International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Hybrid Deep Learning with CSHO based Feature Selection Model for Financial Fraud Detection

V. Rama Krishna [1], Sekharbabu Boddu[2]

Submitted: 26/05/2023 **Revised**: 06/07/2023 **Accepted**: 25/07/2023

Abstract: The progressive implementation of AI-related technology for corporate financial risk management has become a priority for banks and other financial institutions. It would appear that financial risk detection models trained on artificial intelligence perform well on fraudulent business recall. As the volume of financial upsurges, the use of traditional machine-learning algorithms for fraud detection is becoming increasingly challenging. Investigation teams may find it extremely challenging to discover safeties fraud from a mountain of electronic evidence without the aid of mechanisation, statistical methodologies, and analytics. Financial fraud can have devastating consequences for a company's stability, as well as significant losses for shareholders, the industry, and even the entire market. Existing fraud detection studies primarily rely on traditional data sources, which use limited information from financial statements. In this paper, we presented a Batch Normalization Based Auto-Encoded Gated Recurrent Unit (BN-AGRU) approach for financial fraud detection that used an auto encoder to capture local features. Through the use of a pre-trained real word vector, we integrated batch normalisation into the AE-GRU model to produce a unique architecture for financial fraud detection. The process of selecting features is handled by the Seahorse Optimizer (SHO). Chaotic Sea- Horse Optimisation (CSHO) is introduced as a hybrid algorithm that strikes a new balance between the exploration and abuse stages. To mitigate the loss function and capture long-term dependency using the arrangement input approach, we employ lengthy short-term memory as substitutes. Our study adds to existing efforts by expanding on previous modifications to standard GRUs. The experimental findings established that the suggested model outdid state-of-the-art methods across a wide variety of criteria.

Keywords: *Chaotic Sea- Horse Optimization; Financial fraud detection; Auto encoder; Corporate Financial Risk; Machine Learning.*

1. Introduction

Fraud is a worldwide problematic that touches a wide variety of businesses and has major bad effects on those companies and the people who have a stake in them. [1]. The effect on business action includes the company's long-term operations and the stakeholders (shareholders, creditors, customers, workers, and social relationships). Financial accounting and management accounting are important parts of the current method of accounting for businesses [3]. Most insurance companies now use scam detection tools that are based on business rules. These tactics work, but they may be hard to put into place and keep up [4]. The Association (2011) says that the purpose of the statements is to give info about financial flow. This is useful for most people who use financial declarations and shows that the manager is responsible for the resources given to him or her [5].

For economic growth, which is the engine that drives improvements in social benefit, it is important to use wealth in the best way possible. [6]. First, people don't know enough about what financial scam is. Even though there have been a lot of studies on financial scams in the past and some results were made, they were not very useful [7]. This work makes an addition by suggesting a model for detecting fraud that is based on a method for deep learning (Long Short-Term Memory) [8]. Financial theft is a crime that involves getting money by lying to get it for yourself [9]. The method for finding fraud could be a combination of a human process and an algorithm that can find fraud instantly [10]. Fraud detection systems are made to find or track incoming transactions by picking out odd action structures from a huge number of transaction records [11]. Most probes into securities fraud take many hours of carefully looking through proof. In the last few decades, evidence has moved from paper to digital files [12].

To start, we improved the speed of the GRU network by adding an Auto Encoder for fraud detection. This was done to help find financial theft. Second, to speed up the training, we changed the batch normalisation in the suggested BN-AGRU training process to make the nodes in each layer as simple as possible. We used a method to come up with a better plan that works a little bit better but requires a lot less computing power for this reason. In the candidate state, the hyperbolic was replaced with the activation. In particular, the Relu activation function did better than sigmoid non-linearities in deep learning methods. The consequences of the experiments showed

[1]*Phd Research Scholar , Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India*
mekrishnait1984@gmail.com
[2]*Profressor , Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India*
sekharbabu@kluniversity.in

that the new structure does a better job than what was already out there.

Here's how the break of the paper is set up: In Section 2, the linked works are listed, and in Section 3, the problematic description is given. In Section 4, a short description of the suggested method is given. In Section 5, a study of how the proposed model compares to other methods is given. Section 6 is the last part of the study.

We made a new way to find financial scams by adding batch normalisation to the AE-GRU model using a real word vector that had already been trained. Feature selection method uses Seahorse Optimizer (SHO). Chaotic Sea-Horse Optimisation (CSHO) is a hybrid algorithm with a new balance between the discovery and exploitation phases. We use long short-term memory as substitutes to reduce the loss function and the sequence input method to capture long-term relationships. Our study helps these attempts by making GRUs even better than they already are. The results of the experiments showed that the recommended model performed better than current methods in a number of ways.

2. Literature Survey

Owiti, S.O., et al. [13] looked into what makes mobile financial theft happen in Kenya. The study used both qualitative and quantitative tools to collect data. The design was made with Fraud Triangle Theory (FTT) as a help. The results showed that the fraud triangle is very useful when applied to factors that lead to mobile financial theft. Lastly, the results will have big effects on the Central Bank of Kenya, leaders in charge of financial institutions, university researchers, groups that fight fraud, and other groups.

In this work, Schneider, M., and Brühl, R. [14] use machine learning to look at how well CEO traits can predict financial theft. We show that mostly forgotten CEO traits can be used to find accounting fraud using machine learning, both on their own and as part of a new grouping with raw financial data items. Our work is based on the idea of the upper levels. We use five well-known machine learning models in the area of financial theft. In contrast to previous studies, we bring brand-new model-agnostic methods to the field of accounting fraud, which further clears up the question of how accurate individual indicators of accounting fraud are. We look at CEO predictors in particular in terms of feature importance, functional link, marginal predictive power, and how features interact with each other. We find that when CEO data and financial data are used together, the models do a lot better than when they are used separately or compared to a "no-skill" standard. Extreme Gradient Boosting and Random Forest both do a lot better than linear models, which suggests that there is a more complicated relationship between accounting fraud, CEO traits, and financial data. Also, CEO Network Size and CEO Age are right behind CEO Duality in terms of how much they help the best model predict the future. Our results are in line with our improved nonlinear models and show U-shaped, L-shaped, and weak L-shaped links between CEO Age, CEO Network Size, CEO Tenure, and financial fraud. Last but not least, our empirical data shows that financial fraud is more likely to be linked to top CEOs who are not also the chairman and CEOs with a big network and a lot of stock.

