International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

A CLOUD-NATIVE APPROACH TO STATISTICAL ANOMALY DETECTION AND AUTOMATED DATA QUALITY VALIDATION

Chalapathi Koneni
Independent Researcher, USA.
Sanjay Lokula
Independent Researcher, USA.
Lalan Panjiyar
Independent Researcher, USA.

Abstract
The research proposes a statistical anomaly detection cloud architecture and automatic data quality validation with multi-tenant cloud resource utilization data. A model, based on utilization of XGBoost, is utilized to determine hidden overutilization of the resources with high accuracy and with appropriate support of precision-recall analysis and confusion-matrix analysis. Pre-model data quality checks are automated to ensure that there is an audit trail of reliability. SHAP-based explainability leads to better transparency and governance, which proves the efficiency of the framework at the scale of reliable cloud monitoring and data quality assurance.

Keywords:
Anomaly Detection, Cloud-Native Analytics, Multi-Tenant Cloud Monitoring, Explainable AI (SHAP), Data Quality Validation, Resource Usage Governance

Received: 10-11-2025 Accepted: 24-12-2025 Published: 31-12-2025

I. INTRODUCTION

Cloud-native frameworks are progressively dependent on troves of constantly streaming operational data in order to ensure reliability of the service, performance optimization, and cost efficiency. Dynamism and multiple tenancy characteristics of cloud environments make them vulnerable to the problems of data integrity and the misuse of resources that can go unnoticed through conventional monitoring strategies. Failure to have valid or accurate data can result in poor capacity planning and security weaknesses as well as service delivery deficiencies [1]. The research centers on the utilization of statistical anomaly detection and automated data quality validation as requisite processes of upholding credible data on cloud monitoring. The research analyzes normal resource consumption patterns and abnormal resource consumption patterns, as analyzed utilizing cloud resource usage. This contains CPU, disk I/O, memory, and network I/O usage statistics recorded through a shared multi-tenant environment. This dataset contains a variety of workload types and labelled anomalies, involving undetected consumptive scenarios involving illegal compute-intensive work. Outcomes are useful in enhancing real-time observability, governance, and reliability within contemporary cloud frameworks [2]. The research proves a cloud-native architecture to enhance the quality of data and anomaly awareness through the usage of statistical methods to recognize deviations from the expected behavior and incorporate automated rules of validation.

Aims and Objectives

Aim:

The purpose of this research is to test statistical anomaly detection and automated data quality validation against the multi-tenant cloud resource usage data to improve governance, reliability, and observability of the cloud-native settings.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

Objectives:

- To evaluate cloud resource metrics to recognize normal and anomalous usage trends within a multi-tenant environment.
- To implement statistical anomaly identification approaches for identifying concealed resource overuse through various workload types.
- To develop automated data quality validation rules for assuring consistency, accuracy, and reliability of the cloud monitoring data.
- To examine the usefulness of a cloud-native framework in enhancing anomaly identification and data quality assurance.

II. LITERATURE REVIEW

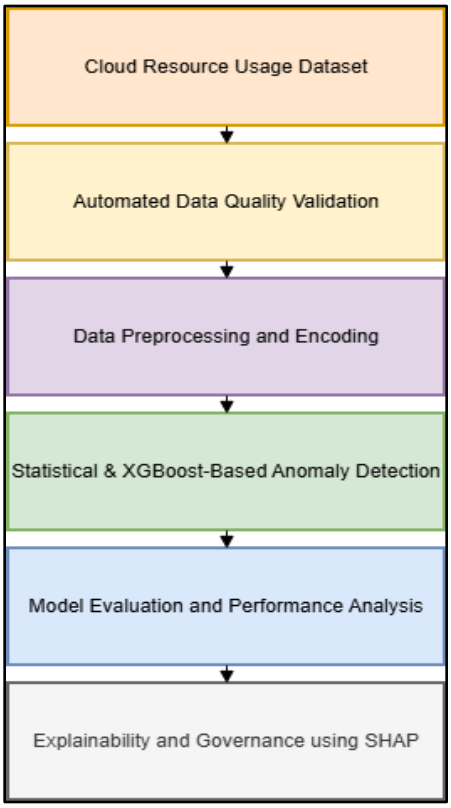


Fig 1: Research Flow

A. The Goal of the Review:

The review aims at discussing the means to utilize statistical anomaly detection as well as automated data quality validation methods for multi-tenant cloud resource usage data to enhance governance, reliability, and effectiveness of real-time monitoring.

