

¹Ruifen Zhao*

Construction and Optimization of Financial Transformation Decision Model Based on Data Mining



Abstract: - Real-time data processing and analysis is crucial in the financial markets since transactions in real time are directly linked to profit. Understanding the integration of finance and business allows the organization to enhance its core competitiveness by improving management and enabling it to lead business expansion more effectively. This manuscript proposes a Sparse Oblique Trees Algorithm (SOTA) optimized with the Waterwheel Plant Algorithm (WWPA) for the construction of financial transformation. The data are collected from multi-national organization dataset. Afterward, the data's are fed to pre-processing. In pre-processing segment; it removes the noise and enhances the input data's utilizing Data-Adaptive Gaussian Average Filtering (DAGAF). The outcome from the pre-processing data is transferred to the SOTA. The investment and dividend are successfully classified by using SOTA. The WWPA is used to optimize the weight parameter of SOTA. The proposed SOTA-WWPA is applied in python working platform. The proposed technique was computed by examining performance indicators like precision, F1-score, accuracy, sensitivity and recall. The proposed SOTA-WWPA technique yields improved results in terms of accuracy (16.65%, 18.85%, and 17.89%), sensitivity (16.34, 12.23%, and 18.54%), and precision (14.89%, 16.89%, and 18.23%). The proposed FT-SOTA-WWPA method is contrasted with the existing methods like FT-CNN, FT-MBLSTMNN, and FT-SVM models respectively.

Keywords: Waterwheel Plant Algorithm, Sparse Oblique Trees Algorithm, Data-Adaptive Gaussian Average Filtering, Financial transformation.

I. INTRODUCTION

Many aspects of the 21st century world include economic globalization, data informatization, and financial internationalization. People can find crucial information through extensive analysis of informatized financial data, which is becoming more and more important in daily life [1, 2]. Data mining is a technique that will likely be developed in the future to help with the processing of connected financial services [3]. People can successfully receive helpful knowledge ahead of time through data mining. Here's an a summary of data mining: data mining is the process of processing data from a significant amount of random data in large databases in order to extract hidden and maybe relevant information [4, 5]. This procedure must deal with a significant volume of business data, therefore artificial intelligence, statistics, automation, and the ability to condense and apply vast amounts of data are all necessary [6]. Finding directions, seizing market chances, and maximizing profits are all made possible with the help of data mining [7]. The financial sector requires extensive data collection and processing due to its complexity [8]. Due to the complexity of these transactions, information asymmetry, and the volume of people conducting related business on a daily basis, Financial services are provided by most financial banks and financial institutions like loan, personal deposit, credit card, investment business. As a result, these financial services generate enormous amounts of data [9, 10].

The article's financial analysis technology, which has particular theoretical and practical relevance and complies with both the evolving body of accounting knowledge and the real-world business requirements of organizations, serves as the foundation for the data mining platform [11, 12]. Errors and delays resulting from mergers plague traditional financial analysis. Real-time, multidimensional, and quick financial analysis is now possible because to the growth of large data and the use of data mining techniques [13]. Financial analysis systems will be established and real-time integration will be facilitated by data mining technology. In addition to establishing a theoretical framework for management accounting and monitoring its general development, it filters and analyzes both external and internal business data [14, 15]. Big data has made decision-making more challenging for businesses, which has given managers new ideas for making decisions. Data loss is one such challenge.

Data mining technology can create an intelligent analysis system, filter and extract relevant information from vast amounts of data, and generate insightful insights for decision-making. Particularly in recent years, a number of advancements in data modeling and algorithm optimization have accelerated the adoption of data mining technology across a wide range of academic disciplines and industries, including financial analysis [16]. Financial analysis's goal is to evaluate a company's operations and financial standing using logical and scientific analysis on a basis of pertinent accounting and business data

¹Associate Professor, Department of International Trade, Shanxi Vocational College of Tourism, Taiyuan, Shanxi, 030031, China

*Corresponding author e-mail: 13834668985@163.com

Copyright © JES 2024 on-line: journal.esrgroups.org

It guarantees that the company gets assessed and aids in high-level decision-making [17]. From its inception at the start of the 20th century to the major financial operations and decision support for businesses, financial analysis has evolved over almost a century. Businesses are gathering and storing an increasing amount of data as the economy and technology evolve [18]. The limitations of conventional financial analysis techniques are progressively becoming apparent, necessitating and pressing modern organizations' demand for decision-making.

"The following is an outline of the paper's main contributions:"

- At first, the data are gathered via the dataset of multi-national organization, anonymized here as 'Oilco'.
- Using a Data-Adaptive Gaussian Average Filtering to eliminate the noise of multi-national organization dataset in the pre-processing segment.
- The pre-processed data's are fed into the Sparse Oblique Trees Algorithm, in order to effectively categorize the investment, and dividend of the financial data.
- The proposed SOTA-WWPA technique is implemented, and performance metrics like precision, F1-score, sensitivity, accuracy, and Recall are analyzed.

The rest portions of this manuscript are organized as follows: segment 2 examines a survey of the literature; part 3 explains the proposed method; part 4 describes the results and discussion; and part 5 concludes.