The goal of Suryani, E.'s study [15] was to find out how the size and length of KAP affect how easy it is to spot financial statement fraud. The study group was made up of 140 manufacturing companies that were listed between 2014 and 2015. SPSS version 20 was used to look at the data. In descriptive analysis, multiple regression was used. Based on the Asset Quality Index (AQI), the study found that the size of accounting companies and the length of time their auditors have worked for them are major factors in the proof of false financial statements. made no real difference. , Total Incidence against Daily Sales Total Assets (TATA), and Messod D Beneche Score (M-score), but decline As measured by the Depreciation Index Index (DEPI), it has a big effect on how much false financial information is reported.

Alwadain, Ali, and Muneer [16] came up with a unique way to spot financial scams using machine learning. In a simulated dataset, users have given different machine learning models transaction-level features from 6,362,620 deals. The connections between different features are also looked at. Also, about 5000 more data samples were made with the help of a Conditional Generative Adversarial Network for Tabular Data (CTGAN). When the suggested prediction was tested, it was found to be more accurate than the other machine-learning-based methods. With an accuracy score of 0.999, XGBoost did better than all of the other 27 algorithms. But accuracy score of 0.998. The results are very important for officials and politicians who want to come up with new, effective ways to reduce risks of financial crime. They are especially important for banks and other financial institutions.

Nahri Aghdam Ghalejoogh, J's recent work [17] uses support vector machine regression and boosted regression tree, two types of machine learning, to find financial fraud on the Iranian stock market. The model with the lowest RMSE is the boosted regression tree machine model. The enhanced regression tree model, which also looks at how sensitive the models are, has the highest sensitivity because it correctly found that there was no financial crime on the Tehran Stock Exchange market. The boosted regression tree has the best kappa coefficient compared to

the other models used in the study. This means that it works as predicted.

The company's financial data has been portrayed as a time- and money-themed three-dimensional (3-D) data cube by Chen, Z.Y., and Han, D. [18]. A two-stage mapping approach may be utilized to manage all of this data concurrently. As a preliminary step, we sparsely pattern the data from our common domain. The output from the first stage is then used in the second stage, where a high-level, low-dimensional representation of the characteristics is learned. We use actual financial data from Research (CSMAR) database to assess the effectiveness of the proposed two-stage mapping technique. We discover that the processing gap has to be at least three years, or twelve quarters, for identifying corporate financial fraud (CFF), and that detection effectiveness improves considerably as the number of quarters increases. This discovery is significant because it makes logic and because a comparable shared processing delay has been seen in other data sets. We also discover that jointly processing the monetary and temporal information sets aids in CFF identification. The projected two-stage mapping technique for combined processing in the domains of temporal and financial information is straightforward to implement, allowing for the development of intelligent CFF detection systems.

The authors of this study, Yadiati, W. and Rezwiandhari, A. [19], set out to determine whether or not the hexagonal fraud hypothesis is useful for detecting instances of money misappropriation at BUMN. Using the Indonesia Stock Exchange's (IDX) publicly available annual financial statements as well as multiple regression analysis, this study provides quantitative insights into the company's performance. Financial predictability (stimulus), environmental influence (stimulus), and the character of the market (opportunity). Changing the auditor (Capability), the board of directors (Rationalisation), the sum of CEO photographs (Arrogance), and collaboration with the government on projects (Collusion) were all shown to be effective means of identifying SOEs that were providing misleading financial statements in the research. Finding SOEs with fabricated financials may be done by examining the company's financial health, external pressure, the nature of the industry, director turnover, and collaboration with the government on initiatives. How simple it is to recognize phony financial accounts in BUMN has not changed from 2012 to 2019 despite changes in the number of auditors and CEO photos. The study's findings might provide a general description of the factors that may lead SOEs to report inflated financials. The study's findings, it is believed, would aid interested parties in making informed decisions.

3. Problem Statement

A recurrent neural network (RNN) is a convenient design to represent the components involved in performing financial fraud detection. However, the problem of vanishing gradients and exploding is effectively addressed in regular RNN designs, while this problem is solved by well-known gated RNNs like GRU. The main purpose of using GRU models is to extract invariant information from continuous input by deriving higher-level translations and assembling many layers. According to recent literature, GRU models are gradually gaining popularity due to their excellent patterning capabilities for extracting long-term semantic relationships within sequential components.

However, the research indicates that two restrictions have been the subject of prior research. The inability to dimension-reduce data based on input data is the primary shortcoming of classical GRU. This means it's not a good option for detecting financial fraud. The second restriction is associated with RNN networks in general, not only GRU. Even when used in conjunction with other classification strategies, this strategy may not yield the best accurate results when detecting financial fraud. As the number of possible successive inputs grows, the current setup of RNNs is becoming increasingly problematic due to the removal of crucial properties during training.

Therefore, we put forth the BN-AGRU, a model that employs the identical weights parameters (W, U) as the conventional GRU. In most cases, the AE model also produces a more accurate illustration of the input than the raw input itself, and it constantly decreases the input data while picking relevant features for training. To further decrease the layer-by-layer complexity, we employed Batch Normalisation (BN) in the BN-AGRU training procedure.

4. Proposed Methodology

4.1 Dataset Description

The original untried dataset was obtained from a chief Chinese Internet financial service provider. Following data pre-processing, the dataset comprises 192586 data samples, with 4375 fraud samples. The dataset contains over 60 data fields, including initial amount, financial status, balance sheet, currency, records, sale status, and so on. To maintain the data confidentiality of sensitive info, not all learning models are divided into 8 -validation. Each time, the training-to-testing data ratio is nearly 4:1.

4.2. Data pre-processing

Data should all be performed before smearing the perfect to the dataset.

