ESP-JETA
ESP Journal of Engineering & Technology Advancements
ISSN: 2583-2646 / Volume 6 Issue 1 February 2026
Paper Id: JETA-V6I1P110 / Page No: 79-88

Original Article

An Intelligent Machine Learning Framework for Optimizing Identity and Access Management (IAM) Policies in Cloud Infrastructure

Jiwan Prakash Gupta
Sr Software Engineer, Davita Inc.

Received Date: 06 February 2026 Revised Date: 15 February 2026 Accepted Date: 19 February 2026

Abstract: *Identity and Access Management (IAM) is an important factor in ensuring secure and effective access control in cloud computing environments. This work presents a sequential machine-learning-based model to optimize IAM policies using a high-dimensional Cloud Access Control Parameter Management dataset. The paradigm incorporates systematic preprocessing of data, feature engineering, label encoding, feature selection using Boruta, feature scaling, and hybrid data balancing, 80: 20 train -test split and 5-fold cross-validation. Together with various ML algorithms, such as Random Forest Classifier (RFC), LightGBM, Gradient Boosting Classifier (GBC), XGBoost, and a soft Voting ensemble, multiple machine learning models are trained and tested based on the standard performance metrics and ROC analysis. The experiment's findings show that the Voting Classifier has the best accuracy (98.19), followed by LightGBM (98.01), RFC (97.24), and XGBoost (90.95). The results indicate that ensemble-based and boosting models offer strong, precise, and generalized IAM security predictions, and hence the suggested framework could be effectively used to improve control over access to clouds and security policy issues.*

Keywords: *IAM, Cloud, Voting Classifier, Artificial Intelligence, Machine Learning.*

I. INTRODUCTION

Personal data has become an economic asset and a lucrative resource in recent years. The significance of data regulation is evident in the advent of data protection laws. Digital identity is the representation of an entity as one or more characteristics. An individual may maintain one or more digital identities. Consumers are becoming more aware of how much data organizations store about them in recent years [1], [2]. There is increasing demand for greater transparency in data collection, use, and security.

Organizational data management, storage, and accessibility techniques have been significantly transformed by the widespread adoption of cloud computing across industries. The foundation for comprehending the implications of this change for security and data governance may be traced back to early studies on cloud computing [3], [4]. One significant study emphasized the unique security issues that arise when businesses transition to cloud environments, particularly regarding data privacy and the need for effective access control to mitigate risks in cloud storage environments.

IAM is a crucial component of an organization's overall IT security infrastructure, since it oversees the management of digital identities and user access [5]. A hybrid method for industry-specific access control has been developed within the environment of information and communication technology (ICT) [6], [7]. A framework for corporate procedures that makes managing electronic identities easier is the IAM system [8]. The technology required to provide identity management is part of the framework. The automated process of establishing records and managing user identities and the associated access permissions is made possible by IAM technology [9], [10]. This guarantees that all services and persons are appropriately verified, approved, and audited, and that access rights are issued in accordance with corporate policy [11]. IAM policies are set up by the organization's IT administrators, who often lack access to automated decision support tools to optimize them. Therefore, since IAM policy optimization is inherently complicated, human error and security lapses may often result in incorrectly designed IAM policies and data breaches [12]. An emerging field with many opportunities and threats is the use of AI into IAM systems, especially in cloud environments. As cloud computing becomes more commonplace, the relevance of AI's involvement in IAM is growing [13]. ML and AI have emerged as key technologies in cybersecurity, offering sophisticated tools for analyzing complex data and identifying patterns that are difficult to detect with conventional rule-based systems. AI and ML approaches may be used in IAM to improve access control, strengthen authentication, and facilitate proactive threat detection [14]. The following are the primary contributions of this paper:

This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/2.0/)

- Created a complete ML model to optimize the policy of IAM processes on clouds through the use of preprocessing, feature engineering, Boruta-based feature selection, and hybrid data balancing, and ensemble learning.
- High dimensional data of cloud access control and high imbalance between the classes are effectively managed, leading to better quality of data, robustness of the model, and generalization.
- Carried out an extensive analysis of various models (RFC, LightGBM, GBC, XGBoost, and Voting) that showed that the best results were achieved using ensemble techniques.
- Reached high predictive accuracy with the Voting Classifier and the strong ROC-AUC and cross-validation outcomes confirming the validity of the suggested strategy.

A. Justification and Novelty

The current study is based on the fact that the complexity and size of cloud environments continue to increase, and that rule-based IAM mechanisms are not adequate to support high-dimensional access control policies and dynamically changing security threats. This novelty is supported by the integration of an ordered, fully automated machine learning pipeline to optimize IAM policies using Boruta-based feature selection, hybrid data balancing, and highly regularized ensemble models. This framework is proven to be effective in achieving reliable and scalable IAM security assessment in a cloud infrastructure unlike the other methods, which are based on single models or limited preprocessing, providing high accuracy and robustness as well as interpretability.

B. Organization of the paper

The paper's framework is as follows: Section II provides an overview of the literature on IAM in the cloud; Section III delves into the dataset and methods; Section IV presents the experimental results and analysis; and finally, Section V concludes with key findings and potential directions for future study.