II. LITERATURE REVIEW

Several research works were suggested in the literature related to the financial transformation decision model; a few recent works are reviewed here,

Cao and Wang [19] have introduced a stock price forecasting model that uses financial time series analysis and a modified convolution neural network. In order to predict the financial activity' future trend, a CNN is utilized for stock index forecasting. Developing an Index Prediction Model for CNN Stocks comes first, followed by the application of the CNN model algorithm and an analysis of the CNN model's structural parameter relationship. Second, a CNN-support vector machine (SVM) based stock index prediction model is construct, and the influence of model parameters on prediction outcomes is examined. The empirical analysis is finally finished.

Vo et al. [20] suggested the use of deep learning for portfolio optimization and decision-making in socially conscious investing. A 'Multivariate Bidirectional Long Short Term Memory Neural Network' is a component of the Deeply Responsible Investment Portfolio Model, which is used to predict stock returns while building a socially responsible investment portfolio. Neural networks are retrained using the deep reinforcement learning technique, and the portfolio is periodically rebalanced.

Paiva et al. [21] Have suggested a fusion technique of ML and portfolio selection for financial trading decision-making. This makes the model a special decision-making tool for day trading stocks on the stock market. It was created by combining a machine learning classifier with the MV technique for portfolio selection and the SVM method for classifiers. The Sao Paulo Stock Exchange Index's assets served as the basis for the experimental evaluation of the model. The best-performing parameter sets were selected for testing and analyzed using monthly rolling windows. The daily rolling windows with fresh classification algorithm training and portfolio optimization made up the monthly windows. There were 81 different parameter arrangements created .

Ren [22] have suggested using big data to optimize business financial management and decision-making processes. Internet finance firms should create their own integrated business finance system in response to industry need in order to accomplish business finance integration. This will improve enterprise management effectiveness, improve data matching between Internet finance enterprises, and improve business finance integration. Big data applications lead to increased financial gains in capital budgeting, production control, investment decision-making, and procurement management. The conclusion is that large amounts of data can be used to support enterprise decision-making in-depth in the big data era. This can help break down financial and business barriers, increase prediction and early warning capability, optimize organizational structure and personnel, and improve decision-making efficiency and quality. The utilization of big data tools is now crucial for improving corporate value and supporting financial decision-making.

Wang and Yu [23] showed the development and empirical study of the data mining algorithm-based financial early warning model for the company. The three data mining approaches that were used are introduced, along with the research methodology for the financial risk early warning model. 'Lastly, 77 listed manufacturing organizations and their matching companies that were first processed by ST in 2005–2007 were

used as research samples, based on the financial data of the 2.4 years before to being processed by ST and CXISP'. The financial risk early warning index system was developed by combining the concepts of financial risk early warning with data mining technology.

Yang [24] have created an AI analyses approach based on data mining technology, as well as financial big data administration and control. In order to help other businesses improve comparable financial systems, this work examines the company's intelligent financial reengineering approach. Financial processes can be revolutionized by combining IT technology with artificial intelligence (AI) capabilities like knowledge graphs, deep learning, human-computer interaction, data mining, image recognition, natural language processing, intelligent decision-making, and others. This can cut down on the amount of time it takes to process routine financial transactions, lessen the need for manual accounting, and increase the productivity of the finance department. The intelligentization of financial management is achieved and businesses are given access to more precise and efficient financial decision-making support through the self-directed analyses and decision-making of AI.

Wang and Guo [25] have suggested an intelligent method for enterprise cost accounting optimization that is based on data mining and big data. In addition to using big data technology to optimize management costs and apply innovative business management techniques, this article demonstrates how big data technology may be effectively applied to tackle the problem of enterprise cost management. This study elaborates on the goals, sources, and computation techniques of data mining while building a cost management model based on it. This model involved mining the company's data, coming up with a detailed plan, and suggesting an enhanced association algorithm to check for job completion and consolidation. This paper offers a novel approach to cost forecasting that utilizes a fuzzy model to effectively pick and combine data when multiple tasks are involved.

A. Motivation

There are three primary viewpoints for a framework in the financial sector. First, care should be taken while choosing a model and reducing the dimensionality of the data, as it is crucial to identify the factors that influence the analysis's outcome. As a result, wrapper techniques should be employed in conjunction with dimensional reduction algorithms. Second, it is appropriate to employ measurements that accurately reflect numerical assessments of the performance of the financial markets. It may be more useful in some circumstances to estimate the range of a target measure of interest rather than determining exact values. Because of the transaction cost, selling a stock at a specific moment can, for example, be based on a range that includes the target price rather than the precise amount. Therefore, the financial market needs a framework that makes use of a metric that shows how accurate the target value can be predicted as well as an evaluation meter that takes the target value's tolerances into account. Very few approach-based works have been provided in the literature to address this issue; these drawbacks and issues are what have inspired this study effort.

III. PROPOSED METHODOLOGY

In this section, financial transformation using Sparse Oblique Trees Algorithm model along with Waterwheel Plant Algorithm (SOTA-WWPA) is discussed. Fig 1 displays Block diagram of proposed methodology .It contains three stages, like Data acquisition, Data pre-processing, and Classification. Consequently, each step's full description is provided below.