4.2.1. Data Validation

This step is used to validate the data within the dataset, or a negative amount.

4.2.2. Normalization

The model's accuracy relies on rescaling the input variables to the range [-1,1]. The dataset's numeric column values will need to be converted in order to be utilized on a common scale, but this process should not distort the range of values or remove any information. The normalization process use Equation 1.

$$x(i) = \frac{x(i)-\overline{x}}{s(x)} \quad (1)$$

4.2.3. Dataset samples divide

To provide an accurate evaluation of performance, data samples must be split between training and testing. The suggested model trains on 70% of the dataset and tests on the remaining 30%..

4.3. Feature Selection using SHO

4.3.1. Brief Introduction of Sea Horse optimizer

For better results, the authors of this study integrate the latest metaheuristic procedure, Sea-Horse Optimizer (SHO), with a chaotic algorithm [20]. The Chaotic Sea-Horse Optimizer (CSHO) strategy aims to improve the SHO algorithm's performance in the sweet spot between exploration and exploitation by utilizing a mixed approach. The suggested strategy combines chaotic and SHO approaches. The following is a condensed explanation of this study's significance:

1. CSHO is given, a hybrid algorithm with a novel equilibrium between the exploration and exploitation steps.

2. Parameters of the power system stabilizer are adjusted using the suggested CSHO.

A. SHO

The SHO takes cues from the sea-horse's aquatic lifestyle, including its foraging, migration, and reproduction. The SHO algorithm is inspired by the social behavior of moving around and searching for food like a seahorse would. This is achieved using the metaheuristic idea of exploration and exploitation. When the two parts are no longer present, the breeding process reaches its last stage. This is a detailed representation of the SHO procedure:

$$Sh = \begin{bmatrix} x_{1,i} & \cdots & x_{1,Dim-1} & x_{1,Dim} \\ \cdots & : & : & : \\ x_{N.i} & \cdots & x_{N,Dim-1} & x_{N,Dim} \end{bmatrix} (2)$$

$$Sh_{ij} = Rand \times \left(UB_j - LB_j\right) + LB_j \quad (3)$$

$$Sh_{elite} = \arg\min\left(f(X_i)\right) \quad (4)$$

Where Dim is the size of the population, N, and the dimension of the variable. Upper and lower bounds, represented by UB and LB, are the probabilistic outcomes of any solution. Rand represents a random number in the interval [0,1]. $Shelite$ is a icon for people who meet the criteria for elite status in terms of physical fitness. SHO mimics the behaviors of seahorses by swimming, hunting, and reproducing.

B. Seahorse Movement Behavior

Seahorse behavior may be understood in terms of the normal distribution. two example studies illustrating the optimal border between exploration and exploitation.

Case 1: To increase the size of the locally-solvable problem space, the agent spirals towards the Xelite and continuously adjusts the rotation angle. Mathematical expressions for the first case are as follows.:

$$X_n(t+1) = X_i(t) + Levy(\lambda)\left(\left(X_{elite}(t) - X_i(t)\right) \times x \times y \times z + X_{elite}(t)\right) \quad (5)$$

$$\theta = rand \times 2\pi \quad (6)$$

$$x = \times \cos(\theta) \quad (7)$$

$$y = \rho \times \sin(\theta) \quad (8)$$

$$z = \rho \times \theta \quad (9)$$

$$\rho = \mu \times e^{\theta v} \quad (10)$$

$$Levy\,(\lambda) = s \times \frac{w \times \sigma}{|K|^{\frac{1}{\lambda}}} \quad (11)$$

$$\sigma = \left(\frac{\Gamma(1+\lambda) \times \sin\left(\frac{\pi\lambda}{2}\right)}{\Gamma\left(\frac{1+\lambda}{2}\right) \times \lambda \times 2^{\left(\frac{\lambda-1}{2}\right)}}\right) \quad (12)$$

Where the constants u $(default = 0.05)$ and v $(default = 0.05)$ define the logarithmic spiral that determines the rod's length. by ρ. λ is a random sum [0, 2]. k and w are random statistics [0, 1]. s is a fixed endless of 0.01.

Case 2: In order to improve its traversal, a seahorse will make a brownian motion in response to ocean waves that simulates seahorse. The correct formulation of this is as follows.:

$$X_n(t+1) = X_i(t) + rand \times l \times \beta_i \times \left(X_i(t) - \beta_i \times X_{elite}(t)\right) \quad (13)$$

$$\beta_i = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (14)$$

$$X_n(t+1) = \left\{X_i(t) + Levy(\lambda)\left(\left(X_{elite}(t) - X_i(t)\right) \times x \times y \times z + X_{elite}(t)\right) \; if \; r_1 > 0 \; X_n(t+1) = X_i(t) + rand \times l \times \beta_i \times \left(X_i(t) - \beta_i \times X_{elite}(t)\right) \; if \; r_1 \leq 0\right\} (15)$$

where βi is the Brownian motion random walk coefficient. The value of l remains unchanging. (default=0.5). $r1$ is symbolized as a accidental value.

C. Seahorse Foraging Behavior

There are only two outcomes when seahorses search for food: success and failure. With r2 > 0.1, the requirement has been met. When the seahorse is quicker than its prey, it is in this state. When the reaction is not what was expected, however, we have a failure state. When searching for food, seahorses have the following success and failure criteria:

$$X_n(t+1) = \begin{cases} \alpha \times \left(\left(X_{elite}(t) - rand \times X_{new}(t) + (1-\alpha) \times X_{elite}(t)\right)\right) if\ r_2 > 0 \\ (1-\alpha) \times \left(\left(X_{new}(t) - rand \times X_{elite}(t)\right) + (\alpha) \times X_{new}(t)\right) if\ r_2 \le 0 \end{cases} \quad (16)$$

$$\alpha = \left(1 - \frac{t}{T}\right)^{\frac{2t}{T}} \quad (17)$$

where $Xnew$ new is the novel position of the seahorse. $r2$ is a accidental sum [0, 1]. T is the maximum iteration.