II. LITERATURE REVIEW

The field of IAM is widely recognized as a foundational element of enterprise security, with its traditional role encompassing the authentication and authorization of digital identities.

Bora, Silva Weber and Zincir-Heywood (2025) present these techniques use Machine Learning (ML) models to analyze encrypted traffic metadata. The ML models are trained on data generated by our scalable, cloud-native Android traffic-generation framework. For its primary task, our method, using emulated traffic from eight IMAs, yielded a 98.6% F1-score for application identification with Gradient Boosting [15]. AbouElabbas, Abdel-Gawad and Fahmy (2025) propose a tabular deep learning-based approach that utilizes TabNet, a deep learning architecture for tabular data, to automatically generate access control decisions. TabNet was evaluated on the Amazon Kaggle and Amazon UCI datasets, achieving an F-score of over 96% on both. These results outperform the state-of-the-art [16].

Demirsoy et al. (2024) introduce an AI-driven IAM system that enhances security protocols through real-time anomaly detection. By leveraging a hybrid architecture consisting of CNN and LSTM layers, the system provides real-time analysis of user behavior to detect identity-related anomalies. The proposed model achieved test Acc of 85.44%, Prec of 87.95%, Rec of 85.44%, and AUC score of 0.8578 [17]. Sivaraman (2023) introduces a Zero Trust IAM paradigm for multi-cloud environments that includes features such as identity verification, unified governance, continuous authentication within Watchful Security Zones (WSZs), and micro-segmentation to isolate different parts of identities. It creates a common identity layer across cloud providers for operational convenience, security, and compliance [18].

Van Ede et al. (2022) provide a new method for identifying misconfigurations in AWS's IAM policies. Real IAM policy data from three business cloud environments are used to evaluate our approach. They evaluate the effectiveness of our method for detecting misconfigurations and show that, although it has a somewhat lower precision than rule-based systems, it can correctly identify between 3.7 and 6.4 times as many misconfigurations [19].

III. METHODOLOGY

This research suggested implements an ordered machine learning pipeline to optimize IAM policies in the cloud. Once all data is collected, feature engineering and label encoding are used for preprocessing, followed by an 80:20 train-test split and 5-fold cross-validation. Boruta-based feature selection, feature scaling and hybrid data balancing are used to enhance quality of learning. Several models such as the RF, LightGBM, GB, XGBoost and a Voting Classifier, are trained and tested based on the conventional performance measurements as well as ROC analysis to get valid security predictions. The whole process is shown in Figure 1.

A. Data Collection

The Cloud Access Control Parameter Management dataset is a high-dimensional tabular one with 51 records and 100,000 attributes, which is the different cloud access control and security configuration parameters. It is mostly used to analyze and evaluate access control mechanisms, permission settings, and security policies in cloud computing systems.

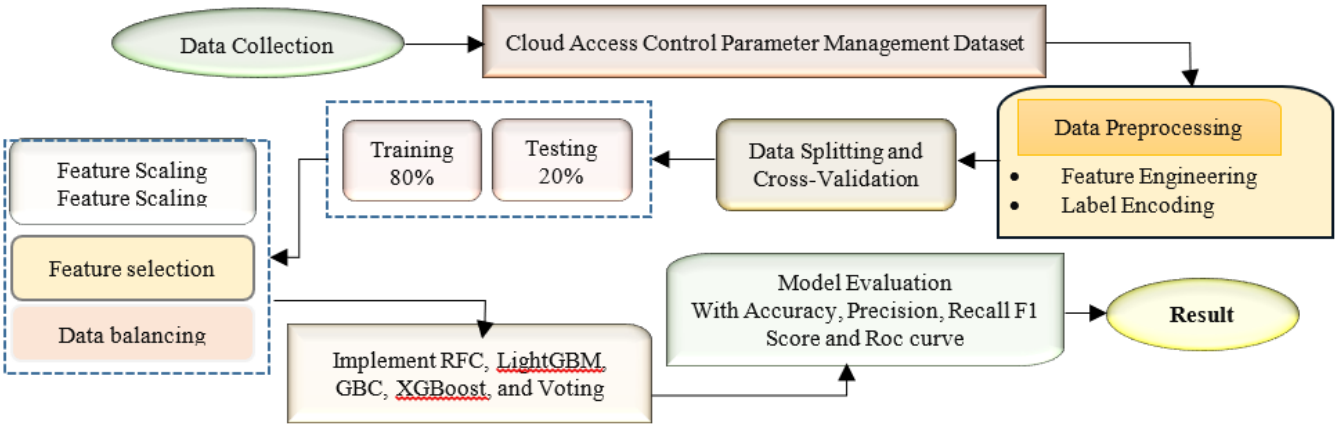

Figure 1 : Flowchart for IAM Policies in Cloud Infrastructure using Machine Learning.