B. Study of Previous Literature

1. Trends of Cloud Resource Utilization in Multi-Tenant Environments

Cloud computing systems are multi-tenant, signifying that several users and workloads utilize the basic infrastructure. This defines cloud resource utilization as having dynamic and heterogeneous trends depending on the nature of workloads, user behavior, and changes in demand over time. Established patterns are deviated; a sign of inefficiency, mal-configurations, or hidden overuse of resources can be detected [3]. Measures like memory, CPU use, disk I/O, and network I/O are predictable based on trends while operating in normal conditions and are indicative of routine activities on the applications, such as database queries, web services, and media streaming.

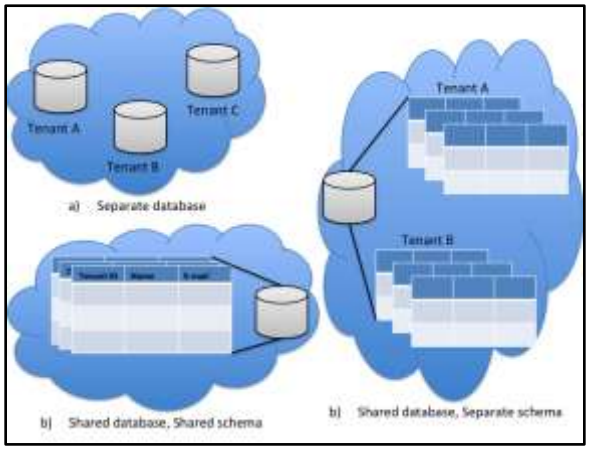


Fig 2: Multi Tenancy in cloud computing

Anomalies tend to be highly difficult to notice in a multi-tenant environment, as normal workloads that are a result of high demand can obscure the abnormal behavior. Patterns of utilization at an individual user and aggregated tenant level can be vital in adequate monitoring and governance [4]. The adequate resource metric analysis allows identifying the routine use patterns, cyclicality, as well as the workload-related signature. This understanding of the basis

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

can be utilized to assess the statistical anomaly detection approaches and automated data quality validation. This develops statistically strong baselines where abnormal resource consumption can be identified within the cloud-native monitoring rules.

2. Statistical Approaches for Detecting Anomalies in Cloud Metrics

The methods of statistical anomaly detection are vital in detecting abnormal behavior within the cloud resource monitoring data. CPU Usage, DiskI/O, MemoryUsage, and NetworkI/O are all continuously produced resource metrics as data within cloud-native systems. As a result, statistical approaches are highly appropriate in order to recognize an anomaly against an expected behavior. Normal operation limits have been set by approaches involving threshold-based evaluation, calculation of Z scores, moving averages, and the utilization of "interquartile range (IQR)" [5]. The values of resources are more than statistically established limits, and the possible anomalies, like hidden utilization of resources, can be determined.

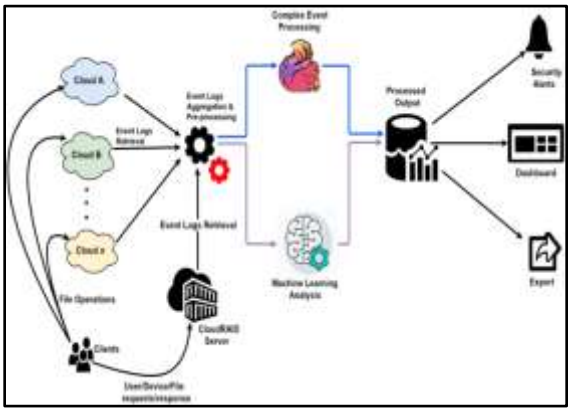


Fig 3: Anomaly Detection within the Cloud Storage

Strategies are specifically applicable where there is a workload diversity within a multi-tenant setting, where there is inherent variation in the utilization patterns. Statistical techniques are transparent and interpretable, as cloud operators know the reason that a data point is considered anomalous [6]. Moreover, they facilitate real-time detection, which has a low level of computational overhead, and can be utilized within a scalable cloud-native framework. Anomaly detection utilizing statistical methods is perceived to enhance the accuracy of anomaly detection as well as improve the overall quality of the data when combined with automated validation mechanisms.

3. Automated Data Quality Validation within Cloud Monitoring Systems

This is significant to do automated validation of the data quality to be sure that cloud monitoring systems are reliable as well as trustworthy. Through cloud-native infrastructure, resource metrics in great volumes constantly accumulate, and the manual verification of the data is not feasible [7]. The automated validation mechanisms are utilized to perform predefined rules and statistical validations to evaluate the accuracy, completeness, consistency, and validity of the data on a real-time basis. In the case of cloud resource measurements, involving memory utilization, CPU utilization, disk I/O, and network I/O, validation rules can be utilized to detect the absence of values. Values are the out-of-range, time anomalies, and the labeling of workload resources.