D. Seahorse Breeding Behavior

The seahorse population is evenly split between males and females at breeding season (50 percent each).

$$Father = X_{sort}\left(1{:}\frac{pop}{2}\right) \quad (18)$$

$$Mother = X_{sort}\left(\frac{pop}{2} + 1{:}pop\right) \quad (19)$$

Sorted $Xsort$ values return the consequence in ascending order. $Father$ and $Mother$ were chosen arbitrarily. In the SHO procedure, each pair crops one child.

$$X_i = (1 - r_3)X_{mother} + r_3 X_{father} \quad (20)$$

Where $r3$ is a random sum [0, 1].

E. The Innovative Chaotic Sea-Horse Optimizer (CSHO)

Several chaotic maps of varying forms have been utilized in several studies to optimize algorithms. The statistics of a chaotic map are inherently unpredictable and reflect its dynamic nature. Parameter adjustments effect behavior in the future. In such a way that subtle alterations to the inputs provide varying results. In this piece, we substitute the logistic type chaotic map for the rand in Eq. (6). The logistic type chaotic map has the following mathematical equation:

$$y\log_{(i+1)} = a \times y\log_{(i)}\left(1 - y\log_{(i)}\right) \quad (21)$$

Where a is 4. So, the map adjustable with variety [0, 1] is found. So, Equation (6) turns into Equation (22) as shadows:

$$\theta = y\log \times 2\pi \quad (22)$$

Pseudo code can be understood in Procedure 1.

Algorithm 1: Pseudo Code of Chaotic Sea − Horse Optimizer

Input: *population size pop, maximum iteration T and variable dimension Dim*

Output: *Optimal search agent Xbest*

1: *procedure CSHO*

2: *Initialize search agent Xi*

3: *Calculate the fitness value of each search agent*

4: *Determine the best search agent Xe*

/∗ Movement behavior ∗/

5: ***While*** $(t < Max\_iteration)$ ***do***

6: ***if*** $r1 = randn > 0$ ***do***

7: $u \leftarrow 0.05$

8: $v \leftarrow 0.05$

9: $l \leftarrow 0.05$

10: $y\,log(i+1) \leftarrow a \times y\,log_{(i)}\left(1 - y\,log_{(i)}\right)$ *using Eq.* (21)

11: $\theta \leftarrow ylog \times 2\pi$ *using Eq.*(22)

12: $x \leftarrow cos(\theta)\ using\ Eq.(7)$

13: $y \leftarrow \rho \times sin(\theta)\ using\ Eq.(8)$

14: $z \leftarrow \rho \times \theta\ by\ using\ Eq.(9)$

15: $\rho \leftarrow \mu \times e\ by\ using\ Eq.(10)$

16: ***else if do***

17: $update\ positions\ of\ the\ search\ agent\ by\ using\ Eq.(11)$

18: $end\ if$

$/*\ Foraging\ behavior\ */$

19: $update\ positions\ of\ the\ search\ agent\ by\ using\ Eq.(16)$

20: $Calculate\ the\ fitness\ value\ of\ each\ search\ agent$

$/*\ Foraging\ behavior\ */$

21: $Select\ fathers\ and\ mother\ by\ using\ Eq.(18)\ and\ Eq.(19)$

22: $Breed\ Offsping\ by\ using\ Eq.(20)$

23: $Calculate\ the\ fitness\ value\ of\ each\ offspring$

24: $Select\ the\ next\ offspring\ and\ parent\ ranked\ top\ pop\ in\ fitness\ values$

25: $Update\ elite\ (Xelite)\ position$

26: $t \leftarrow t + T$

27: ***end while***

28: ***end procedure***

4.4. Proposed Classifier

In the suggested procedure [21], we go into detail about the created model, which combines GRU with normalisation. The suggested BN-AGRU model for financial fraud detection in NLP seeks to solve the issues with dimensionality reduction and complexity present in the standard GRU method. In order to enhance the dimensionality reduction capacity of the projected architecture, auto encoder layers were additional to the GRU network. In addition, the learning procedure for a standard GRU design with AE layers is executed in two stages: 1) It starts by compressing the raw features of the input data by minimizing the input features vector x.

Training of network are separated into four discrete steps in the proposed BN-AGRU approach's execution technique. All computational limits of the GRU and auto encoder were set before training began. threshold value, encoded and decoded activation function, and so on are all parameters of the autoencoder that have been selected. In the second stage, computations are performed between the input and hidden layers of AE during the encoding phase, and between the hidden and output layers of AE during the decoding phase, with reconstruction error taken into account. When calculating the rebuilding error, the network employs a threshold value. If the error in reconstruction is below a certain level, the data is advanced.

Our improved performance can be attributed to the higher threshold value, and the network continues to improve accuracy by increasing the threshold price at the conclusion of each epoch. In contrast, the third stage operates the mapping apparatus of the created BN-AGRU model. In which the output of the AE layer is sent into a GRU network to learn useful features. The GRU hidden layers' output feature vectors are then used as input vectors for the Softmax layer, which performs the final classification. The designed GRU architecture with a combined auto encoder was shown off. By constructing BN-AGRU, two major modifications were made to the suggested design to boost the efficiency of the current GRU.

The training method of the new BN-AGRU model has been modified in several ways, including the incorporation of AE and GRU and the adoption of a batch normalisation technique.

4.4.1 Batch Normalization

Adding Batch Normalisation (BN) to the training process

of the well-known BN-AGRU model has increased its productivity. Incorporating the batch normalisation approach before the encoded and decoded function needed is the primary technological improvement described in BN-AGRU. There is a considerable danger of oscillations in the values of input features from training without the usage of the BN approach.

The BN method was used to enhance the GRU training procedure during the development of the BN-AGRU model. As the AE input vector, the Batch has been functional to features, and the encoded vector of the hidden layer is the consequence of the input layer's computations. Finally, the decreased input vector is the AE output vector, which is sent to the first GRU layer. This illustration shows how the planned BN-AGRU paradigm works in practise. In addition, incorporating batch normalisation in GRU layers accelerates training without adding further computational complexity to the network. Following is a description of the equations used to normalise BN nodes, which can be found in Equations (23) and (24).