Each step of proposed flowchart of methodology is briefly explained below:

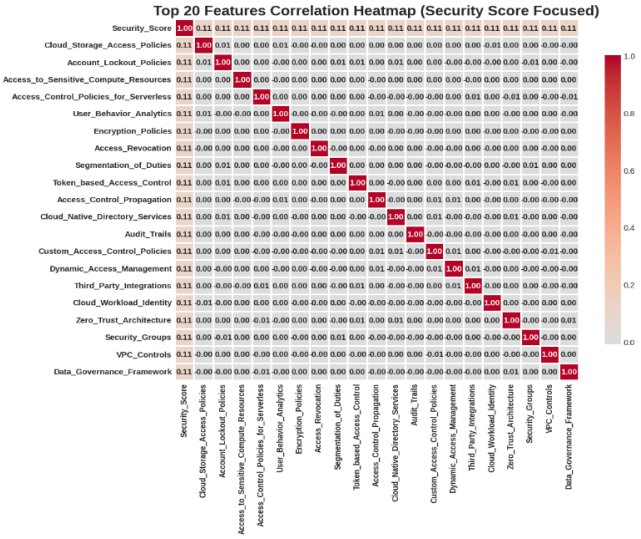

Figure 2 : Correlation Heatmap of Top 20 Features with 'Security Score Focused.

A heatmap of the top 20 features with the strongest correlation with the Security Score is shown in Fig. 2. The diagonal values indicate perfect self-correlation, whereas the off-diagonal values indicate low correlation between features, suggesting low multicollinearity. Attributes like access control policies, authentication mechanisms, encryption policies, monitoring and governance-related attributes are positively correlated with the Security_Score indicating their relevance in the context of determining the overall IAM security strength.

B. Data Preprocessing

The Preprocessing of data consisted of thorough analysis of the data to make sure that it is of good quality and is appropriate to feed the machine learning. The number of rows and columns was determined, and all feature names were displayed, along with missing values and data types. Lastly, duplicate records were verified to assess data cleanliness and reliability, and an appropriate, model-ready dataset was the outcome.

C. Feature Engineering

To transform the IAM configuration attributes (which are in the form of a Boolean) into meaningful security-based metrics, feature engineering was done. Once all the features of the boolean type were converted to numerical forms, the related controls were summed to form composite indicators such as Total_Security_Features, Total_Access_Controls, and Total_Features Enabled. The domain-specific scores were also created in order to reflect management, authentication strength, monitoring, data protection, network security, identity, and compliance. These aggregated characteristics streamline complex IAM environments to a higher level.

D. Label Encoding

All categorical (object-type) features are converted into numeric representations that can be used in machine learning by using the label encoding. A separate encoder is used to encode each feature and is stored so it can be decoded later if needed. Also, all the unspent attributes that are of a type of a boolean are transformed into integer values. This means the entire dataset is numeric and prepared for effective training and model evaluation.

E. Data partitioning and Cross-Validation

The data is divided into training (80%) and testing (20%) sets to evaluate model performance on unseen data. Also, 5-fold cross-validation has been used in training and average cross-validation accuracy is reported to give a robust and reliable performance estimate.

F. Feature Scaling using StandardScaler

The StandardScaler is used to scale features by centering them and scaling them to unit variance. The scaler will be attached on training data to avoid data leaking and applied on both training and testing data. The results in scaled form are returned to a DataFrame format to maintain names of features and indexing. Such normalization facilitates consistent feature effects and enhances the stability and performance of ML models.

G. Feature selection with Boruta Algorithm

The selection of features is done with the help of the Boruta algorithm, which is a wrapper-based algorithm based on a Random Forest classifier, to determine the most significant IAM features in prediction of security. Boruta repeatedly compares original features with randomized shadow features and classifies them as confirmed, tentative, or rejected based on their importance. Using a Random Forest with a fixed depth and a specified number of estimators is intended to produce stable importance estimates. Confirmed and tentative features are maintained in the modeling process to maintain informative characteristics which leads to a smaller and more meaningful feature set that increases the robustness of the model, its interpretability and generalization capabilities.

H. Data Balancing

In order to deal with the issue of class imbalance in the security level labels, a hybrid data balancing approach was adopted with the use of the imblearn library. The first step was to use Random Under Sampler to halve the number of majority, retaining all critical patterns with minimum dominance bias. Then, the samples of the minority class were enlarged with the help of RandomOverSampler to the size of the diminished majority one. These two steps were combined into an imblearn Pipeline to ensure a uniform, repeatable resampling strategy. Following the hybrid balancing, the security classes were balanced with the same amount of samples and the model training was unbiased and had equal stability.

```
After Hybrid Balancing:
 Security_Level
0    32892
1    32892
2    32892
Name: count, dtype: int64
```

Figure 3 : Class Distribution of After Hybrid Balancing.

Figure 3 shows the class distribution of the Security Level variable after applying the hybrid balancing technique. It shows a perfect balance among the three classes (0, 1, and 2), with each class containing the same number of samples.