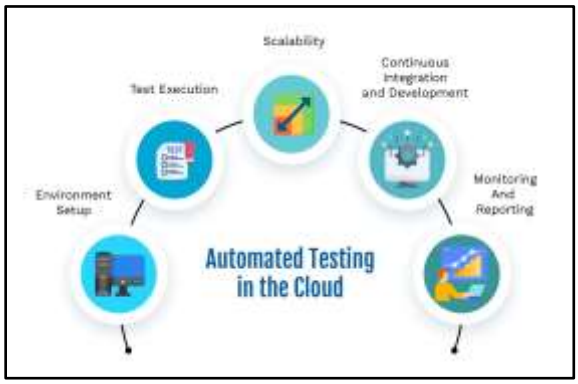


Fig 4: Automated Testing within the cloud environment

Significant in multi-tenant settings, where noise or inaccurate data can occur within flawed anomaly detection or the wrong interpretation that there is overutilization of resources.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

Ensuring early detection of data ingestion errors, as well as the early monitoring of faults before spreading into analytical models, is also aided by automated validation [8]. An enhanced and cost-effective statistical anomaly detection can be obtained by integrating data quality checks into pipelines of cloud-native monitoring. This ensures scalable, resilient, and governance-ready cloud data ecosystems.

4. Cloud-Native Frameworks for Scalable Anomaly Detection and Governance

Cloud-native architectures have the architectural support that is needed to obtain scalable anomaly detection and proper governance in contemporary cloud settings. Frameworks utilize distributed processing, containerization, and micro services to process high velocity monitoring data developed by multi-tenant cloud infrastructures [9]. The principles of cloud-native, the system of anomaly detection, and data quality validation can be implemented as a modular network of scalable services able to work in close real-time. This allows consistent monitoring of resource measures, involving memory utilization, CPU utilization, disk I/O, and network I/O without impacting overall system performance.

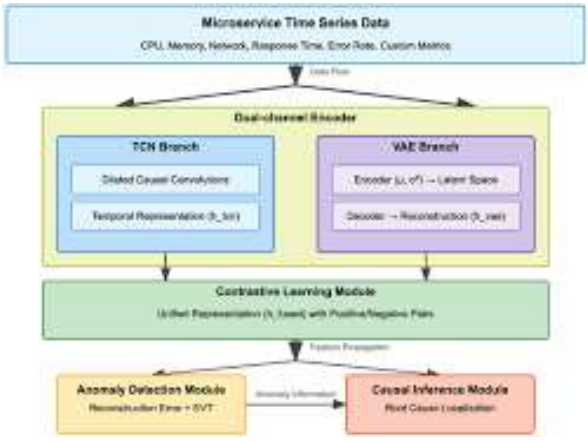


Fig 5: Multi-Dimensional Anomaly Detection

Cloud-native frameworks assist in enforcing policies automatically, provide auditability and observability, keys that are needed to govern and comply. Through a statistical anomaly detector incorporated with frameworks, it is possible to capture both abnormal resource utilization and hidden patterns of overuse stringently through both tenants and workloads [10]. Moreover, cloud-native designs enable easy integration into monitoring, alerting, and orchestration tools to make it resilient and tolerant to faults. The cloud-native frameworks are vital in ensuring the integrity, transparency of the operation, and confidence in the cloud monitoring systems.

Literature gap

Literature discusses in detail cloud monitoring and anomaly detection strategies, but little has combined statistical anomaly detection and automated data quality validation in a cloud-native, multi-tenant environment. The literature can be limited to a few studies covering concealed resource overuse using labelled time-series data, and a gap is observed in unified, scalable frameworks that assure data governance, integrity, and real-time awareness of anomalies.