:

$$\overline{x} = (x - M[x])/sqrt(var(x)) \quad (23)$$

$$\overline{\overline{x}} = (\gamma * \overline{x} + \beta) \quad (24)$$

Where $\overline{x}$ is the novel rate of a solitary node, $M[x]$ is the batch $mean, and\ var(x)$ is the batch variance. And $\overline{\overline{x}}$ is the final regularized value, γ, β are layer. Finally, this research modified equation by replacement the tanh with Relu activation in the applicant state, as shown in equation (25).

$$\hat{h}_t^i = Relu\left(Wx_t + U(r_t * h_{t-1})\right)^i + b_h \quad (25)$$

Final output

$$h_t = (1 - z_t) * h_{t-1} + z_t * h_t \quad (26)$$

Relu activation has been stated to outperform nonlinearities in deep learning algorithms. The building of the BN-AGRU approach is demonstrated in Figure 1

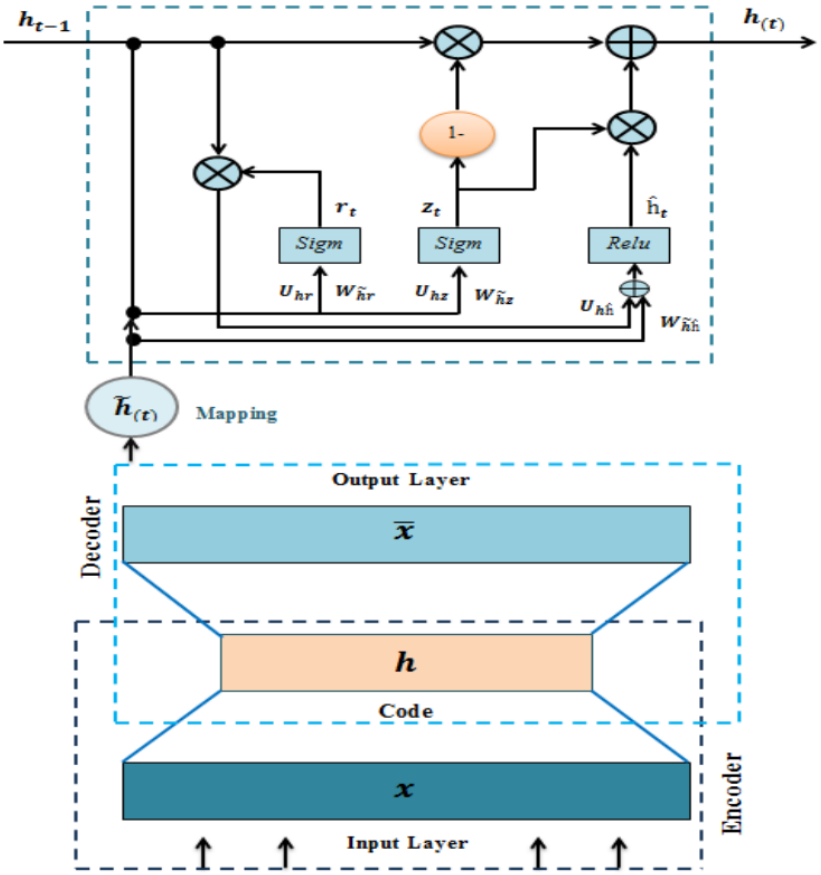

Fig 1: Architecture of the projected perfect

5. Results and Discussion

Clusters of 30 similar node" and the others as "worker nodes," are used to conduct research in groups. There 8-cores and 64 GB-RAM in each scheme. CentOS 7 with the Java SE Kit 10 and Scala 2.12 is the OS.

5.1. Performance Metrics

Multiple metrics, including runtime, were used to evaluate our models' efficacy. This evaluate metrics included F1, and Cohen's kappa. We also used the k-fold cross-validation test in our analysis. Our performance metrics are derived from Equations (27-31) and focus on execution speed rather than training and prediction period.:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (27)$$

$$Precision = \frac{TP}{TP+FP} \quad (28)$$

$$Recall(True\ positive\ Rate) = \frac{TP}{TP+FN} \quad (29)$$

$$F1Score = \frac{2TP}{2TP+FP+FN} = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (30)$$

$$Cohen\ Kappa\ Score = \frac{Accuracy-P_{random}}{1-P_{random}} \quad (31)$$

In a TP (True Positive), both the prediction and the intended outcome are accurate. Sometimes a prediction will be spot right, yet the target will be off. This condition is commonly referred to as TN. An FP occurs when a prediction is true when the target is false, and a FN (False Negative) occurs when a forecast is false while the target is true. Existing models from [13-19] and ELM are taken into account; these models use a wide variety of datasets to detect fraud. Thus, we apply the available models to our data and present a weighted average of the outcomes in Tables 1 and 2.

Table 1: Validation Investigation of Projected feature selection perfect

Model	Accuracy	Precision	Recall	F1-score	Cohen Kappa score
HHO	87.29	0.861	0.863	0.880	0.7812
Bird Swarm	88.29	0.896	0.872	0.893	0.7996
Whale	89.76	0.903	0.890	0.895	0.8233
Butterfly	91.23	0.925	0.912	0.922	0.8597
Proposed	93.76	0.942	0.932	0.932	0.8809

In above table 1 represent that the Validation Analysis of Proposed feature selection model. In this analysis, we have used different models. in this analysis of HHO model reached the accuracy range of 87.29 and also the precision cost as 0.861 and the recall charge as 0.863 and F1- score value as 0.880 and finally the Cohen Kappa score as 0.7812 respectively. And another model as Bird Swarm model reached the accuracy range of 88.29 and also the precision value as 0.896 and the recall value as 0.872 and F1- score rate as 0.893 and finally the Cohen Kappa score as 0.7996 respectively. Also, another scheme as Whale model reached the accuracy range of 89.76 and also the precision rate as 0.903 and the recall value as 0.890 and F1- score value as 0.895 and finally the Cohen Kappa score as 0.8233 respectively. Also, another model as Butterfly model reached the accuracy range of 91.23 and the precision value as 0.925 and recall value as 0.912 and F1- score value as 0.922 and finally the Cohen Kappa score as 0.8597. And finally, the proposed model reached the accuracy worth as 93.76 and the precision value asv0.942vand recall rate as 0.932 and F1- score value as 0.932 and finally the Cohen Kappa score as 0.8809. In this comparison analysis study, we analysed the performance of some discussed model, the proposed model attained the better performance than other compared model.

