ISSN: 2182-2069 / E-ISSN: 2182-2077

Optimized Cloud Security: An AI-based Data Classification Approach for Financial Cloud Computing

Dr. Pradeep Kanchan[1*], Divya Paikaray[2], Dr. Ashwini Malviya[3], Ramkumar Krishnamoorthy[4], Vivek Saraswat[5], and Sachin S. Pund[6]

[1*]Department of Computer Science and Engineering, NMAM Institute of Technology, Nitte (Deemed to be University), Nitte, Karnataka, India. pradeepkanchan@nitte.edu.in, https://orcid.org/0000-0002-9320-5288

[2]Assistant Professor, Department of Computer Science & IT, ARKA JAIN University, Jamshedpur, Jharkhand, India. divya.p@arkajainuniversity.ac.in, https://orcid.org/0000-0001-7886-1538

[3]Associate Professor, Department of uGDX, ATLAS SkillTech University, Mumbai, Maharashtra, India. ashwini.malviya@atlasuniversity.edu.in, https://orcid.org/0009-0006-9154-4725

[4]Assistant Professor, Department of Computer Science and Information Technology, Jain (Deemed to be University), Bangalore, Karnataka, India. ramkumar.k@jainuniversity.ac.in, https://orcid.org/0000-0003-1164-6504

[5]Centre of Research Impact and Outcome, Chitkara University, Rajpura, Punjab, India. vivek.saraswat.orp@chitkara.edu.in, https://orcid.org/0009-0000-6875-1255

[6]Assistant Professor, Mechanical Engineering, Ramdeobaba University, RBU, Nagpur, Maharashtra, India. pundss@rknec.edu, https://orcid.org/0000-0002-5616-2469

Received: January 17, 2025; Revised: February 26, 2025; Accepted: April 09, 2025; Published: May 30, 2025

Abstract

Data categorization is the technique for classifying and arranging facts based on their value and sensitivity, in addition to offering suitable control and safety precautions. Data categorization is critical in cloud computing for finance because it helps to defend cloud-based monetary data. In this research, we aim to create an Artificial Intelligence (AI)-based totally Data type method for economic cloud computing. We proposed an innovative set of rules named Dynamic Vortex Search-driven Light Gradient Boosting Machine (DVS-LightGBM) for enhancing the method for classifying data in monetary cloud computing. The text cleaning process was conducted to preprocess the collected raw data, and it enhances the quality of the data acquired. We applied Latent Semantic Analysis (LSA) to extract the essential traits from the processed data. The dynamic vortex technique improves accuracy and increases convergence in the data classification process. The data classification procedure is based on cloud sensitivity levels. Our recommended approach consists of three classes, including standard, sensitive and highly sensitive. We implemented our suggested model in Python software. The result analysis phase is done on various parameters such as precision, accuracy, F1 score and recall. We conducted a comparison analysis with other existing methodologies to examine the effectiveness of the suggested model. The results of the trial

Journal of Internet Services and Information Security (JISIS), volume: 15, number: 2 (May), pp. 75-87.
DOI: *10.58346/JISIS.2025.I2.006*

*Corresponding author: Department of Computer Science and Engineering, NMAM Institute of Technology, Nitte (Deemed to be University), Nitte, Karnataka, India.

demonstrate that the recommended strategy outperformed other existing methods in the data classification approach for financial cloud computing.

Keywords: Data Classification, Cloud Security, Artificial Intelligence (AI), Dynamic Vortex Search-driven Light Gradient Boosting Machine (DVS-LightGBM), Financial Cloud Computing.

1 Introduction

The most demanding platform is called cloud computing (CC) and it consists of many innovative technical apps that allow users to securely save, retrieve and preserve their data (Shafiq et al., 2021; Chowdary et al., 2024). Cellular devices, including smartphones, laptops and tablets, could be used to access cloud-based resources and apps. Regarding document security, cloud computing is regarded as the most dependable and safe platform (Ahmad et al., 2020). However, owing to limitations in battery life, performance and storage space, additional storage devices, such as laptops and mobile phones, cannot be able to provide such a safe platform for data storage (Chen et al., 2020).

Cloud storage solutions are widely used to swiftly, cheaply and easily preserve as well as backup any kind of data. They also make it easy to share data and synchronize devices (Vegesna, 2021). Several cloud computing architectural designs differ throughout system configurations, and cloud storage architecture does not have a uniform set of accepted standards (Sehgal et al., 2020). A common file system is used to join millions of storage units in a group, an internet connection and additional storage middleware to offer cloud storage services for customers (Gabi et al., 2020).