I. Machine Learning Model Design

It provides the theoretical explanation of the ML methods for IAM Policies in Cloud Infrastructure.

a) Random Forest Classifier

A kind of ensemble learning, the Random Forest model constructs a number of decision trees during training and uses them to determine the distribution of classes during classification [20]. It efficiently serves as a way of representing intricate interactions

between features and also minimizes overfitting than having an individual decision tree. In this model, 50 decision trees (n_estimators = 50) are used to achieve stable ensemble learning at a moderate computational cost. The maximum depth of each tree has been set to 3 (max depth = 3) which restricts the complexity of the trees and the model does not memorize the training data. In order to promote further generalization, a node may only split when it has at least four samples (min_samples_split = 4), and a leaf node must have at least two samples (min_samples_leaf = 2).

b) LightGBM

It is a boosting approach that uses tree-based learning methods, which are considered as a particularly effective processing method [21]. It is believed to be an effective processing technique. The LightGBM approach develops vertically, which means it grows leaf-wise whereas other algorithms grow level-wise, in contrast to other algorithms, which build their trees horizontally. It runs 15 boosting iterations (nestimators = 15) with exceptionally low learning rate 0.0005, which allows the learning to go smooth and steady. The complexity of the trees is highly restricted by limiting the maximum depth of the tree to 1 and the amount of leaves to 12, and to improve the consistency of the splits, the number of samples per leaf must be at least 30 (min_child_samples = 30).

c) Gradient Boost Classifier

The Gradient Boosting classifier is configured with conservative hyperparameters to enhance generalization and reduce overfitting [22]. This model has 15 boosting stages (n_estimators = 15) that have very small learning rate of 0.0005, which allows slow learning and stability. Complexity of the tree is strictly controlled by the maximum depth to 1, and robustness is enhanced by subsampling (subsample = 0.3) and feature selection (max features = 0.3) both of the training instances and features respectively.

d) XGBoost

XGBoost is a versatile and highly accurate GB system that pushes the limits of processing power for boosted tree algorithms. It may be used to computing infrastructure and offers the benefits of enhancing the algorithm and changing the model [23]. It is used to solve issues such as regression and classification. The model uses 15 boosting rounds (n_estimators = 15) with a small learning rate of 0.004, allowing it to be optimized gradually. The complexity of the trees is tightly regulated by putting a maximum depth of 1 and robustness is improved by using 30% row subsampling (subsample = 0.3) and 30% column subsampling per tree (colsample_bytree = 0.3). In order to prevent splits with a lot of noise, a minimum child weight of 30 (minimum child weight = 30) and a gamma of 5 are used, making conditions of splits stricter.

e) Voting

A soft Voting Classifier is constructed using the RF, LightGBM, Gradient Boosting, and XGBoost models, with highly regularized hyperparameters to improve generalization. Random Forests splits 20 shallow trees (depth = 3) using controlled splits, LightGBM splits 15 depth-1 trees with a very low learning rate and strong L1/L2 regularization, Gradient Boosting splits 15 estimators with subsampling and limited features and XGBoost splits 15 boosting rounds with strict split constraints, subsampling, and regularization with multi-class classification.

J. Model Evaluation

At this stage, we will train and evaluate each model using the same training and test sets, and compare it with other popular regression and classification models. The following lists the descriptions of each performance metric.

Accuracy (Acc) is the percentage of correct predictions, while precision (Prec) is determined by the ratio of true positive predictions to the sum of true positive and false positive predictions. Recall (Rec) quantifies the model's capacity to recognize every actual positive instance [24], [25]. F1score offers a balance between Prec and Rec, calculated as the harmonic mean of Prec and Rec. The performance metrics formulae in Equations (1) to (4):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - Score = 2 * \frac{(Precision * Recall)}{Precision + Recall} \quad (4)$$

AUC: where $ROC(x)$ illustrates the correlation between the TPR and FPR at the threshold value of x. It is derived in equation (5).

$$AUC = \int_0^1 ROC(x)dx \quad (5)$$

A smaller number of True Negatives (TNs) indicates that negative instances were correctly classified as negative, whereas a larger number of True Positives (TPs) indicates that positive instances were correctly detected. False Positives (FPs) are those that are incorrectly categorized as positive, while False Negatives (FNs) are those that are erroneously categorized as negative.

K. Experimental Setup

The hardware specifications for the machine used in the studies include Windows 10 Pro, 16.0 GB of RAM, and an Intel (R) Core (TM) i5 CPU operating at 3.20 GHz. A version of Python 3.7.3 based on the Jupyter Notebook (Anaconda3) software standard has been used to build the proposed model. The model has undergone pre-processing using several tools, such as sklearn and pandas.

IV. RESULT ANALYSIS AND DISCUSSION

In Table 1 is the experimental performance of the proposed machine learning models of IAM in cloud infrastructure.