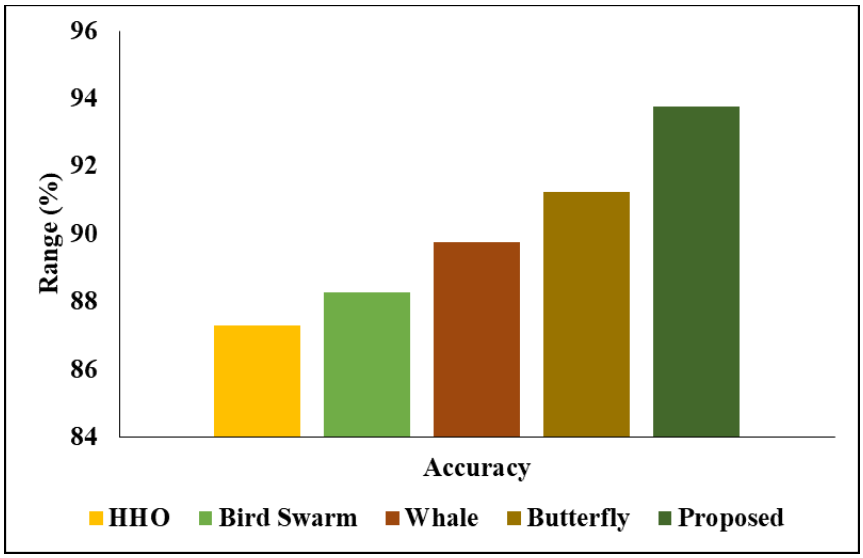

Fig 2: Accuracy Analysis

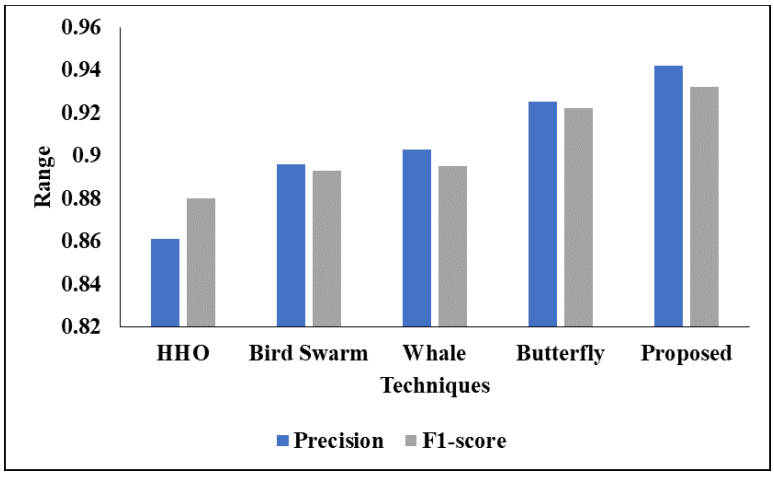

Fig 3: Validation Performance of Proposed Optimization

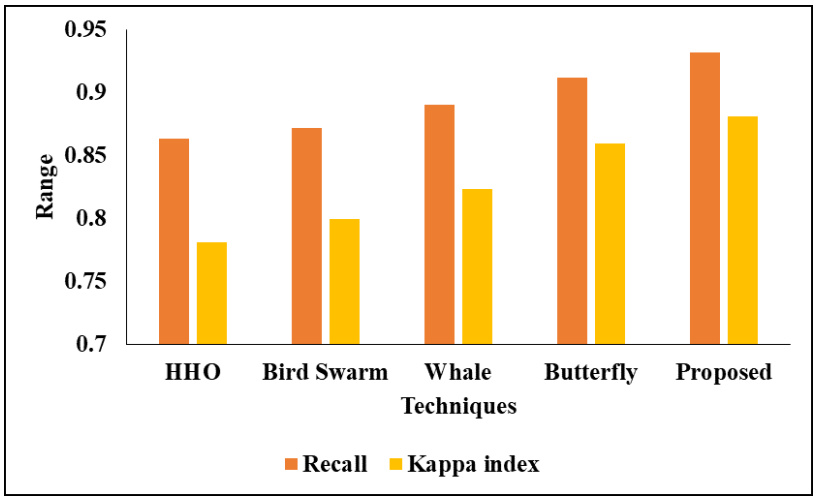

Fig 4: Graphical Representation for proposed CSHO

In below table 2 represent that the Analysis of Proposed Hybrid Classifier.in this analysis, we used different model as LR, DBN, AE, GRU with Proposed model. In this analysis, the LR model reached the accuracy of 92.07 and the precision charge as 0.92 and another metrics of recall value as 0.880 then the F1- score as 0.87 and finally, the Cohen Kappa score as 0.809 respectively. Also, another model of DBN model reached the accuracy of 94.94 and the precision value as 0.93 and another metrics of recall value as 0.898 and the F1-score as 0.89 and finally, the Cohen Kappa score as 0.831 respectively. Also, another model of AE model stretched the accuracy of 95.47 and

the precision charge as 0.93 and another metrics of recall value as 0.916 and the F1- score as 0.91 and finally, the Cohen Kappa score as 0.859 respectively. Also, another model of GRU model reached the accuracy of 96.01and the precision value as 0.94 and another metrics of recall rate as 0.931 and the F1- score as 0.93 and finally, the Cohen Kappa score as 0.885 respectively. Also, another model of Proposed model reached the accuracy of 98.20 then the precision rate as 0.96 and another metrics of recall value as 0.956 And the F1-score as 0.96 And finally, the Cohen Kappa score as 0.928 respectively. In this comparison analysis study, we analysed the performance of some discussed model, the proposed model attained the better performance than other compared model.

