020. *Fake news detection using machine learning ensemble methods*. *Complexity*, 2020, 1–15. https://doi.org/10.1155/2020/8885861

[11]. Paka, W.S., Bansal, R., Kaushik, A., Sengupta, S., Chakraborty, T., 2021. *Cross-SEAN: A cross-stitch semi-supervised neural attention model for COVID-19 fake news detection*. *Applied Soft Computing*, 107, 107393. https://doi.org/10.1016/j.asoc.2021.107393

[12]. Sastrawan, I.K., Bayupati, I.P.A., Arsa, D.M.S., 2021. *Detection of fake news using deep learning CNN–RNN based methods*. *The Korean Institute of Communications and Information Sciences (Elsevier)*, —. https://doi.org/10.1016/j.jksuci.2021.10.003

[13]. Jaybhaye, S.M., Badade, V., Dodke, A., Holkar, A., Lokhande, P., 2023. *Fake news detection using LSTM based deep learning approach*. *ITM Web of Conferences*, 53, 03016. https://doi.org/10.1051/itmconf/20235303016

[14]. Guler, G., Demirci, S., 2022. *Deep learning based fake news detection on social media*. *International Journal of Information Security Science*, 11(3), 191–202.

[15]. Xiang, N., 2022. *Deep learning-based fake information detection and influence evaluation*. *Computational Intelligence and Neuroscience*, 2022, 1–10. https://doi.org/10.1155/2022/4947608

[16]. Balshetwar, S.V., Abilash, R.S., Jermisha, D.R., 2023. *Fake news detection in social media based on sentiment analysis using classifier techniques*. *Multimedia Tools and Applications*, Springer, —. https://doi.org/10.1007/s11042-023-15522-0

[17]. Mishra, S., Shukla, P., Agarwal, R., 2022. *Analyzing machine learning enabled fake news detection techniques for diversified datasets*. *Wireless Communications and Mobile Computing*, 2022, 1–13. https://doi.org/10.1155/2022/9560134

[18]. Aljrees, T., Cheng, X., Ahmed, M.M., Umer, M., Majeed, R., Alnowaiser, K., Abuzinadah, N., Ashraf, I., 2023. *Fake news stance detection using selective features and FakeNET*. *PLOS ONE*, 18(7), e0287298. https://doi.org/10.1371/journal.pone.0287298

[19]. Hansrajh, A., Adeliyi, T.T., Wing, J., 2021. *Detection of online fake news using blending ensemble learning*. *Scientific Programming*, 2021, 1–10. https://doi.org/10.1155/2021/9988856

[20]. Aslam, N., Khan, I.U., Alotaibi, F.S., Aldaej, L.A., Aldubaikil, A.K., 2021. *Fake Detect: A deep learning ensemble model for fake news detection*. *Complexity*, 2021, 1–10. https://doi.org/10.1155/2021/5596782