III. METHODOLOGY

Dataset Description

Timestamp	CPU_Usage	Memory_Usage	Disk_IO	Network_IO	Workload_Type	User_ID	Anomaly_Label
01-07-2025 00:00	18.68	43.19	11.4	6.03	Database_Query	user_1	0
01-07-2025 00:01	25.31	45.43	7.68	17.67	Video_Streaming	user_1	0
01-07-2025 00:02	5.87	49.5	14.08	3.48	Database_Query	user_1	0
01-07-2025 00:03	28.92	25.88	17.33	4.77	Web_Service	user_1	0
01-07-2025 00:04	55.39	43.94	10.61	4.48	Web_Service	user_1	0
01-07-2025 00:05	23.99	17.08	6.99	6.85	Web_Service	user_1	0
01-07-2025 00:06	24.66	39.94	8.85	5.35	Web_Service	user_1	0
01-07-2025 00:07	32.09	30.4	3.36	20.58	Video_Streaming	user_1	0

Fig 6: Dataset Description

The data set contains cloud resource records gathered within a multi-tenant setting, which are CPU utilization, memory utilization, disk I/O, and network I/O. The records are time-stamped and identified by a type of workload and user ID. The anomaly labels reveal the statistical anomaly detection as well as automated data quality validation in cloud-native systems through signaling normal behaviors and hidden resource over-utilization.

Research Methods

The research takes a quantitative as well as data-intensive research perspective in order to investigate statistical anomaly identification and

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

automated data quality validation utilizing a cloud-native setting. The dataset utilized to run the analysis is a cloud resource usage data comprised of CPU usage measurements, memory consumption measurements, disk I/O measurements, and network I/O measurements obtained on a multi-tenant infrastructure. The first step taken is automated data quality checks that can assist in evaluating missing data values and duplicate entries, as well as maintain the reliability of data before modelling.

Categorical variables involving the type of workload and user identifiers are coded, and machine learning analysis can be performed. Stratified sampling is then applied to the dataset to assess anomaly distribution, where training and test sessions are determined [11]. A highly trained XGBoost multicast model is utilized to identify abnormal resource utilization patterns in the system, where the weighted learning is applied in cases of imbalance within the classes. Detection effectiveness is determined by the utilization of precision, recall, F1-score, confusion matrix, as well as precision-recall curves to determine the overall performance of the model. SHAP is utilized to combine explainable AI techniques to comprehend the contribution of features, along with improving transparency [12]. The approach corresponds to a cloud-native analytical pipeline to be used in identifying anomalies on a large scale, controlling them, and ensuring automatic data quality assurance.

IV. DATA ANALYSIS

Data Preprocessing

```
data_cloud.isnull().sum()
```

	0
Timestamp	0
CPU_Usage	0
Memory_Usage	0
Disk_IO	0
Network_IO	0
Workload_Type	0
User_ID	0
Anomaly_Label	0

dtype: int64

Fig 7: Checking the null values

The figure shows that there are no missing values within the dataset of cloud resource usage values among all variables, involving the type of workload, resource metrics, and the label of anomalies. High data completeness is validated and allows automated data quality validation, which is essential in assuring strong statistical anomaly identification along with a correct model [13]. It works well within a cloud-native monitoring framework.

```
data_cloud.duplicated().sum()
```

```
np.int64(0)
```

Fig 8: Checking the duplicates

This can be seen in this particular research that there are no duplicate records within the cloud resource usage database. Data consistency and integrity are maintained, which are necessary in automated data quality validation [14]. The result of statistical anomaly identification lacks the risk of bias through repeated observations within a cloud-native monitoring environment.

```
df = data_cloud.copy()
df['Timestamp'] = pn.to_datetime(df['Timestamp'])
df = pn.get_dummies(df, columns=['Workload_Type', 'User_ID'], drop_first=True)
```

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

```
X = df.drop(columns=['Timestamp', 'Anomaly_Label'])
y = df['Anomaly_Label']
```

The mentioned figure shows the preprocessing stage of data processing, which consists of transforming timestamps as a form of temporal consistency, encoding categorical workload and user characteristics, and extracting features and labels of anomalies. The steps allow validating data quality automatically and preparing cloud resource metrics to identify statistical anomalies in a cloud-native analyst architecture.

Anomaly Detection Model Using XGBoost Classifier Model

```
model = XGBClassifier(
    n_estimators=200,
    max_depth=5,
    learning_rate=0.05,
    subsample=0.8,
    colsample_bytree=0.8,
    scale_pos_weight=(y_train.value_counts()[0]
/ y_train.value_counts()[1]),
    eval_metric='logloss',
    random_state=42
)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
y_proba = model.predict_proba(X_test)[:, 1]
```

The XGBoost classifier is utilized here to recognize the anomaly within a cloud-native framework. This model is set to process class imbalance with the assistance of weighted learning and is optimized by depth and learning rate pressure parameters. Training multi-tenant cloud resource measurements and producing class and anomaly probability allows a model [15]. This appropriately recognizes hidden resource overuse as well as promotes statistical anomaly identification, along with automated data quality assurance purposes.