Table 2: Analysis of Proposed Hybrid Classifier

Model	Accuracy	Precision	Recall	F1-score	Cohen Kappa score
LR	92.07	0.92	0.880	0.87	0.809
DBN	94.94	0.93	0.898	0.89	0.831
AE	95.47	0.93	0.916	0.91	0.859
GRU	96.01	0.94	0.931	0.93	0.885
Proposed	98.20	0.96	0.956	0.96	0.928

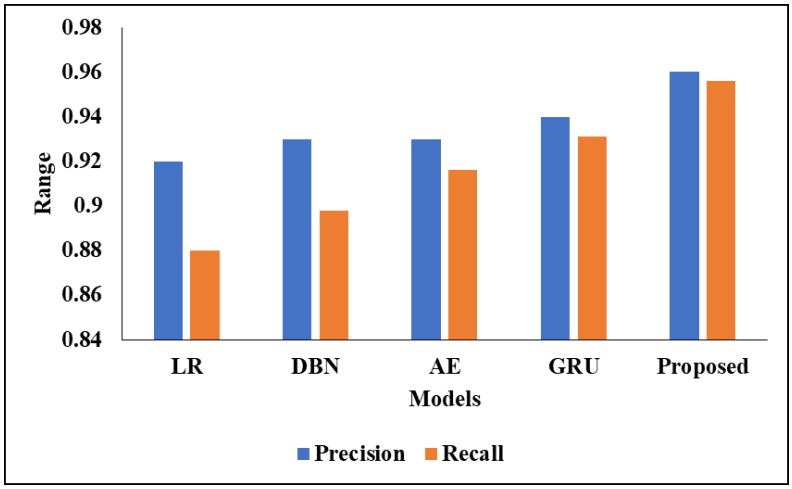

Fig 5: Analysis of Various DL models

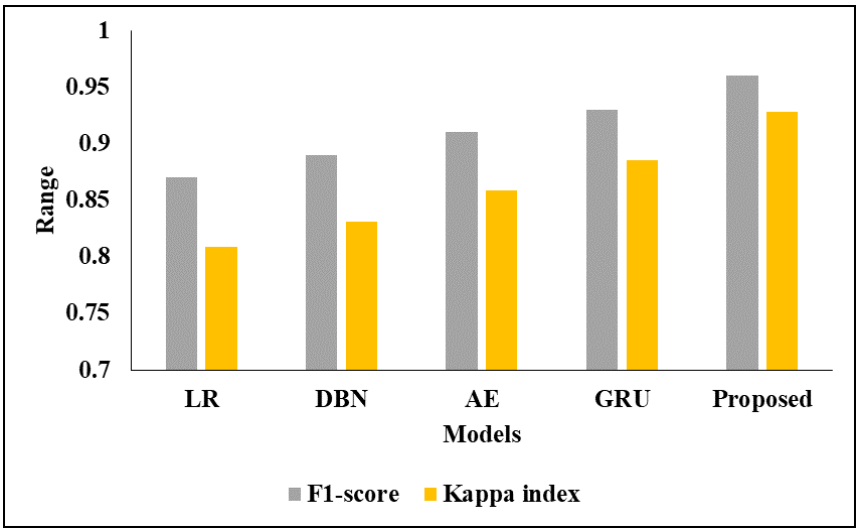

Fig 6: Comparative Performance of Proposed Model

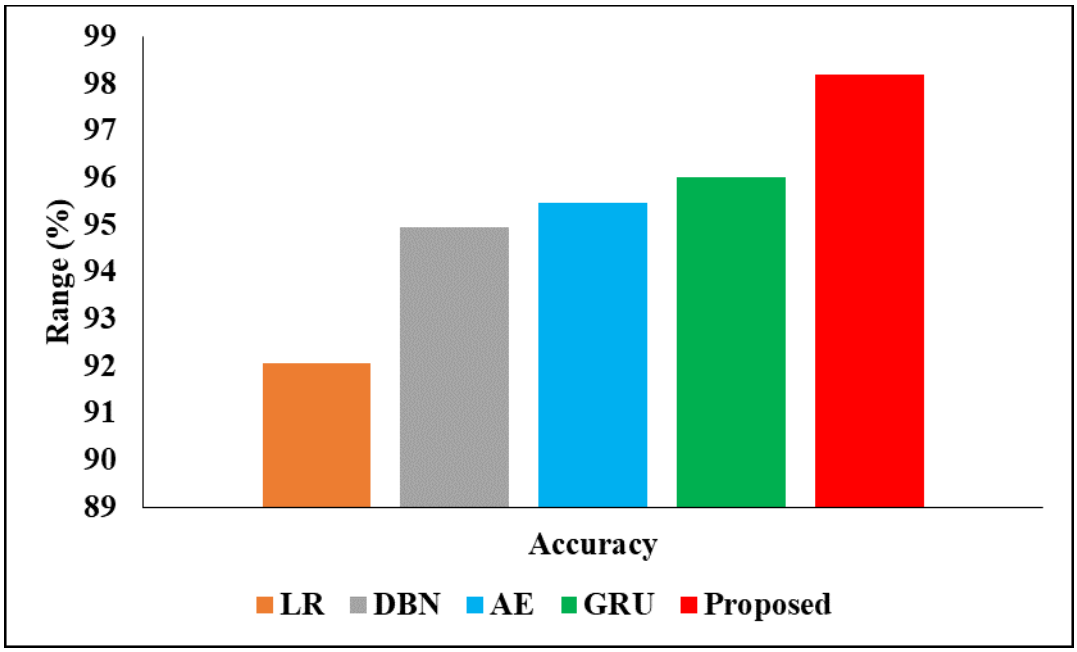

Fig 7: Accuracy analysis

6. Conclusion

Fraud detection on financial data requires constant updating. Financial fraud detection is becoming more difficult as evidence of business grows. We contribute to financial fraud detection with BN-AGRU, a tool for rapid analysis of a specific fraud pattern. Using benchmark financial fraud detection datasets, a brief explanation of the simulation consequences of the proposed BN-AGRU is compared with out-dated deep learning models. The outcomes are analysed using various evaluation metrics. The novel enhanced SHO algorithm is used in this paper as a feature selection technique, which is a hybrid of the SHO process and the chaos map technique. The benchmark function is used to begin testing. The performance of CSHO has improved and has reached a novel balance point among survey and misuse. This article employs integration in conjunction with the chaotic method. The proposed technique for financial fraud detection has been validated with an average accuracy of 98.2%, which is higher than the base line approach, which has an overall accuracy of 89.4%. Finally, we discovered that the BN-AGRU model is capable of detecting financial fraud with a low error rate. The following future work is proposed to implement hyper parameter tuning in deep learning with feature selection for financial fraud detection in order to improve classification accuracy.

References

[1] Craja, P., Kim, A. and Lessmann, S., 2020. Deep learning for detecting financial statement fraud. Decision Support Systems, 139, p.113421.

[2] Jan, C.L., 2021. Detection of financial statement fraud using deep learning for sustainable development of capital markets under information asymmetry. Sustainability, 13(17), p.9879.

[3] Sun, H., 2023. Construction of integration path of management accounting and financial accounting based on big data analysis. Optik, 272, p.170321.

[4] Krishnavardhan, N., Govindarajan, M. and SV, A.R., 2023. Flower Pollination Optimization Algorithm with Stacked Temporal Convolution Network based Classification for financial anomaly Fraud Detection.

[5] Putri, M.P.E., Financial Statement Fraud Detection Using Diamond Theory Analysis and Covid-19.

[6] Tang, J. and Karim, K.E., 2019. Financial fraud detection and big data analytics–implications on auditors' use of fraud brainstorming session. Managerial Auditing Journal, 34(3), pp.324-337.

[7] Liu, C., Chan, Y., Alam Kazmi, S.H. and Fu, H., 2015. Financial fraud detection model: Based on random forest. International journal of economics and finance, 7(7).

[8] Alghofaili, Y., Albattah, A. and Rassam, M.A., 2020. A financial fraud detection model based on LSTM deep learning technique. Journal of Applied Security Research, 15(4), pp.498-516.