A distributed file system, cloud storage consists of storage groups, resource service relationships and service level agreements (SLAs) (Kondori & Peashdad, 2015; Gao, 2022). Delivering storage on request in an extremely adaptable, multi-tanned manner is the main objective of cloud storage solutions (Tan et al., 2022). An application programming interface (API)-equipped front end is a standard feature of cloud storage designs, allowing for storage access. This API, known as the SCSI protocol in conventional storage devices, is becoming increasingly common in the cloud (Surianarayanan & Chelliah, 2019). In this study, we enhanced the process of data classification in financial cloud computing by proposing an innovative DVS-LightGBM (Darapour, 2016).

The remaining study components could be classified. We will discuss the related works in Section 2. The conceptual frameworks are discussed in section 3. Experimental results are presented in Section 4. The last section of this paper, Section 5, is the conclusion.

2 Related Works

Swain et al., (2022) illustrated a data security technique that uses encryption and data categorization. They used the K-Nearest Neighbour (K-NN) classification technique to divide the data according to their needs, as it was inappropriate to implement safety measures before knowing the specifications of the content. The framework was intended to facilitate cloud services for powerful computing, even if its application range was extensive.

Asharaf et al., (2023) introduced a variety of resource-related AI technologies which have to be carefully adjusted when they are merged with large amounts of computer resources, particularly interconnected systems. Any intelligent Internet of Things (IoT) ecosystem in the real world is thought to place the highest priority on security and privacy. IoT-based networks' security vulnerabilities placed digital environments at risk for security breaches.

Thabit et al., (2021) suggested a novel, feasible cryptographic algorithm to boost data security, which could be used to safeguard cloud computing services. When compared to the encryption techniques, a high level of security, in addition to a noticeable reduction in parameters like cipher processor time and safety measures, was shown in the testing results of the recommended method, which was often employed in cloud computing (Falaki, 2019).

Zhang, (2023) suggested using a cloud platform to store company financial data in a safe, dynamic manner (Agnes Pravina et al., 2024). Initially, the grey correlation analysis technique was used to determine the absolute connection degree of the data in the financial database Almaliki & Al-saedi, (2023). Based on the outcomes of that calculation, data extraction and clustering were performed.

Balashunmugaraja & Ganeshbabu, (2022) provided a novel hybrid meta-heuristic idea for creating a corporate data security preservation plan in the cloud industry. The main objective of the study was to create a novel hybrid "red deer-bird swarm algorithm (RD-BSA)" that minimized the need for control parameters throughout the solution generation process while ensuring greater convergence.

Tabassum et al., (2023) suggested a unique hybrid strategy built on machine learning and natural language processing (NLP). The desired outcomes included bug prioritization using trained classifiers and multi-class supervised categorization to handle the concerns. The results of all trials demonstrated the superior performance of the balanced dataset in classification when compared to the unbalanced dataset models (ChithraDevi et al., 2024).

3 Methodology

We collected the "fraudulent transaction" database from Kaggle Depository. We suggested a novel technique called DVS-LightGBM to enhance the process of data classification in financial cloud computing. Preprocessing the raw data that was gathered through text cleaning improves the quality of the data that was obtained. Latent Semantic Analysis (LSA) was used to extract significant features from the processed data.

3.1 Dataset

Data were gathered the "fraudulent transaction" dataset from the Kaggle Depository to identify fraudulent transactions (https://github.com/dj2097/FraudDataDetection). There are 6362620 records in the collection, each with 10 characteristics. In this data collection, the greatest transaction reported amounts to 1991430 USD, while the mean value of every transaction is 144972 USD. The distribution of the financial value of every transaction remains strongly skewed toward as they could have already realized based on the mean and maximum. A fraction of transactions approaches the limit, and the great majority of transactions are quite minor.

3.2 Data Cleaning

To guarantee precision and consistency, the financial dataset is pre-processed. Detecting and correcting outliers resulting from bias evaluation, standardized data types for reliability and reducing numerical parameters to a standard scale are the initial stages in managing missing values. There are more procedures involved as well. To improve the accuracy of models and lower dimensionality, redundant or unnecessary features are removed, and classification variables are encoded into a numerical representation. Validation tests, such as cross-referencing against outside sources, guarantee data integrity. To evaluate the model, the dataset is finally divided into sets for both training and testing.

3.3 Feature Extraction

LSA is a feature extraction method utilized to enhance data classification in cloud computing for finance. Through examining the connections between sentences and texts in a substantial amount, LSA discovers hidden patterns that improve understanding of semantics. LSA makes it feasible to display financial statements more efficiently by decreasing the amount of data while preserving its theoretical foundation.

3.3.1 LSA

LSA is a method based on mathematical statistics to examine the language structure of a text. Each step in the LSA procedure is as follows:

(a) Construct a matrix B, where the cell accommodates the j frequencies of the common terms and j rows of the matrix include the exact phrases in each document. The document's caption is located in the column i.