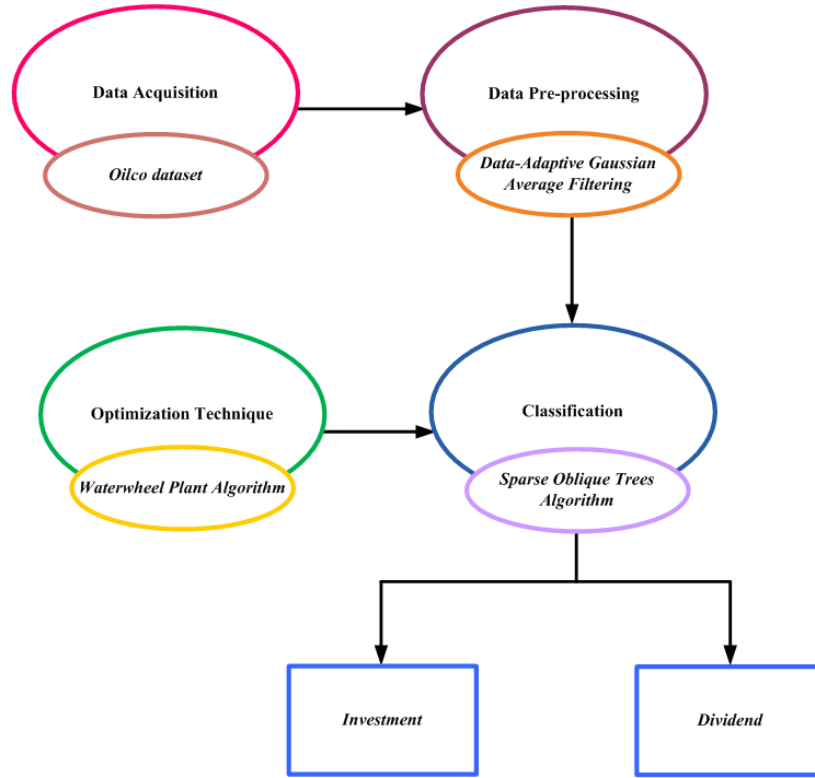


Fig 1: Block diagram of proposed Methodology

A. Data Collection

Anonymized here as "Oilco," a multinational energy and petrochemical firm with over 95,000 uses across more than 70 country's , is the single case multi-national organization dataset from which the data was collected [26].

B. Pre-processing using Data-Adaptive Gaussian Average Filtering (DAGAF)

In this section, the input data are pre-processed utilizing DAGAF [27]. It resolves imbalance ratio through duplicating instances in minority class. Next, the mean function of the lower and upper envelopes—the instantaneous mean—is eliminated from the data. This filtering approach is moving average process performed to data, average filter weights are created using a Gaussian window. Gaussian function with a standard deviation on its distribution is given in Eqn (1).

$$L_q(\varpi) = \rho \cdot (2\Omega)^{1/2} \cdot b^{-\frac{(\rho\varpi)^2}{2}} \tag{1}$$

Where ρ is the analogous continuous-time Fourier transform (CTFT). ‘Gaussian function’ in the discrete time domain becomes L_x assuming sample interval ρ equal to 1, Ω is integer which is followed by Fourier transform shown in Eqn (2).

$$l[G] = l[-G] = b^{-\frac{1}{2}\alpha^2 < 0.05} \tag{2}$$

Where, the Gaussian distribution's standard deviation and the parameter α are inversely related. Making values at end points (G) lesser than 5% of maximal window value yields parameter α in this research. The maximum may be seen at b , when $L[0]$ has a value of 1, and is shown in the following Eqn (3).

$$L(\varpi) = \frac{G}{\alpha} (2\Omega)^{1/2} \cdot b^{-\frac{1}{2}\left(\frac{G}{\alpha}\varpi\right)^2} \tag{3}$$

The spectrum obtained from above equation has a bell-shaped center at $L(\varpi)$ can be considered as a low pass filter when given the reasonable values G and α . DAGAF uses normalized discrete truncated Gaussian window for average filtering the spatial data, provided in order to preserve the data during the filtering process and given in Eqn (4).

$$l_M[g] = \frac{l[g]}{\sum_{w=-G}^G L[w]} \quad (4)$$

Where, the above equation defines Gaussian average filter and it is in fact low pass filter. Where, is then assuming Gaussian average filter $l_M[g]$ has been found and that the missing temporal h to be analyzed is represented as $L[w]$. The moving-average technique that follows can be used to get the instantaneous mean in DAGAF and shown in Eqn (5).

$$g_i[h] = \sum_{g=-G}^G l_M[g] \cdot r[g+h], \quad \text{for } 0 \leq h \leq H-1 \quad (5)$$

But there is still more work to be done on this equation. The algorithm must first take into account how the Missing Temporal and Spatial Data are corrected, as the raw data $r[g+h]$ is only specified in the interval $0 \leq h \leq H-1$. The value of G is equal to the length of extension across every boundary. The processed data in one decomposition iteration is indicated as $l_M[g]$, while the data to be analyzed is represented by $g_i[h]$. This method involves first extending the data by a "reflection" extension, and then reflecting the expanded segment in a filtering process with regard to a fixed value. Then the contemporaneous mean in DAGAF is then obtained by extending the data using the following Eqn (6).

$$g_i[h] = \sum_{g=-G}^G l_M[g] \cdot r_b[g+h], \quad \text{for } 0 \leq h \leq H-1 \quad (6)$$

Due to the symmetric structure of the Gaussian average filter $l_M[g]$, the instantaneous mean calculation in equation is essentially the convolution total of $r_b[\cdot]$. The Missing Temporal hidden in the traffic data will be recovered in the data domain by directly multiplying the spectrum of $r_b[\cdot]$ and the data bank of $l_M[\cdot]$.