Explainability Using SHAP

```
explainer = shap.TreeExplainer(model)
shap_values = explainer.shap_values(X_test)
shap.summary_plot(shap_values, X_test, plot_type="bar")
shap.summary_plot(shap_values, X_test)
```

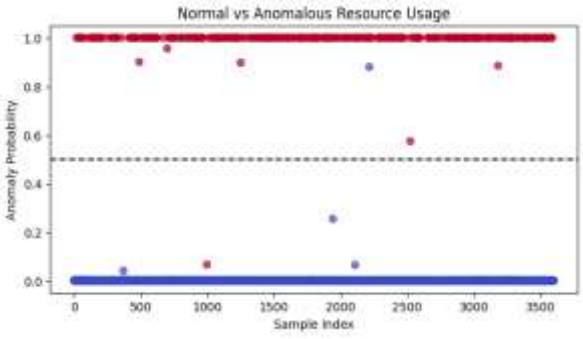
SHAP is applied to explainable anomaly detection in a data quality system on a cloud-native platform. SHAP can measure the influence of each cloud resource measure and a workload characteristic on providing anomalies to ensure identification of significant drivers involving the consumption of CPU and high-risk workloads [16]. This explainability is also significantly enhancing the principles of governance, auditability, and trust since automated decisions in anomaly detection are explainable and consistent with the prerequisites of data quality validation in multi-tenant cloud contexts.

Evaluation and Model Performance

```
print("Classification Report:\n")
print(classification_report(y_test, y_pred, digits=4))
```

The classification report is applicable to discriminate the model of anomaly identification. This assesses precision, recalls, and the F1-score, which is significant in measuring the detection accuracy of unbalanced cloud data collections. These measures are significant to prove the correctness of statistical anomaly detection and to ensure the automated data validation of data quality within a cloud-native monitoring and governance framework.

V. RESULTS AND DISCUSSION



International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

Fig 9: Anomaly Detection Visualization

The figure assesses the visualization of the probability of anomalies projected by the XGBoost model in the cloud monitoring records. Normal instances cluster around zero probability, whereas anomalous instances have high probability above the decision threshold. The sharp delineation portrays that there is high model confidence in differentiating between the hidden overuse of the resource and normal behaviors. Minimal bordering cases assess realistic workload variation within the multi-tenant cloud environment [17]. This assists in evaluating statistical anomaly identification as a real-time monitoring and governance tool of the cloud.

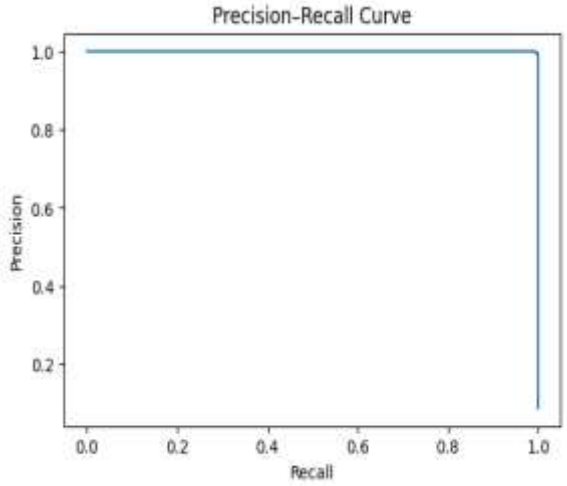


Fig 10: Displaying the Precision–Recall Curve

The accuracy of the precision is high during most recall values, which shows that the detection of abnormal cloud resource utilization is reliable. The extreme recall values decrease drastically, and this is the trade-off of trying to capture all the anomalies. This action is anticipated in unbalanced data and proves that the model fits best to anomaly detection, where precision and recall values are more informative than accuracy is in cloud-native data quality assurance settings.

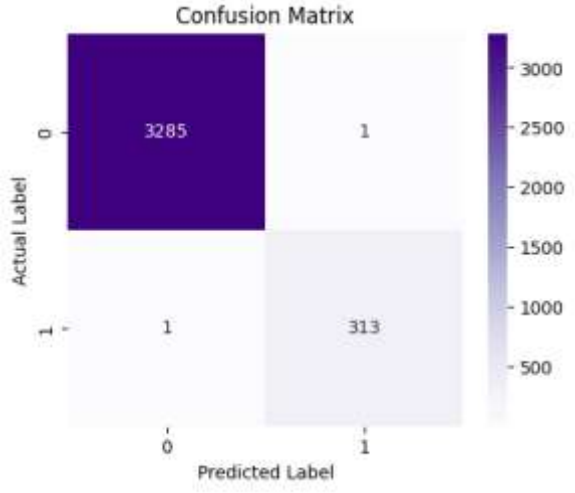


Fig 11: Confusion Matrix

The confusion matrix features about 3285 normal cases along with 313 anomalous cases, which have been rightfully classified with one false positive and one false negative. This very low false separation rate shows that there is adequate division between the normal and unknown cloud resource usage. Balanced detection behavior reveals the performance of the model being effective in detecting the hidden resource overutilization, as well as high reliability [18]. This is crucial to automated overall data quality validation along with real-time cloud monitoring.