(b) Singular value decomposition (SVD) is used with matrices B, causing the breakdown of the matrix into three forms: V, T, and U matrix.

$$B = V \times T \times U^S \quad (1)$$

Details:

$B \epsilon Q^{n,m}$

V: Matrix orthogonal, $V \epsilon Q^{n,\min(n,m)}$

T: Matrix diagonal, $T \epsilon Q^{\min(n,m)\min(m,n)}$

U: Matrix orthogonal, $U \epsilon Q^{n,\min(n,m)}$

(c) Decreases the matrix by keeping all of the rows in the first row and l columns V and U and the first rows and l columns t.

$$B_l = V_l \times T_l \times U_l^S \quad (2)$$

Matrix minimization criteria are denoted by l.

3.4 DVS-LightGBM

The unique DVS-LightGBM strategy attempts to enhance data classification in a financial cloud computing environment. Combining LightGBM with dynamic vortex detection, this technology offers outstanding efficiency. When analyzing intricate financial data. Considering the ability of DVS, the process of the data structure varies dynamically, making it more accurate and adaptable in applications involving categorization. Furthermore, LightGBM, which is well known for its adaptability and effectiveness, further optimizes the process. This allows for rapid and accurate distribution of financial data across cloud computing environments.

3.4.1 Dynamic Vortex Search

The VS algorithm generates possible solutions around a point. In the first iteration, the upper boundary of the current case is used to determine the central origin or μ_0 at this point. In subsequent iterations, the center is shifted to the best position observed so far. The VS method stops working at local minimum points for certain functions, as previously noted, due to this process.

It presents a dynamic VS algorithm (DVS) to address the aforementioned defects. Every time the DVS algorithm runs, numerous centers are the focus of potential solutions that are created. The DVS algorithm's search process could be conceptualized as many concurrent vortices with various centers at each iteration. The centers of these many vortices have to be selected in the same way as in the VS algorithm. Let's examine the total number of centers, also known as vortices, that n represents. Assume that s is the iteration index and

$N_s(\mu)$ is the matrix that holds the values of these n centers at each iteration pass. Therefore, the starting coordinates of these centers are estimated as in equation (3) and $N_0(\mu) = \{\mu_0^1, \mu_0^2, \cdots, \mu_0^k\}$ $k = 1,2, \cdots, n$.

$$\mu_0^1 = \mu_0^2 = \cdots = \mu_0^k = \frac{Upperlimit+lowerlimit}{2} \quad k = 1,2, \cdots, n \tag{3}$$

Next, using the starting radius value q_0Several possible solutions are constructed with a Gaussian distribution around these initial centers. Here, the total number of potential solutions is once again chosen to be m. Note, however, that these m solutions are produced in the vicinity of n centers. Therefore, around each center, m/n solutions must be chosen.

Let's state that the subset of solutions created around the center $k = 1,2, \ldots, n$ for the repetition s is represented by $DT_s^k(t) = \{t_1, t_2 \cdots, t_l\}$ $l = 1,2, \cdots \cdots m/n$. The whole set of solutions produced for the iteration t = 0 could therefore be expressed as follows: $D_0(t) = \{DT_0^1, DT_0^2, \cdots, DT_0^k\}$ $k = 1,2, \cdots \cdots n$. During the selection step, the best solution $t_k \in DT_0^1(t)$, is chosen for each subset of solutions. Make sure that the potential subsets of solutions are within the search bounds before moving on to the selection step. According to equation (3), the solutions that are beyond the borders are moved within them for this reason. Suppose that, for every iteration run, the best solution for each subset is kept in a matrix $OBest_s(t)$. $OBest_0(t) = \{t_1, t_2, \cdots t_k\}$ $k = 1,2, \cdots n$, thus for $s = 0$. Be aware that, for the current iteration, Itr_{best} represents the best solution of the whole candidate solution set $D_0(t)$, which is also the best solution for this matrix ($OBest_0(t)$). The starting point of the VS algorithm is moved to the best solution that is identified so far, sbest, during each iteration run. However, the DVS algorithm has n centers, the locations of which must be changed for the next iteration. That is where the most significant distinction between the VS and DVS algorithms appears. One of these centers is once again been moved to the best response so far, t_{best}, in the DVS algorithm. However, as illustrated in equation (4), the other $n - 1$ centers are moved to a new location based on the best locations created around each center at repetition s and the best position discovered to this point.

$$\mu_s^k = t_k + rand\ (t_1 + t_{best}) \tag{4}$$

A uniformly distributed random number, rand, is given by $k = 1,2, \cdots \cdots, n - 1$ and (

$t_1 \epsilon\ obest_{s-1}(t)$ in Equation (4). Consequently, for $s = 1$, $N_1(\mu) = \{\mu_1^1 \mu_1^2, \cdots \mu_1^k\}$ $k = 1,2, \cdots n - 1$ is obtained using the ($t_1 \epsilon\ obest_0(t)$ positions and the best position discovered so far, t_{best}.