[9] Singh, A. and Jain, A., 2019. Financial fraud detection using bio-inspired key optimization and machine learning technique. International Journal of Security and Its Applications, 13(4), pp.75-90.

[10] Mehbodniya, A., Alam, I., Pande, S., Neware, R., Rane, K.P., Shabaz, M. and Madhavan, M.V., 2021. Financial fraud detection in healthcare using machine learning and deep learning techniques. Security and Communication Networks, 2021, pp.1-8.

[11] Megdad, M.M., Abu-Naser, S.S. and Abu-Nasser, B.S., 2022. Fraudulent financial transactions detection using machine learning.

[12] Krishnan, S., Shashidhar, N., Varol, C. and Islam, A.R., 2022. A Novel Text Mining Approach to Securities and Financial Fraud Detection of Case Suspects. International Journal of Artificial Intelligence and Expert Systems, 10(3).

[13] Owiti, S.O., Ogara, S. and Rodrigues, A., 2023. CONTRIBUTING FACTORS TO MOBILE FINANCIAL FRAUD WITHIN KENYA. EPRA International Journal of Research and Development (IJRD), 8(1), pp.32-39.

[14] Schneider, M. and Brühl, R., 2023. Disentangling the black box around CEO and financial information-based accounting fraud detection: machine learning-based evidence from publicly listed US firms. Journal of Business Economics, pp.1-38.

[15] Suryani, E., Winarningsi, S., Avianti, I., Sofia, P. and Dewi, N., 2023. Does Audit Firm Size and Audit Tenure Influence Fraudulent Financial Statements?. Australasian Accounting, Business and Finance Journal, 17(2), pp.26-37.

[16] Alwadain, A., Ali, R.F. and Muneer, A., 2023. Estimating Financial Fraud through Transaction-Level Features and Machine Learning. Mathematics, 11(5), p.1184.

[17] Nahri Aghdam Ghalejoogh, J., Rezaei, N., Aghdam Mazarae, Y. and Abdi, R., 2023. Detecting financial fraud using machine learning techniques. International Journal of Nonlinear Analysis and Applications.

[18] Chen, Z.Y. and Han, D., 2023. Detecting corporate financial fraud via two-stage mapping in joint temporal and financial feature domain. *Expert Systems with Applications*, p.119559.

[19] Yadiati, W. and Rezwiandhari, A., 2023. Detecting Fraudulent Financial Reporting In State-Owned Company: Hexagon Theory Approach. JAK (Jurnal Akuntansi) Kajian Ilmiah Akuntansi, 10(1), pp.128-147.

[20] Aribowo, W., 2023. A Novel Improved Sea-Horse Optimizer for Tuning Parameter Power System Stabilizer. Journal of Robotics and Control (JRC), 4(1), pp.12-22.

[21] Zulqarnain, M., Alsaedi, A.K.Z., Sheikh, R., Javid, I., Ahmad, M. and Ullah, U., 2023. An Improved Deep Neural Network Based on Combination of GRU and Auto Encoder for Sentiment Analysis.

[22]

[23] Dr. Antino Marelino. (2014). Customer Satisfaction Analysis based on Customer Relationship Management. International Journal of New Practices in Management and Engineering, 3(01), 07 - 12. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/26

[24] Ahire, P. G. ., & Patil, P. D. . (2023). Context-Aware Clustering and the Optimized Whale Optimization Algorithm: An Effective Predictive Model for the Smart Grid. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1), 62–76. https://doi.org/10.17762/ijritcc.v11i1.5987

[25] Dhabliya, D. (2021). Delay-tolerant sensor network (DTN) implementation in cloud computing. Paper presented at the Journal of Physics: Conference Series, , 1979(1) doi:10.1088/1742-6596/1979/1/012031 Retrieved from www.scopus.com