The data are transmitted to SOTA for classification following noise reduction.

C. Classification Using Sparse Oblique Trees Algorithm (SOTA)

Let y be a vector of k softmax values representing the class distribution given y . Let us consider a trained deep net $x = g(y)$ that is used to classify an input instance $y \in M^G$ into several categories [28]. $z = G(y) \in M^G$ are the deep net features and $x = h(z)$ is the classifier layer that consists of the remaining net1. Write the net ' $g(y) = h(G(y))$ ' as the composition of a feature-extraction layer G . Every neuron in that layer may be thought of as a feature detector that encodes a property or concept of the input pattern y . This information can be used to support or refute one or more classes when combined with the concepts of other neurons. In contrast, will usually be more interested in features at an intermediate layer. This covers the class label or softmax outputs and the raw inputs as specific examples of features.

Imagine you have an input instance dataset $\{(y_n, x_n)\}_{n=1}^N \subset M^G \times \{1, \dots, k\}$ with labels (this is the one used to train the net). Next:

1. Utilizing the training set $\{(G(y_n), x_n)\}_{n=1}^N \subset M^g \times \{1, \dots, k\}$, train a sparse oblique tree $x = L(z)$ using TAO. Examine the trade-off between interpretability and accuracy within a meaningful range of the sparsity hyperparameter $\lambda \in [0, \infty]$, then select a preferred tree. Typically, this will be a sparse tree with almost the highest validation accuracy.

2. Examine the tree to find interesting trends about the deep net.

A tree that is as simple as feasible and accurately replicates the deep net is the aim. The tree can be trained using the same training set as the net to achieve this.

A feature extraction part $z = g(y)$ and a classifier part $x = h(z)$ combine to form the neural net $x = g(x) = h(G(y))$. For instance, the first four layers (convolutional and subsampling) of the LeNet5 neural net shown in the diagram correspond to these, and the final two levels are fully-connected. The activations (outputs) of g neurons create up the "neural net feature" vector z , which is representative of features that the neural net derived from the initial features y . By training the tree with the neural net features z as input and the

matching ground-truth labels as output, one can simulate the classifier portion $x = h(z)$ using a sparse oblique tree.

Step 2 is purposely vague. The tree contains valuable information on the meaning of features and their impact on classification, both for specific input instances and broadly.

Finally, SOTA is processed to classify the data based on investment and dividend.. The optimization procedure must optimize the weight parameter g of SOTA since SOTA does not offer optimization strategies for determining the optimal variables for verifying precise detection.

D. Waterwheel Plant Algorithm (WWPA)

The large petioles of waterwheel plants (*Aldrovandavesiculosa*) have traps that resemble tiny, transparent flytraps [29]. In order to guard against damage or accidental triggers from other water plants, a ring of hair-like bristles encircles the trap. ‘The outside edges of the trap are coated in many hook-like teeth that, like flytrap teeth, interlock when the trap closes around its prey. A particular kind of carnivorous plant known as a "flytrap" has developed into a crafty and extremely skilled insect hunter. Roughly forty long trigger hairs are responsible for the clamshell closing (a Venus flytrap trap contains just about 6–8 trigger hairs). To help with meat digestion, predators have glands that secrete acid and trigger hairs. The victim is drawn into the trap by the interlocking teeth and mucus sealant, which pins them to the floor at the hinge. The majority of the water has been driven out of the trap and is being replaced by digestive juices.

Step 1: Initialization

Initialize the input parameters to their initial values. The SOTA weight parameters, represented by g , are the input parameters in this case..

Step 2: Random Generation

In a matrix that depicts, the input parameter is generated at random.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,d} \\ x_{2,1} & x_{2,2} & \dots & x_{2,d} \\ \dots & \dots & \dots & \dots \\ x_{n,1} & x_{n,2} & \dots & x_{n,d} \end{bmatrix} \tag{7}$$

Where, X is the population of waterwheel location, n indicates the count of waterwheels, and d is the number of variables.