3.4.2 LGBM

LightGBM is a gradient-based learning system that utilizes decision trees and boosting. It is explained in depth here due to its novelty. The primary distinction of this model from XGBoost is the use of histogram-based methods to reduce memory use, speed up training and apply a depth-limited leaf-wise development approach. The core principle of the histogram technique involves transforming continuous float-point eigenvalues into a discrete representation by partitioning them into k bins. Subsequently, a histogram is generated with a dimension of k. The histogram-based approach doesn't need extra space

for pre-sorted results. Additionally, it can retain the value following the process of discretizing features, which are commonly kept as an 8-bit integer. This significantly decreases memory usage to one-eighth of the initial size. The algorithm's efficiency remains unaffected by this rudimentary differentiation.

Considering the limited effectiveness of the decision tree as a research model, the correctness of the categorization point is irrelevant. The segmentation points with a less detailed level have a regularization effect, thus reducing the risk of overfitting. The growth methodology of a hierarchy of choices is usually known as layer-wise, however, it is considered ineffective because it handles parts on a single degree in a manner that leads to unnecessary storage use. A more efficient approach is to prioritize leaf-wise traversal, where the leaves with the greatest branching benefit are identified and traversed before moving on to the next branching cycle. Thus, when compared to the horizontal direction, the blade can minimize a greater number of mistakes and achieve superior precision with an equivalent number of segmentations. Leaf orientation has a drawback in that it can lead to the growth of more complex decision trees, which in turn can result in overfitting. LGBM incorporates a maximum depth constraint at the leaf node to minimize overfitting and maintain high efficiency. Fig.1 displays the schematic diagram illustrating the strategies for tree growth, specifically the level-wise and leaf-wise approaches.

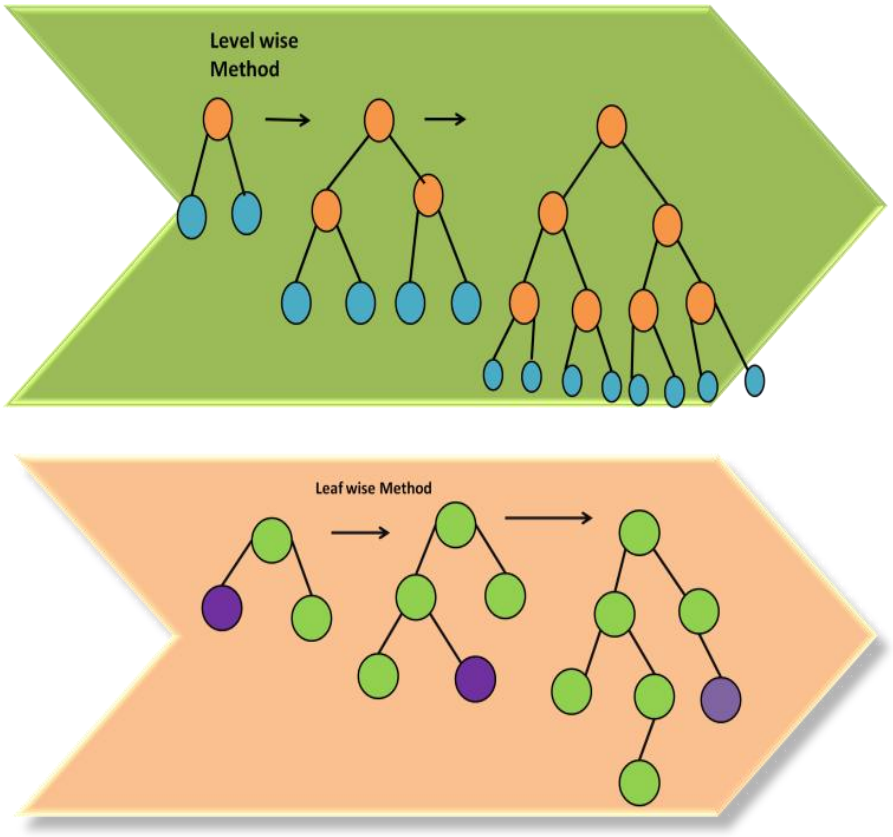

Figure 1: Strategies for Tree Growth

(source: https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2FLevel-wise-tree-growth-VS-leaf-wise-tree-growth_fig3_362591649&psig=AOvVaw39oerOChtILMIvOGLIDSHj&ust=1711007445515000&source=images&cd=vfe&opi=89978449&ved=0CBQQ jhxqFwoTCMDM5vKtgoUDFQAAAAAdAAAAABAQ)

The computation procedures of LGBM are discussed in detail as follows: Given the training dataset $W = \{(w_j, z_j)\}_{j=1}^{n}\}$ finding an approximate solution is the goal of LGBM $\hat{e}(w)$ to the function $e^*(w)$ to minimize the predicted values of certain losses. Life $K(z, e(w))$ LGBM will incorporate several tree regressions.