Table 1 : Experiment Results of Proposed Models for IAM in Cloud Infrastructure

Model	Accuracy	Precision	Recall	F1 Score	CV accuracy
RFC	97.24	99.89	97.24	98.49	98.90
LightGBM	98.01	99.89	98.01	98.89	99.27
GBC	96.03	99.89	96.03	97.86	97.43
XGBoost	90.95	98.34	90.95	94.30	96.22
Voting	98.19	99.89	98.19	98.98	99.26

The Voting Classifier has the highest overall accuracy of 98.19, most importantly it has a high prec of 99.89, rec of 98.19, F1score of 98.98 and a high cross-validation (CV) accuracy of 99.26 which indicates it is able to generalize. LightGBM model provides also competitive results, achieving the acc of 98.01 and F1score of 98.89, and the best CVaccuracy of 99.27. RFC has the best performance with an acc of 97.24 and F1score of 98.49 and GBC has an accuracy of 96.03. Conversely, XGBoost has a relatively low performance with an acc of 90.95% and F1score of 94.30. Overall, the findings indicate that ensemble models, specifically Voting and LightGBM, are very promising for IAM classification in the cloud environment.

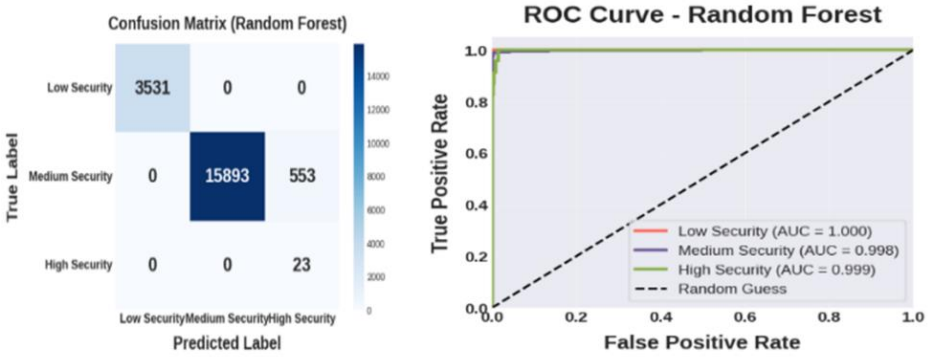

Figure 4 : Confusion Matrix and ROC Curves of the Random Forest Classifier.

Figure 4 demonstrates the RFC's performance using the confusion matrix and ROC curves for the Low, Medium, and High security classes. Based on the confusion matrix, the model correctly classifies 3,531 cases of Low Security with none case misclassified as High Security, 15,893 cases of Medium Security with 553 cases misclassified as High Security and 23 cases of High Security with no misclassification as the rest of the classes thus, the model has very reliable separation of the classes. The ROC analysis also indicates the strength of the model with AUC of 1.000 in Low Security, 0.998 in Medium Security, and 0.999 in the High Security, which are much higher than the random guess baseline.

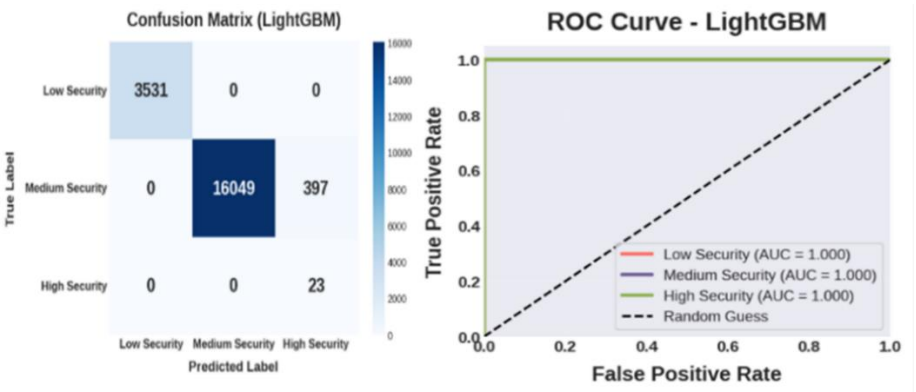

Figure 5 : Confusion Matrix and ROC Curves for Lightgbm.

In Figure 5 shows that the LightGBM model correctly classifies 3,531 Low Security, 16,049 Medium Security, and 23 High Security instances, with only 397 Medium Security samples misclassified as High Security and no other errors. The ROC analysis achieves a perfect AUC of 1.000 for all three classes, demonstrating excellent discriminative performance and high reliability of the LightGBM classifier.

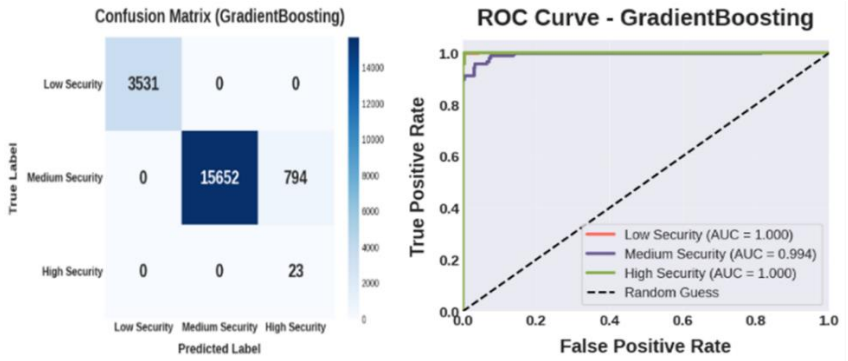

Figure 6 : ROC Curves and Confusion Matrix of the Gradient Boosting Classifier.