Step 3: Fitness Calculation

Using F to determine the fitness value

$$Fitness\ function = Optimizing(g) \tag{8}$$

Step 4: Exploration Phase: Position Identification and Hunting of Insects

With their excellent sense of smell, waterwheels can identify the exact location of pests, making them formidable predators. Any bug inside the waterwheel's attack range will be targeted. Once it has found its victim, it attacks and keeps searching. WWPA simulates this waterwheel action in order to mimic the first step of its population updating procedure. Enhance WWPA's exploration ability to find the ideal region and escape from local optima by mimicking the waterwheel's attack on the insect, which causes significant fluctuations in the waterwheel's location in the look for space. Eqn (9) is utilized to determine the waterwheel's new location by simulating its approach to the insect.

$$w = v_1 \cdot (X(t) + 2R) \tag{9}$$

$$X(t+1) = X(t) + w \cdot (2R + v_2) \tag{10}$$

Where w indicates a vector representing the diameter of the circle the waterwheel plant will search for possible locations, R is an exponential variable with values in the range $[0, 1]$, and $X(t)$ indicates the current solution at iteration t . v_1 and v_2 indicates the random variables with values among $[0, 2]$ and $[0, 1]$.

If after three iterations the solution doesn't change, Eqn (11) can be used to modify the waterwheel's location.

$$X(t+1) = Gaussian(\mu p, \sigma) + v_1 \left(\frac{X(t) + 2R}{w} \right) \tag{11}$$

Step 5: Exploitation Phase: Carrying the Insect in the Suitable Tube

A waterwheel draws in an insect, which is then moved to a feeding tube. ‘The 2nd stage of the WWPA population update is informed by this simulated waterwheel operation. Better solutions are found close to the ones that have already been found, and the WWPA's ability to exploit the local search is strengthened by the model of moving the insect to the proper tube, which causes minor adjustments to the location of the waterwheel in the search space’. For every waterwheel in the population, the WWPA's designers determine a fresh, arbitrary location that is a "excellent position for devouring insects" in order to mimic the natural behavior of waterwheels.

$$w = v_3 \cdot (RX_{best}(t) + v_3 X(t)) \tag{12}$$

$$X(t+1) = X(t) + R w \tag{13}$$

Where, v_3 indicates a variable that is random and has values between 0 and 2, X_{best} indicates the best solution.

Like in the exploratory stage, the following change is applied to ensure that local minima are avoided if the solution does not improve after three iterations:

$$X(t+1) = (v_1 + R) \sin\left(\frac{E}{B} \theta\right) \tag{14}$$

Where, E and B are the arbitrary variables with values between -5 and -5.

Furthermore, using the following equation, the value of falls exponentially’:

$$R = \left(1 - \frac{2 * t^2}{T_{max}} + E\right) \tag{15}$$

Step 6: Update the Best Solution

In each space, the algorithm explores new solutions, evaluates their fitness or quality, and updates the current solution if a better one is found.

Step 7: Termination

If the best solution is identified after reviewing the stopping criteria, the process ends; if not, proceed to step 3. There

IV. RESULT AND DISCUSSION

The experimental outcome of the assessment model construction of international students’ intercultural adaptation technique using STOA-WWPA method is covered in this part. The simulations are done in python working platform. The Python platform is used to simulate the proposed strategy under various performance criteria. Outcome of STOA analyzed with existing approaches such as, CNN, MBLSTMNN, and SVM.

A. Performance measures

This is an important step in choosing the optimal classifier. The performance is evaluated by looking at performance indicators like recall, sensitivity, F1-score, accuracy, and precision. The confusion matrix is considered in order to scale the performance indicators. ‘The True Positive (TP), and True Negative (TN), False Negative (FN) and False Positive (FP) values are required in order to scale the confusion matrix’.

1) Accuracy

It is the ratio of the total number of predictions made for a dataset to the number of exact predictions made. It is calculated using Equation (16).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{16}$$

2) F1-Score

F1-score is a metric used to analyze the performance of proposed FT-SOTA-WWPA technique. It is computed in Eqn (17),

$$F1score = \frac{TP}{TP + \frac{1}{2}[FP + FN]} \tag{17}$$

3) Precision (P)

A measure called precision counts how many of a positive prediction were executed correctly. This is computed via following Eqn (18)

$$P = \frac{TP}{TP + FP} \tag{18}$$

4) Recall (R)

A measure of prediction accuracy called recall determines the number of accurate forecasts there were overall. By using Equation (19), it is measured,

$$R = \frac{TP}{TP + FN} \tag{19}$$

B. Performance Analysis

Fig 2 to 6 shows the simulation outcomes of FT-SOTA-WWPA. Then the outcomes are analyzed with existing FT-CNN, FT-MBLSTMNN, and FT-SVM methods.