$$\hat{e}(w) \arg\min_{e} F_{z,W}\, K(z, e(w)) \tag{5}$$

LGBM utilizes a collection of S regression trees to approximate the completed model, which is shown as $\sum_{s=1}^{S} e_s(W)$.

$$e_S(W) = \sum_{s=1}^{S} e_s(W) \tag{6}$$

Regression trees are defined as a set of trees, denoted as w_{qw}, where $q \in \{1,2,..M\}$ each tree has a decision rule q and a vector w representing the sample weights of the leaf nodes. The total number of tree leaves is represented by M. The experiment is presented in a combination method each step s as follows:

$$\Gamma_s \cong \sum_{i=1}^{M} K\left(z_j, E_{s-1}(w_j) + e_s(w_j)\right) \tag{7}$$

The goal function is efficiently calculated using Newton's method. Equation (8) is simplified by eliminating the variable term.

$$\Gamma_s \cong \sum_{i=1}^{M} \left(h_j e_s(w_j) + \frac{1}{2} g_j e_j^2(w_j)\right) \tag{8}$$

Where $h_j$and g_j represent the statistical outcomes of loss operates' 1st and 2nd order inclinations. If a sample set of leaf j is represented by J_i, Equation (7) is subsequently transformed into Equation (9).

$$\Gamma_s \cong \sum_{i=1}^{M} \left(\left(\sum_{j\in J_i} h_j\right)\omega_i + \frac{1}{2}\left(\sum_{j\in J_i} h_j + \lambda\right)\omega_i^2\right) \tag{9}$$

The equations (10) and (11) are used to determine the optimal leaf weights ω_i^* and extreme values for Γ_s^* in the framework of the tree q_w.

$$\omega_i^* = \frac{\sum_{j\in J_j} h_j}{\sum_{j\in J_j} h_j + \lambda} \tag{10}$$

$$\Gamma_s^* = -\frac{1}{2}\sum_{i=1}^{I} \frac{\left(\sum_{j\in J_j} h_j\right)^2}{\sum_{j\in J_j} h_j + \lambda} \tag{11}$$

The weight function q_w quantifies the overall quality of the tree structure. The objective function is eventually achieved by including the division procedure.

$$H = \frac{1}{2}\left(\frac{\left(\sum_{j\in J_j} h_j\right)^2}{\sum_{j\in J_j} h_j + \lambda} + \frac{\left(\sum_{j\in J_q} h_j\right)^2}{\sum_{j\in J_q} h_j + \lambda} + \frac{\left(\sum_{j\in J} h_j\right)^2}{\sum_{j\in J} h_j + \lambda}\right) \tag{12}$$

Where J_j and J_q represent samples from the left and right branches, correspondingly.

4 Experimental Results

The proposed solution is executed using Python version 3.10.1 on a laptop running Windows 10, provided with an Intel i7 core processor and 8GB of RAM. Utilize libraries such as Scikit-Learn or Tensor Flow/Keras to train our suggested model on the training data. The existing methods are the K-

Nearest neighbor (KNN) classifier (Kumar et al., 2023), support vector machine (SVM) classifier (Kumar et al., 2023) and the proposed method is DVS-LightGBM. The performance of these methods is examined based on accuracy, execution time, precision and recall.

Reliability of financial data classification into pre-established groups or categories is referred to as accuracy. By accurately classifying data points according to different characteristics or properties, it assesses whether the classification strategy works.

$$Accuracy = \frac{TP+FP}{(TN+TP+FN+FP)} \quad (13)$$

Figure 2 and Table 1 show the accuracy performance. The proposed DVS-LightGBM system achieves an accuracy of 94%, outperforming the existing systems KNN classifier, and SVM classifier, which has 90 % and 84 %, respectively. As a result, the proposed scheme is more accurate than existing data distribution techniques for financial cloud computing.

Table 1: Values for Accuracy, Execution Time, Precision, and Recall (Source: Author)

Methods	Accuracy (%)	Execution Time (Second)	Precision (%)	Recall (%)
SVM classifier	84	0.05	87	86
KNN classifier	90	0.02	92	93
DVS-LightGBM [Proposed]	94	0.011	95	96

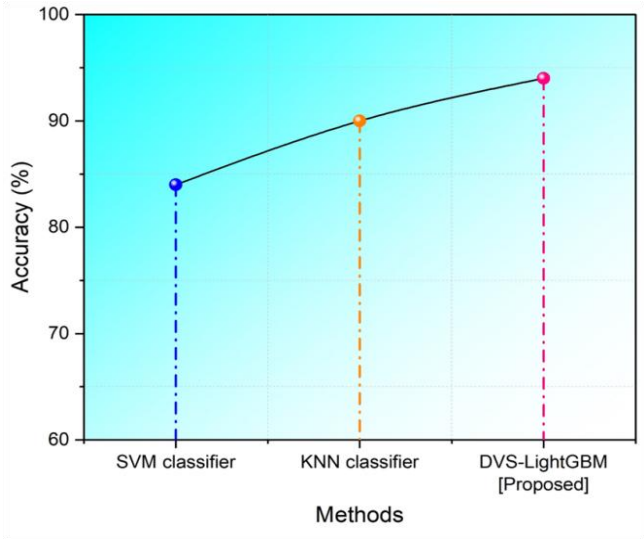

Figure 2: Accuracy Performance (Source: Author)

In the discussion of the data classification approach for financial cloud computing, time consumption refers to the amount of time required for classification algorithms to search and organize financial data in a cloud computing system. This is an important exception importance of the efficiency and effectiveness of the distribution system.