Figure 6 shows Gradient Boosting model performs, with 3,531 Low Security, 15,652 Medium Security, and 23 High Security instances correctly classified, and 794 of the Medium Security instances, and no other errors in classification. The analysis of ROC shows that Low security has an AUC of 1.000, Medium Security has 0.994 and High Security has 1.000 meaning that it has strong discriminative ability and acceptable classification performance at all levels of security.

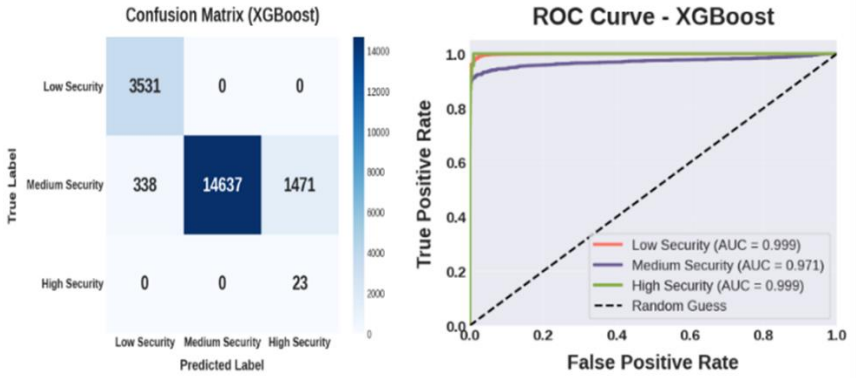

Figure 7 : ROC Curves and Confusion Matrix of the Xgboost Classifier

Figure 7 shows that the XGBoost model correctly labels 3531 Low Security, 14637 Medium Security, and 23 High Security samples, and 338 Medium Security samples are incorrectly labeled as Low Security, and 1471 High Security samples as Medium Security. The ROC analysis provides AUC values of 0.999 with Low Security, 0.971 with Medium Security and 0.999 with High Security so the overall discrimination is strong although the separability is less of the Medium Security class than with other models.

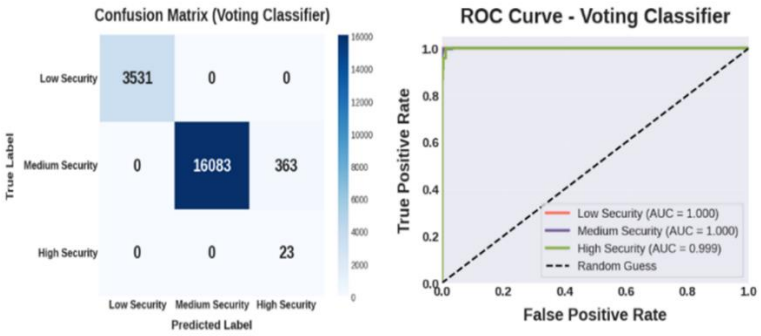

Figure 8 : Voting Classifier Confusion Matrix and ROC Curves.

Figure 8 illustrates the workings of the Voting Classifier that correctly classifies 3,531 Low Security, 16,083 Medium Security, and 23 High Security samples, only 363 medium Security samples are incorrectly identified as Medium Security and no other classification errors are observed. The ROC analysis shows that there is great discrimination with AUC values of 1.000, 1.000 and 0.999 in Low Security, Medium Security, and High Security respectively, which proves the strength and the reliability of the ensemble-based Voting Classifier.

Figure 9 : Contribution Analysis and Probability Distribution of Prediction of Security Classification.

Figure 9 demonstrates the prediction probabilities of the three security classes as predicted by the model with the scores of probabilities being 0.30, 0.43, and 0.27 with the ultimate decision being with the best probability score. The figure also shows the feature contribution analysis which shows the most influential factors contributing to the prediction, the Total Features Enabled (1.14), Compliance Score (1.33), Monitoring Score (1.32), Total Security Features (−0.34), and Management Score (0.26) features, policy- and access-control features.

A. Comparative Analysis

Table 2 evaluates the performance of different models in assessing IAM policy using the cloud setting. The proposed model achieves a high accuracy of 97.24, indicating reliable performance. Most other methods are outperformed by ensemble-based methods, including Voting (98.19%) and LightGBM (98.01), which have better predictive performance. The reported models of previous studies, such as the Isolation Forest, LSTM with Isolation Forest and GRN have lower accuracies and the proposed and ensemble models are effective.

Table 2 : Model Performance Comparison of Different Evaluation Metrics for IAM Policies in Cloud

Model	Reference	Accuracy
RFC	Proposed	97.24
LightGBM		98.01
GBC		96.03
XGBoost		90.95
Voting		98.19
Isolation Forest	[26]	89.9
LSTM + Isolation Forest	[27]	91.7
GRN	[28]	89.5

To implement the proposed framework, five ML models are used such as RFC, LightGBM, GBC, XGBoost and Voting ensemble to test IAM policies in cloud environments. RFC provides resilience to noise and strong support for the high-dimensional feature, whereas LightGBM can train quickly and support the use of large cloud data. GBC is better at predictive performance by means of sequential learning and error correction, whereas XGBoost is better at making a model more stable by regularization and efficient tree optimization. The Voting model is an amalgamation of the individual learner, which results in more dependable and stable IAM policy choices in extreme cases of cloud security.