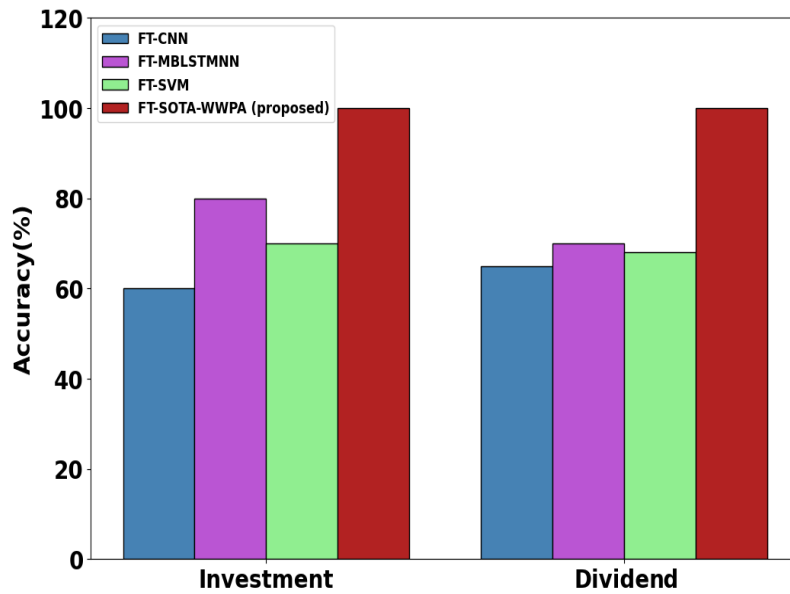


Fig 2: Accuracy value comparison with proposed and existing methods

Accuracy value comparison with proposed and existing methods is displays in Fig 2. The performance of the proposed FT-SOTA-WWPA technique results in accuracy that are 50.56%, 20.76%, 35.97% higher for the classification of language proficiency, 20.46%, 35.58%, 23.54% higher for the classification of cultural awareness when evaluated to the existing FT-CNN, FT-MBLSTMNN, and FT-SVM models correspondingly.

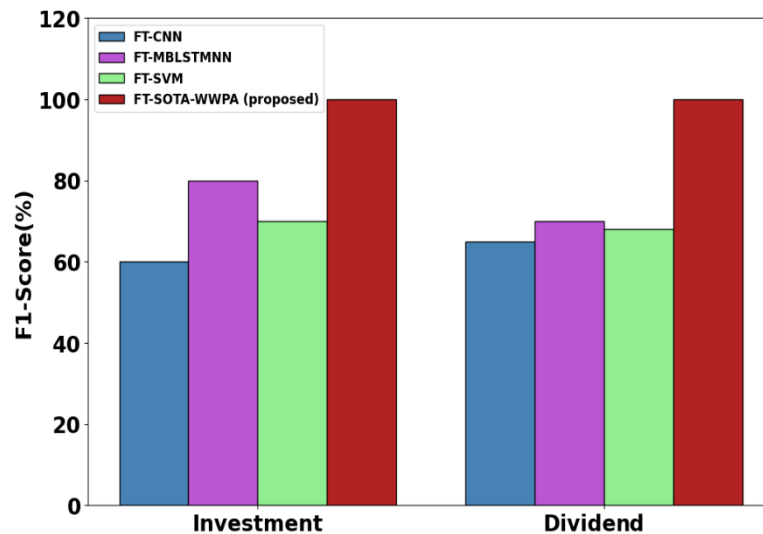


Fig 3: Analysis the performance of F1-score value with proposed and existing method

Analysis the performance of F1-score value with proposed and existing method is displays in Fig 3. The performance of the proposed FT-SOTA-WWPA technique results in F1-score that are 22.56%, 21.76%, 33.97%, higher for the classification of language proficiency, 21.46%, 33.58%, 23.54% higher for the classification of cultural awareness when evaluated to the existing FT-CNN, FT-MBLSTMNN, and FT-SVM models correspondingly.

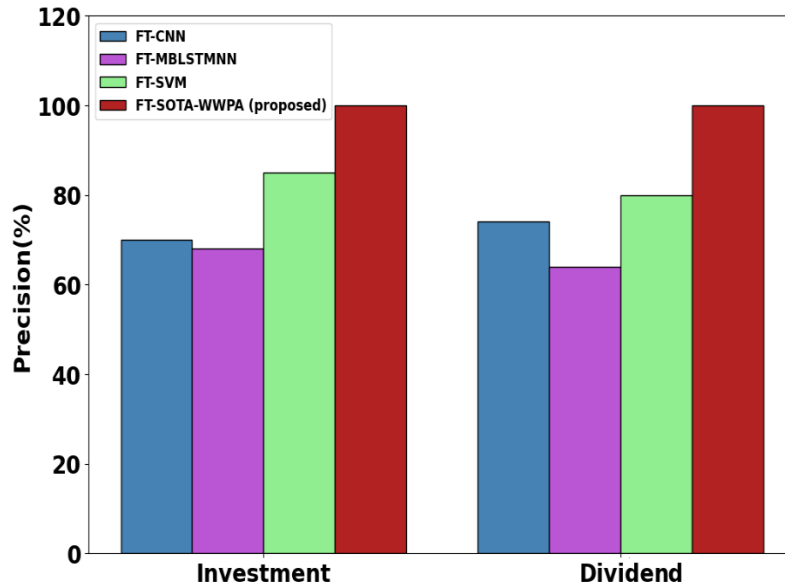


Fig 4: Comparison of precision value with proposed and existing method

The comparison of precision value with proposed and existing system is displays in Fig 4. Here, a direct comparison with proposed approaches is offered to show how the suggested method’s precision is higher. The proposed method provides for a more extensive analysis of a proposed and has higher precision than existing methods due to its wider consideration of factors. The performance of the proposed FT-SOTA-WWPA technique results in precision that are 30.56%, 21.76%, 35.97%, higher for the classification of language proficiency, 21.46%, 33.58%, 23.54% higher for the classification of cultural awareness when evaluated to the existing FT-CNN, FT-MBLSTMNN, and FT-SVM models correspondingly.