Figure 3 and Table 1 illustrate execution time overall performance. In comparison to the existing methods KNN, SVM classifiers, which have 0.02s, 0.05s, the proposed technique, DVS-LightGBM, received 0.011s. The suggested method consumes less execution time when compared to the existing methods used for data categorization in financial cloud computing.

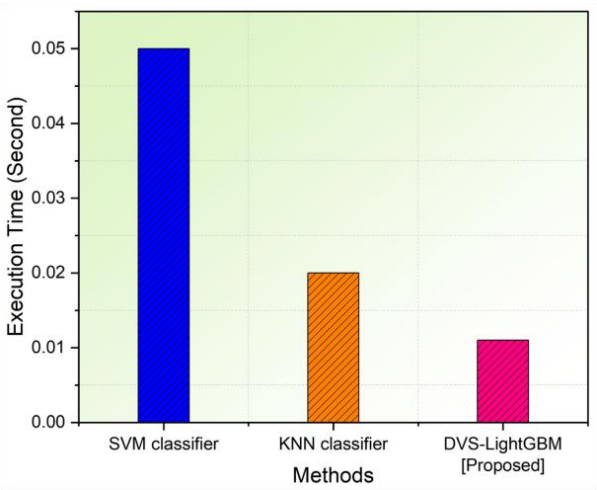

Figure 3: Execution Time Performance (Source: Author)

When it comes to data categorization for Financial Cloud Computing, precision refers to the method's correctness and dependability in recognizing and classifying various financial data types. It entails reducing false positives and false negatives while optimizing the accurate categorization of private data.

$$\text{Precision} = \frac{\text{TP}}{TP+FP} \tag{14}$$

The precision performance is shown in Table 1 and Figure.4. The accuracy of DVS-LightGBM of the proposed method is 95%, which is higher than the KNN and SVM classifiers of the existing methods, which have 92 % and 87%. Thus, the proposed scheme exhibits higher performance compared to existing data categorization methods in financial cloud computing.

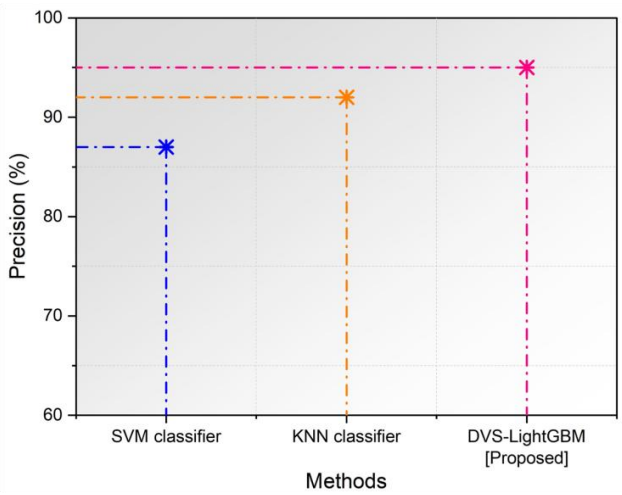

Figure 4: Precision Performance (Source: Author)

The term "data partitioning methodology for financial cloud computing" describes the process of classifying data according to relevance, sensitivity, and regulatory requirements in a financial cloud environment. This method classifies data and determines data-based criteria by pre-defined types, including availability, integrity, and confidentiality.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{15}$$

Table 1 and Figure 5 show the recall patterns. Compared with the existing algorithms, KNN and SVM classifiers with 93% and 86%, respectively, the recommended DVS-LightGBM algorithm has 96%, and as a result, the method shows better performance as compared to the existing methods for data categorization in financial cloud computing.