B. IAM Policies and Cloud Computing

One of the very basic aspects of cloud computing security is the IAM policies that can control how the users, services, and applications can be granted access to cloud resources. IAM establishes authentication, authorization, and permission control systems that will make sure that only authorized players can carry out authorized operations. The dynamism of users, distributed resources, and multi-tenant environments have complicated the IAM policies with the fast adoption of cloud services. IAM policy management is thus critical towards reducing unauthorized access, imposing compliance and ensuring confidentiality, integrity and availability of cloud-based systems.

V. CONCLUSION AND FUTURE SCOPE

The transition of digital enterprises necessitates that IAM systems transition from static, policy-based access control to intelligent, proactive threat mitigation. The AI-driven anomaly detection framework introduced in this study improves IAM capabilities by continuously analyzing identity behavior, identifying anomalies, and enabling real-time responses to potential threats. The study proposed a well-designed and strong ML framework to optimize IAM policies in clouds. Through the use of extensive preprocessing, feature selection using Boruta and hybrid data balancing and ensemble learning, the given approach successfully addressed high-dimensional cloud access control data and class imbalance. The experimental findings showed that ensemble-based models, the Voting Classifier and the LightGBM, performed better according to accuracy, F1-score, cross-validation stability, and ROC-AUC, indicating that they are reliable in cloud security classification. The framework represents valuable security insights with the identification of powerful IAM attributes and facilitates informed and data-driven policy decisions in complex cloud environments.

Future research can focus on testing the framework on larger real-world cloud datasets and on dynamic or streaming IAM data. The use of deep learning models, explainable AI methods, and real-time mechanisms of policy adaptation would be even more effective in terms of scalability, interpretability, and practical implementation in changing cloud security contexts.

VI. REFERENCES

[1] A. Wairagade and S. Ranjan, “AI in Identity and Access Management (IAM) for Enterprise Systems: A Comparative Analysis,” Procedia Comput. Sci., vol. 263, pp. 167–174, 2025, doi: 10.1016/j.procs.2025.07.021.

[2] S. Amrale, “Proactive Resource Utilization Prediction for Scalable Cloud Systems with Machine Learning,” Int. J. Res. Anal. Rev., vol. 10, no. 4, pp. 758–764, 2023.

[3] S. K. Chintagunta, “Enhancing Cloud Database Security Through Intelligent Threat Detection and Risk Mitigation,” TIJER – Int. Res. J., vol. 9, no. 10, pp. 49–55, 2022.

[4] D. Patel, “Leveraging Blockchain and AI Framework for Enhancing Intrusion Prevention and Detection in Cybersecurity,” Tech. Int. J. Eng. Res., vol. 10, no. 6, 2023, doi: 10.56975/tijer.v10i6.158517.

[5] M. Menghnani, “Modern Full Stack Development Practices for Scalable and Maintainable Cloud-Native Applications,” Int. J. Innov. Sci. Res. Technol., vol. 10, no. 2, 2025, doi: 10.5281/zenodo.14959407.

[6] M. R. R. Deva, “Advancing Industry 4.0 with Cloud-Integrated Cyber-Physical Systems for Optimizing Remote Additive Manufacturing Landscape,” in 2025 IEEE North-East India International Energy Conversion Conference and Exhibition (NE-IECCE), IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/NE-IECCE64154.2025.11182940.

[7] A. Parupalli and H. Kali, “An In-Depth Review of Cost Optimization Tactics in Multi-Cloud Frameworks,” Int. J. Adv. Res. Sci. Commun. Technol., vol. 3, no. 5, pp. 1043–1052, Jun. 2023, doi: 10.48175/IJARSCT-11937Q.

[8] P. Chandrashekar and M. Kari, “Design Machine Learning-Based Zero-Trust Intrusion Identification Models for Securing Cloud Computing System,” Int. J. Res. Anal. Rev., vol. 11, no. 4, pp. 901–907, 2024.

[9] S. Phalke, Y. D. Athave, and B. N. Ilag, “A Multi-Layered Approach to IT Infrastructure Governance and Compliance: Security, Hardening, and Audit Readiness,” Int. J. Comput. Appl., vol. 187, no. 12, p. 9, 2025, doi: 10.5120/ijca2025925133.

[10] G. Maddali, “An Efficient Bio-Inspired Optimization Framework for Scalable Task Scheduling in Cloud Computing Environments,” Int. J. Curr. Eng. Technol., vol. 15, no. 3, pp. 229–238, 2025.

[11] V. Shah, “Traffic Intelligence in IoT and Cloud Networks: Tools for Monitoring, Security, and Optimization,” Int. J. Recent Technol. Sci. Manag., vol. 9, no. 5, 2024, doi: 10.10206/IJRTSM.2025894735.

[12] V. Shewale, “Beyond EDR: Exploring the rise of XDR for unified threat detection and response,” World J. Adv. Eng. Technol. Sci., vol. 15, no.