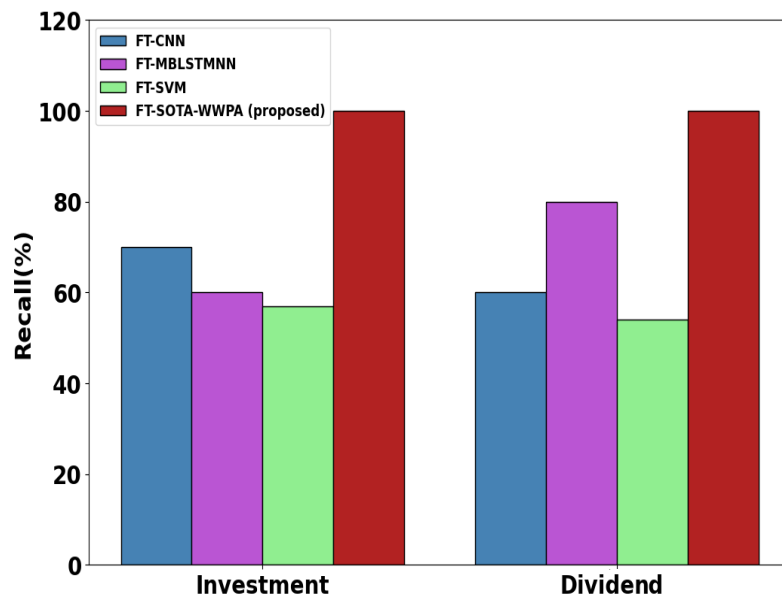


Fig 5: The recall value comparison between the proposed and existing systems

The recall value comparison between the proposed and existing systems is displays in Fig 5. The performance of the proposed FT-SOTA-WWPA technique results in recall that are 23.56%, 22.76%, 31.97% higher for the classification of language proficiency, 22.46%, 31.58%, 22.54% higher for the classification of cultural awareness when evaluated to the existing FT-CNN, FT-MBLSTMNN, and FT-SVM models correspondingly.

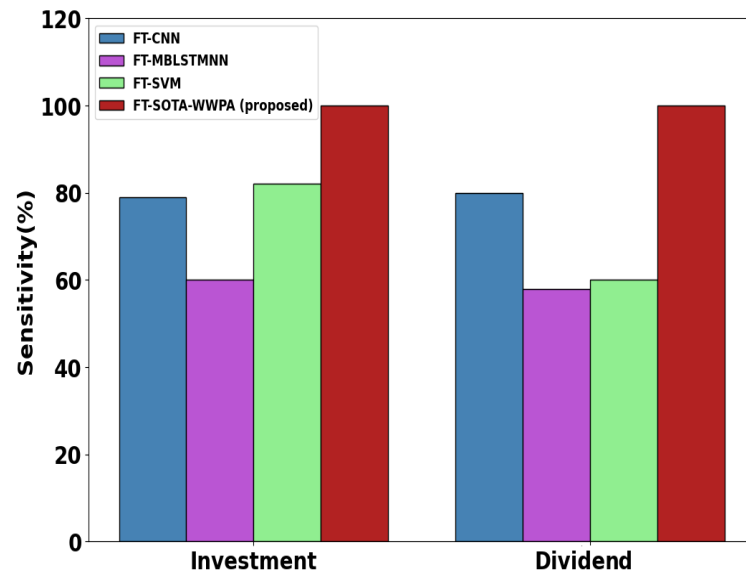


Fig 6: The sensitivity value comparison between the proposed and existing systems

The sensitivity value comparison between the proposed and existing systems is displayed in Fig 6. The performance of the proposed FT-SOTA-WWPA technique results in sensitivity values that are 30.56%, 21.76%, 35.97%, higher for the classification of language proficiency, 21.46%, 33.58%, 23.54% higher for the classification of cultural awareness when evaluated to the existing FT-CNN, FT-MBLSTMNN, and FT-SVM models correspondingly.

V. CONCLUSION

In conclusion, this research harnesses the power of Sparse Oblique Trees Algorithm to significantly model the construction of financial transformation. An essential first step in data is collected through the multi-national organization dataset. The Data-Adaptive Gaussian Average During pre-processing, filtering is utilized to process the finance data. The Sparse Oblique Trees Algorithm receives the pre-processed result and uses it to efficiently classify the investment and dividend. The suggested method is assessed using the Python working platform and contrasted with other methods that are currently in use. The suggested method is examined in several scenarios, including those involving accuracy, precision, sensitivity, F1-score, and memory. Its accurate identification of the recommended optimal region expanding technique has increased the accuracy of the DAGAF classifier to 98%, which is utilized in the last phase of the system and shows that it can correctly identify investment and dividend.