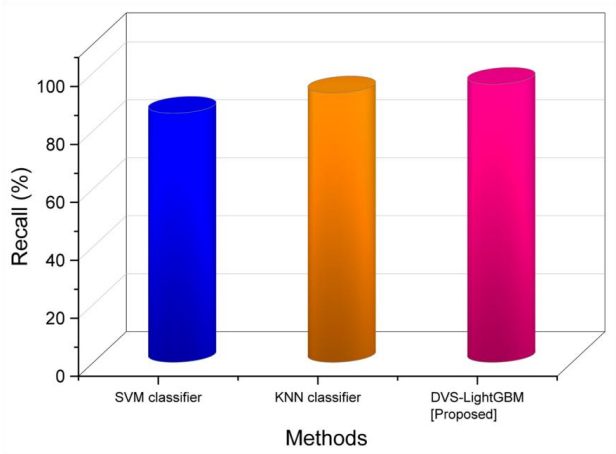

Figure 5: Recall Performance (Source: Author)

5 Conclusion

Data categorization was the process of grouping and arranging statistics consistent with their usefulness and severity while offering appropriate management and safety controls. In this study, research created an economic cloud computing statistics categorization method that was based on AI and operates successfully. To improve the data categorization approach in financial cloud computing, this study recommended a novel approach referred to as DVS-LightGBM. The collected raw data was pre-processed using text cleaning, which improves the quality of the data that was ultimately retrieved. Research extracted the important characteristics from the processed data using LSA. The effectiveness of the recommended method was analyzed by means of accuracy (94%), execution time (0.011s), precision (95%) and recall (96%). Robust privacy in financial computing environments could require ongoing monitoring and improvement due to reliance on high-quality data and possible biases in AI models. Accuracy and flexibility could be improved by integrating new developments in AI, and scaling to bigger datasets remains an exciting possibility.

References

[1] Agnes Pravina, X., Radhika, R., & Ramesh Palappan, R. (2024). Financial Inclusiveness and Literacy Awareness of Fisherfolk in Kanyakumari District: An Empirical Study. *Indian Journal of Information Sources and Services, 14*(3), 265–269. https://doi.org/10.51983/ijiss-2024.14.3.34

[2] Ahmad, N., Hoda, N., & Alahmari, F. (2020). Developing a cloud-based mobile learning adoption model to promote sustainable education. *Sustainability, 12*(8), 3126. https://doi.org/10.3390/su12083126

[3] Almaliki, O. J., & Al-saedi, M. O. (2023). The Impact of the Qualitative Peculiarities of Accounting Information Based on the Financial Reports of Commercial Banks. *International Academic Journal of Social Sciences, 10*(1), 49–56. https://doi.org/10.9756/IAJSS/V10I1/IAJSS1006

[4] Asharaf, Z., Ganne, A., & Mazher, N. (2023). *Artificial Intelligence in Cloud Computing Security. Research Gate.* https://www.doi.org/10.56726/IRJMETS33029

[5] Balashunmugaraja, B., & Ganeshbabu, T. R. (2022). Privacy preservation of cloud data in business application enabled by multi-objective red deer-bird swarm algorithm. *Knowledge-Based Systems, 236,* 107748. https://doi.org/10.1016/j.knosys.2021.107748

[6] Chen, R. C., Dewi, C., Huang, S. W., & Caraka, R. E. (2020). Selecting critical features for data classification based on machine learning methods. *Journal of Big Data, 7*(1), 52.https://doi.org/10.1186/s40537-020-00327-4

[7] ChithraDevi, S. A., Mahendravarman, I., Ragavendiran, A., Selvam, M. M., Yamuna, K., & Vibin, R. (2024, July). Optimal configuration of radial distribution networks with stud krill herd optimization. *In 2024 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT)* (pp. 1-6). IEEE.

[8] Chowdary, P. B. K., Udayakumar, R., Jadhav, C., Mohanraj, B., & Vimal, V.R. (2024). An Efficient Intrusion Detection Solution for Cloud Computing Environments Using Integrated Machine Learning Methodologies. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15*(2), 14-26. https://doi.org/10.58346/JOWUA.2024.I2.002

[9] Darapour, M. (2016). The role of toxic assets in the formation of financial crisis (Case study: Iran central bank's monetary system). *International Academic Journal of Organizational Behavior and Human Resource Management, 3*(2), 1–7.

[10] Falaki, A. A. (2019). A Proposed Framework to Use Cloud Computing Power in Business Intelligence. *International Academic Journal of Science and Engineering, 6*(1), 85–89. https://doi.org/10.9756/IAJSE/V6I1/1910008

[11] Gabi, D., Ismail, A. S., Zainal, A., Zakaria, Z., Abraham, A., & Dankolo, N. M. (2020). Cloud customers service selection scheme based on improved conventional cat swarm optimization. *Neural Computing and Applications, 32,* 14817-14838. https://doi.org/10.1007/s00521-020-04834-6

[12] Gao, J. (2022). Analysis of enterprise financial accounting information management from the perspective of big data. *International Journal of Science and Research (IJSR), 11*(5), 1272-1276. https://doi.org/10.21275/SR22514203358.

[13] Kondori, M. A., & Peashdad, O. H. (2015). Analysis of challenges and solutions in cloud computing security. *International Academic Journal of Innovative Research, 2*(1), 20–30.