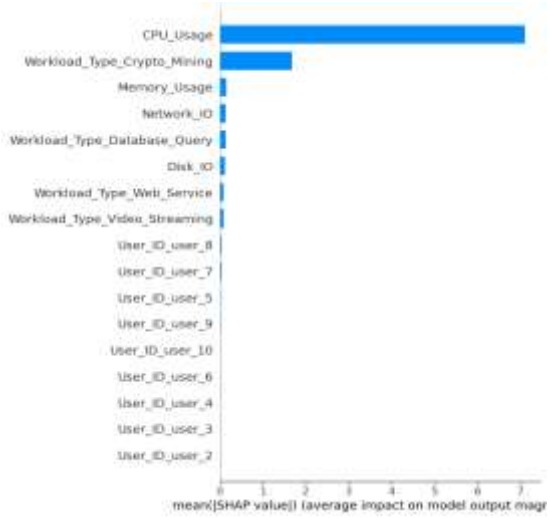


Fig 12: SHAP Feature Importance

This demonstrates the most adequate attributes that support the decision towards anomaly

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

detection. The leading indicators are CPU usage, crypto-mining workload, and memory usage. This is consistent with the typical cloud abuse trends, which feature heavy workloads of compute resources leading to uncharacterized usage. Ranking assesses that a model is learning indicative, real-world-relevant signals instead of the user identifiers [19]. This facilitates readable log controllers, considering other regimes and anomaly identification.

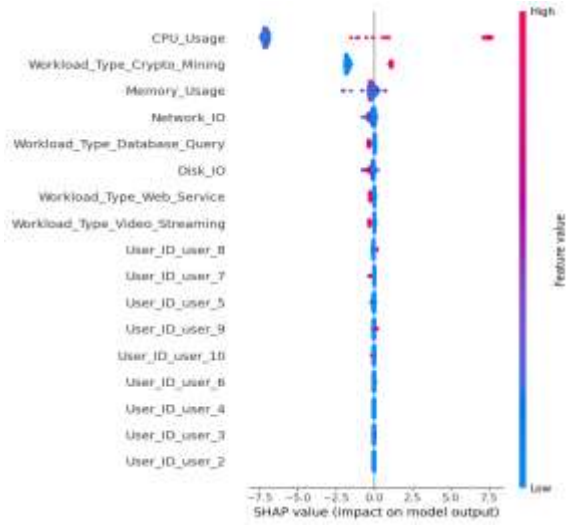


Fig 13: SHAP Details

The mentioned SHAP summary plot offers an adequate overview of feature values, assisting in predicting anomalies occurring in every observation. The high CPU load and crypto-mining workloads, in a continuous nature, skew the predictions to anomaly disclosures, whereas the low values yield a normal category. Heterogeneity in interactions between features, as well as variability of SHAP values, influences the overall behavior of tenants and workloads [20]. This explainability makes the model unwarranted, and is a facilitator of automated data quality verification through giving feature-level interpretable outcomes.

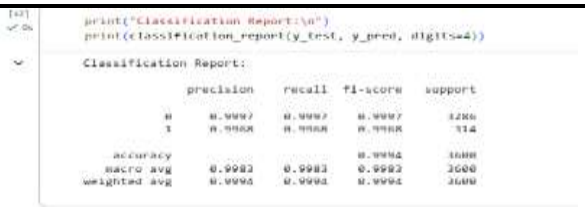


Fig 14: Summary of the classification report

Recall, precision, and F1-values of the two classes as reported by the classification report are greater than 0.996, and the accuracy is 0.9994. The recall to anomalies is high, which proves the skills of the model to identify the misuse of resources, and the percentage of false alarms is low. These findings confirm the strength of the anomaly detection model and its adaptability to scalable, adequate data quality assurance and faithful data management tools, which can be customized and deployed on cloud solutions.