REFERENCES

- [1] Liu, B. (2021, April). Research on Enterprise Financial Management Supervision Mode Based on Data Mining. In *Journal of Physics: Conference Series* (Vol. 1881, No. 4, p. 042074). IOP Publishing.
- [2] Huang, J., & Zhang, X. (2023). Dynamic Allocation Method of Incentive Pool for Financial Management Teaching Innovation Team Based on Data Mining. *International Journal of Advanced Computer Science and Applications*, 14(7).
- [3] Guo, Y. (2022). Design of University Financial Information System Based on Data Mining Technology and IoT. *Wireless Communications and Mobile Computing*, 2022.
- [4] Guo, Y. (2021). CNS: Interactive Intelligent Analysis of Financial Management Software Based on Apriori Data Mining Algorithm. *International Journal of Cooperative Information Systems*, 30(01n04), 2150008.
- [5] Gao, B. (2021). The use of machine learning combined with data mining technology in financial risk prevention. *Computational Economics*, 1-21.
- [6] Yi, X. (2023). Application of data mining in enterprise financial risk prediction based on genetic algorithm and linear adaptive optimization. *Soft Computing*, 27(14), 10305-10315.
- [7] Li, K., & Chen, Y. (2022). Fuzzy clustering-based financial data mining system analysis and design. *International Journal of Foundations of Computer Science*, 33(06n07), 603-624.
- [8] Cheng, J. (2022). Financial data analysis and application based on big data mining technology. *Computational intelligence and neuroscience*, 2022.

- [9] Chang, Z., & Yang, H. (2021). Financial Management Early-Warning Mechanism Construction and Decision Analysis Research Based on Wireless Sensor Network and Data Mining. *Wireless Communications and Mobile Computing*, 2021, 1-13.
- [10] Kim, M. (2021). A data mining framework for financial prediction. *Expert Systems with Applications*, 173, 114651.
- [11] Shan, R., Xiao, X., Che, J., Du, J., & Li, Y. (2022). Data Mining Optimization Software and Its Application in Financial Audit Data Analysis. *Mobile Information Systems*, 2022.
- [12] Zhou, H., Sun, G., Fu, S., Liu, J., Zhou, X., & Zhou, J. (2019). A big data mining approach of PSO-based BP neural network for financial risk management with IoT. *IEEE Access*, 7, 154035-154043.
- [13] Feng, R., & Qu, X. (2022). Analyzing the Internet financial market risk management using data mining and deep learning methods. *Journal of Enterprise Information Management*, 35(4/5), 1129-1147.
- [14] Li, H. (2022). Analysis and design of financial data mining system based on fuzzy clustering. *Expert Systems*, e13031.
- [15] Li, W., Zhou, Q., Ren, J., & Spector, S. (2019). Data mining optimization model for financial management information system based on improved genetic algorithm. *Information Systems and eBusiness Management*, 1-19.
- [16] Song, Y., & Wu, R. (2022). The impact of financial enterprises' excessive financialization risk assessment for risk control based on data mining and machine learning. *Computational Economics*, 60(4), 1245-1267.
- [17] Prabagar, S., Tiwari, S. K., Jagannayaki, K., Jain, A., Singh, V., & Vidhya, K. (2023). Data mining based framework for organisational financial management. *Journal of Informatics Education and Research*, 3(2).
- [18] Jin, X., & Hu, H. (2022). Research and implementation of smart energy investment and financing system design based on energy mega data mining. *Energy Reports*, 8, 1226-1235.
- [19] Cao, J., & Wang, J. (2019). Stock price forecasting model based on modified convolution neural network and financial time series analysis. *International Journal of Communication Systems*, 32(12), e3987.
- [20] Vo, N. N., He, X., Liu, S., & Xu, G. (2019). Deep learning for decision making and the optimization of socially responsible investments and portfolio. *Decision Support Systems*, 124, 113097.
- [21] Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. *Expert Systems with Applications*, 115, 635-655.
- [22] Ren, S. (2022). Optimization of enterprise financial management and decision-making systems based on big data. *Journal of Mathematics*, 2022, 1-11.
- [23] Wang, A., & Yu, H. (2022). The construction and empirical analysis of the company's financial early warning model based on data mining algorithms. *Journal of Mathematics*, 2022.
- [24] Yang, N. (2022). Financial Big Data Management and Control and Artificial Intelligence Analysis Method Based on Data Mining Technology. *Wireless Communications and Mobile Computing*, 2022.
- [25] Wang, W., & Guo, J. (2022). Based on Data Mining and Big Data Intelligent System in Enterprise Cost Accounting Optimization Application. *Scientific Programming*, 2022.
- [26] Lambert, S. A., Herbert, I. P., & Rothwell, A. T. (2020). Rethinking the Career Anchors Inventory framework with insights from a finance transformation field study. *The British Accounting Review*, 52(2), 100862.
- [27] Lin, Y. D., Tan, Y. K., & Tian, B. (2022). A novel approach for decomposition of biomedical signals in different applications based on data-adaptive Gaussian average filtering. *Biomedical Signal Processing and Control*, 71, 103104.
- [28] Hada, S. S., Carreira-Perpiñán, M. Á., & Zharmagambetov, A. (2023). Sparse oblique decision trees: A tool to understand and manipulate neural net features. *Data Mining and Knowledge Discovery*, 1-40.
- [29] Abdelhamid, A. A., Towfek, S. K., Khodadadi, N., Alhussan, A. A., Khafaga, D. S., Eid, M. M., & Ibrahim, A. (2023). Waterwheel Plant Algorithm: A Novel Metaheuristic Optimization Method. *Processes*, 11(5), 1502.