[14] Kumar, A., Khan, S. B., Pandey, S. K., Shankar, A., Maple, C., Mashat, A., & Malibari, A. A. (2023). Development of a cloud-assisted classification technique for the preservation of secure data storage in smart cities. *Journal of Cloud Computing, 12*(1), 92. https://doi.org/10.1186/s13677-023-00469-9

[15] Sehgal, N. K., Bhatt, P. C. P., & Acken, J. M. (2020). Cloud computing with security. Concepts and practices. *Second edition. Switzerland: Springer.* https://doi.org/10.1007/978-3-030-24612-9

[16] Shafiq, D. A., Jhanjhi, N. Z., Abdullah, A., & Alzain, M. A. (2021). A load balancing algorithm for the data centres to optimize cloud computing applications. *IEEE Access, 9,* 41731-41744. https://doi.org/10.1109/ACCESS.2021.3065308

[17] Surianarayanan, C., & Chelliah, P. R. (2019). Storage Fundamentals and Cloud Storage. In Essentials of Cloud Computing: A Holistic Perspective (pp. 133-182). *Cham: Springer International Publishing.* https://doi.org/10.1007/978-3-031-32044-6_5

[18] Swain, T., Singh, A., Verma, K., Sahoo, A. K., Samaddar, S. G., & Barik, R. K. (2022, November). Machine learning based data classification methods in cloud security using cloudlightning framework. *In 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 55-60). IEEE.

[19] Tabassum, N., Namoun, A., Alyas, T., Tufail, A., Taqi, M., & Kim, K. H. (2023). Classification of Bugs in Cloud Computing Applications Using Machine Learning Techniques. *Applied Sciences, 13*(5), 2880. https://doi.org/10.3390/app13052880
[20] Tan, W., Zhu, H., Tan, J., Zhao, Y., Xu, L. D., & Guo, K. (2022). A novel service level agreement model using blockchain and smart contract for cloud manufacturing in industry 4.0. *Enterprise Information Systems, 16*(12), 1939426. https://doi.org/10.1080/17517575.2021.1939426
[21] Thabit, F., Alhomdy, S., Al-Ahdal, A. H., & Jagtap, S. (2021). A new lightweight cryptographic algorithm for enhancing data security in cloud computing. *Global Transitions Proceedings, 2*(1), 91-99. https://doi.org/10.1016/j.gltp.2021.01.013
[22] Vegesna, V. V. (2021). Analysis of Data Confidentiality Methods in Cloud Computing for Attaining Enhanced Security in Cloud Storage. *Middle East Journal of Applied Science & Technology, 4*(2), 163-178.
[23] Zhang, L. (2023). The dynamic and secure storage of enterprise financial data based on cloud platform. *International Journal of Critical Infrastructures, 19*(5), 449-462. https://doi.org/10.1504/IJCIS.2023.133285

Authors Biography

Dr. Pradeep Kanchan is working as Assistant Professor in the Department of Computer Science and Engineering, NMAM Institute of Technology, Nitte (Deemed to be University). He has done his BE from NMAMIT, MTech from NITK, Surathkal and PhD from NITK, Surathkal. His areas of interest include: Wireless Sensor Networks, Quantum Computing and Nature Inspired Algorithms. He has a rich teaching experience of 25 years.

Divya Paikaray is an Assistant Professor in the Department of Computer Science & IT at ARKA JAIN University, Jamshedpur, Jharkhand, India. Her areas of interest include computer networks, software development, and emerging technologies. She is actively involved in academic research and student mentoring.

Dr. Ashwini Malviya is an Associate Professor in the Department of uGDX at ATLAS SkillTech University, Mumbai, Maharashtra, India. She specializes in design thinking, innovation, and interdisciplinary education. With a strong academic and creative background, she actively contributes to curriculum development and research in design education.

Ramkumar Krishnamoorthy is an Assistant Professor in the Department of Computer Science and Information Technology at Jain (Deemed-to-be University), Bangalore, Karnataka, India. His academic and research interests include data analytics, machine learning, and cloud computing. He is actively involved in teaching, research, and mentoring students in cutting-edge technologies.

Vivek Saraswat is affiliated with the Centre of Research Impact and Outcome at Chitkara University, Rajpura, Punjab, India. His work focuses on enhancing the visibility, quality, and societal relevance of academic research. He is actively engaged in supporting research initiatives and fostering collaborative projects.

Sachin S. Pund has received B.E. in Industrial Engineering from Shri Ramdeobaba College of Engineering and Management, Nagpur University, in 2000 and M. Tech degree in Industrial Engineering from Shri Ramdeobaba College of Engineering and Management, Nagpur University, 2008. He has awarded with Ph.D. degree at the Department of Mechanical Engineering, in 2024 from SSSUTMS, Sehore, Bhopal (MP), India. His research interests include Optimization Techniques, Facilities Planning, Operations Management and Project Management.