Discussion:

The outcomes illustrate that the suggested cloud-native anomaly identification framework can be efficient in identifying hidden cases of resource overuse and ensure the data quality and dependability. The overall accuracy of the XGBoost model is 99.94%, and the recall, precision, and F1-value are higher than 0.996 in the normal and anomalous classes, signifying that it demonstrated strong results in performance despite the class imbalance. This reliability is further supported by the confusion matrix, which indicates that there is a single false positive and a single false negative out of 3600 records, in confirmation of this reliability.

The precision-recall plot demonstrates a high ability to separate the classes and assesses that anomalies are detected correctly in a multi-tenant setup. SHAP explainability indicates that the computation of CPU usage and crypto-mining workloads are the most prevalent in anomaly prediction, which assures the relevance of the domain and transparency of models [21]. Results validate previous evidence that automated data quality validation characteristics

implement statistical anomaly detection, versatile in promoting trust and governance, as well as real-time observability in cloud-native monitoring frameworks.

Research Limitations:

The research is based on one data set on labelled cloud resource utilization, which cannot be generalized to different cloud systems and real-life working conditions. Presence of well-defined anomalies that can be widely separated can make detection less challenging as compared to much more complicated, unmarked as an anomaly, or dynamic situations of resource abuse [22]. Additionally, an analysis is not performed regarding real-time deployment limitations involving streaming latency, peak load scalability, and compatibility with cloud infrastructures used in production. Statistical and supervised machine learning require quality labelled data, which is not present in real-time cloud monitoring frameworks [23]. This can affect the viability of a real-world implementation.

VI. CONCLUSION AND FUTURE RESEARCH

The research shows that statistical anomaly identification can be utilized in combination with automated data quality validation to offer a powerful cloud-native tool to recognize hidden overuse of resources. The presented XGBoost-animated framework was diagnosed with an extremely dependable performance in detection and can be easily approached with the assistance of SHAP explainability. The accuracy, explainability, and governance-readiness of data assist in establishing the usefulness of scalable, explainable analytics in maximizing trust, observability, and reliability in operations with current multi-tenant cloud environments.

Future Research:

The future research can build on the study by assessing unsupervised and semi-supervised anomaly detection methods to deal with situations in which there are only a few or changing labels. Innovative possibilities include real-time streaming based on cloud-native technologies like serverless applications as well as Kubernetes [24]. Furthermore, adding adaptive thresholds, concept drift detection, and cross-cloud data can lead to better scalability, resilience, and generalization of automated data quality and anomaly identification frameworks. Federated learning strategies can be evaluated for preserving data privacy through tenants while offering collaborative anomaly identification [25]. Integration of cost-aware optimization can additionally assess anomaly identification outcomes.

VII. REFERENCE

[1] SAMUEL, A., 2021. Cloud-Native AI solutions for predictive maintenance in the energy sector: A security perspective. *Available at SSRN 5290068*.

[2] Keane, A.F.G., 2021. AI-Enhanced Cloud Security with ERP Integration: MFA, Multivariate Classification, and Transformer-Based Threat Detection. *International Journal of Computer Technology and Electronics Communication*, *4*(6), pp.4288-4295.

[3] Veluru, S.P., 2021. Leveraging AI and ML for Automated Incident Resolution in Cloud Infrastructure. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, *2*(2), pp.51-61.

[4] Hagemann, T. and Katsarou, K., 2020, December. A systematic review on anomaly detection for cloud computing environments. In *Proceedings of the 2020 3rd Artificial Intelligence and Cloud Computing Conference* (pp. 83-96).

[5] Elsayed, M.A. and Zulkernine, M., 2020. PredictDeep: security analytics as a service for anomaly detection and prediction. *IEEE Access*, *8*, pp.45184-45197.

[6] Feng, L., Xu, S., Zhang, L., Wu, J., Zhang, J., Chu, C., Wang, Z. and Shi, H., 2020.

Anomaly detection for electricity consumption in cloud computing: framework, methods, applications, and challenges. *EURASIP Journal on Wireless Communications and Networking*, *2020*(1), p.194.

[7] "Scalable Suspicious Activity Detection Using Teradata Parallel Analytics And Tableau Visual Exploration," Inaternational Journal of Communication Networks and Information Security, vol. 8, no. 3, Jun. 2025, doi: 10.48047/ijcnis.8.3.236..

[8] Malaiya, R.K., Kwon, D., Suh, S.C., Kim, H., Kim, I. and Kim, J., 2019. An empirical evaluation of deep learning for network anomaly detection. *IEEE Access*, *7*, pp.140806-140817.