2, pp. 380–386, May 2025, doi: 10.30574/wjaets.2025.15.2.0551.
[13] N. K. Prajapati, "Cloud-based serverless architectures: Trends, challenges and opportunities for modern applications," World J. Adv. Eng. Technol. Sci., vol. 16, no. 1, pp. 427–435, Jul. 2025, doi: 10.30574/wjaets.2025.16.1.1225.
[14] A. Syed, AI-Powered Threat Detection and Mitigation. 2024.
[15] C. B. Bora, J. Silva Weber, and N. Zincir-Heywood, "Network Identity Management: Application, Action and Device Aware Monitoring," in 2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC), IEEE, Jul. 2025, pp. 2269–2274. doi: 10.1109/COMPSAC65507.2025.00319.
[16] A. AbouElabbas, A. H. Abdel-Gawad, and Y. Fahmy, "A Tabular Deep Learning Approach for Access Control Using TabNet," in 2025 7th Novel Intelligent and Leading Emerging Sciences Conference (NILES), IEEE, Oct. 2025, pp. 134–138. doi: 10.1109/NILES68063.2025.11231988.
[17] H. B. Demirsoy, E. N. Kose, F. Aydogan, M. H. Ezgin, and M. A. Akcayol, "Hybrid Deep Learning Model Based Advanced AI-Driven Identity and Access Management System for Enhanced Security and Efficiency," in 2024 8th International Symposium on Innovative Approaches in Smart Technologies (ISAS), IEEE, Dec. 2024, pp. 1–4. doi: 10.1109/ISAS64331.2024.10845215.
[18] H. Sivaraman, "Zero Trust Identity and Access Management (IAM) in Multi-Cloud Environments," ESP J. Eng. Technol. Adv., vol. 3, no. 2, pp. 135–139, 2023, doi: 10.56472/25832646/JETA-V3I6P108.
[19] T. Van Ede, N. Khasuntsev, B. Steen, and A. Continella, "Detecting Anomalous Misconfigurations in AWS Identity and Access Management Policies," in CCSW 2022 - Proceedings of the 2022 Cloud Computing Security Workshop, co-located with CCS 2022, 2022. doi: 10.1145/3560810.3564264.
[20] S. Chatterjee, "A Data Governance Framework for Big Data Pipelines: Integrating Privacy, Security, and Quality in Multitenant Cloud Environments," Tech. Int. J. Eng. Res., vol. 10, no. 5, 2023, doi: 10.56975/tijer.v10i5.158181.
[21] M. Mehmood, R. Amin, M. M. A. Muslam, J. Xie, and H. Aldabbas, "Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning," IEEE Access, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3273895.
[22] Balaji, R. (2024) 'Threat-Model-Driven Zero-Trust Authorization for Microservice APIs in Kubernetes-Based Cloud-Native Applications', ESP Journal of Engineering & Technology Advancements, 4(4), pp. 223–234. doi: 10.56472/25832646/JETA-V4I4P128.
[23] G. Sarraf and V. Pal, "Autonomous Threat Detection and Response in Cloud Security: A Comprehensive Survey of AI-Driven Strategies," Int. J. Emerg. Res. Eng. Technol., vol. 6, no. 4, 2025, doi: 10.63282/3050-922X.IJERET-V6I4P114.
[24] S. Srinivasan, R. Sundaram, K. Narukulla, S. Thangavel, and S. B. Venkata Naga, "Cloud-Native Microservices Architectures: Performance, Security, and Cost Optimization Strategies," Int. J. Emerg. Trends Comput. Sci. Inf. Technol., vol. 4, no. 1, pp. 16–24, 2023, doi: 10.63282/3050-9246.ijetcsit-v4i1p103.
[25] V. Verma, "Big Data and Cloud Databases Revolutionizing Business Intelligence," TIJER – Int. Res. J., vol. 9, no. 1, 2022.
[26] S. Narang and A. Gogineni, "Zero-Trust Security in Intrusion Detection Networks: An AI-Powered Threat Detection in Cloud Environment," Int. J. Sci. Res. Mod. Technol., vol. 4, no. 5, pp. 60–70, Jun. 2025, doi: 10.38124/ijsrmt.v4i5.542.
[27] I. Bansal, "Digital Transformation using Artificial Intelligence and Machine Learning for Secure Enterprises for Secure Enterprise Applications: A Framework using Cloud IAM Security," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 11s, pp. 815–821, 2023.
[28] K. C. Wannere, "AI-Augmented Threat Detection and Policy Drift Remediation in Hybrid Cloud Network Security Architectures," Int. J. Eng. Res. Technol., vol. 14, no. 05, 2025.
[29] B. Rajak, N. Kumaresh, N. K. Hamid, M. B. Alazzam, S. I. Hassan, and S. V, "AI-Driven Anomaly Detection for Secure Identity and Access Management in Cloud Platform," in 2025 Global Conference in Emerging Technology (GINOTECH), IEEE, May 2025, pp. 1–5. doi: 10.1109/GINOTECH63460.2025.11076807.