[9] N. Naga Charan, "Predictive Sql Injection Detection And Prevention Using Machine Learning Across Aws, Azure, And Google Cloud Platforms," International Journal of Engineering Science and Advanced Technology, vol. 22, no. 8, pp. 68–74, Aug. 2022, doi: 10.64771/ijesat.2022.v22.i08.pp68-74.

[10] Enokkaren, S.J., Bitkuri, V., Kendyala, R., Kurma, J., Mamidala, J.V. and Attipalli, A., 2021. Enhancing Cloud Infrastructure Security Through AI-Powered Big Data Anomaly Detection. *International Journal of Emerging Research in Engineering and Technology*, *2*(2), pp.43-54.

[11] Liu, Y., Pang, Z., Karlsson, M. and Gong, S., 2020. Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. *Building and Environment*, *183*, p.107212.

[12] S. Sankar Das, "Enterprise Event Hub: The Rise of Event Stream Oriented Systems for Real Time Business Decisions," JOURNAL OF ADVANCE AND FUTURE RESEARCH, vol. 1, no. 10, Dec. 2023, doi: 10.56975/jaafr.v1i10.500878.

[13] El-Shamy, A.M., El-Fishawy, N.A., Attiya, G. and Mohamed, M.A., 2021. Anomaly detection and bottleneck identification of the distributed application in cloud data center using software–defined networking. *Egyptian informatics journal*, *22*(4), pp.417-432.

[14] Oktay, T., Yoğurtçuoğlu, E., Sarıkaya, R.N., Karaca, A.R., Kömürcü, M.F. and Sayar, A., 2021. Multimodel anomaly detection on spatio-temporal logistic datastream with open anomaly detection architecture. *Expert Systems with Applications*, *186*, p.115755.

[15] Lou, P., Yang, Y. and Yan, J., 2019, June. An anomaly detection method for cloud service platform. In *Proceedings of the 2019 4th International Conference on Machine Learning Technologies* (pp. 70-75).

[16] S. R. Nelluri and K. M. Poluri, "Real-Time Marketing Optimization through Scalable Telemetry Data Engineering: A Framework for Enhanced Engagement and ROI," International Journal of Computer Trends and Technology, vol. 73, no. 1, pp. 26–31, Jan. 2025, doi: 10.14445/22312803/ijctt-v73i1p103.

[17] Guezzaz, A., Asimi, Y., Azrour, M. and Asimi, A., 2021. Mathematical validation of proposed machine learning classifier for heterogeneous traffic and anomaly detection. *Big Data Mining and Analytics*, *4*(1), pp.18-24.

[18] Nassif, A.B., Talib, M.A., Nasir, Q. and Dakalbab, F.M., 2021. Machine learning for anomaly detection: A systematic review. *Ieee Access*, *9*, pp.78658-78700.

[19] S. K. Immadi, "Optimizing ERP for Human Capital Management," Applied Research for Growth, Innovation and Sustainable Impact, pp. 377–384, Aug. 2025, doi: 10.1201/9781003684657-63.

[20] Ahmed, A., Hameed, S., Rafi, M. and Mirza, Q.K.A., 2020. An intelligent and time-efficient DDoS identification framework for real-time enterprise networks: SAD-F: Spark based anomaly detection framework. *IEEE Access*, *8*, pp.219483-219502.

[21] Prodduturi, S.M.K. (2025). Opportunities and Challenges for iOS Developers in Exploring

the Integration of Augmented Reality Technologies. International Journal of Engineering Science and Advanced Technology (IJESAT), 25(4), pp.200–207. ISSN 2250-3676.

[22] Yahyaoui, A., Abdellatif, T., Yangui, S. and Attia, R., 2021. READ-IoT: Reliable event and anomaly detection framework for the Internet of Things. *IEEE Access*, *9*, pp.24168-24186.

[23] Mason, A., Zhao, Y., He, H., Gompelman, R. and Mandava, S., 2019, June. Online anomaly detection of time series at scale. In *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)* (pp. 1-8). IEEE.

[24] Abreu, F.H., Soares, A., Paulovich, F.V. and Matwin, S., 2021. A trajectory scoring tool for local anomaly detection in maritime traffic using visual analytics. *ISPRS International Journal of Geo-Information*, *10*(6), p.412.

[25] Panicucci, S., Nikolakis, N., Cerquitelli, T., Ventura, F., Proto, S., Macii, E., Makris, S., Bowden, D., Becker, P., O'Mahony, N. and Morabito, L., 2020. A cloud-to-edge approach to support predictive analytics in robotics industry. *Electronics*, *9*(3), p.492.