iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.83038

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

AI-Driven Detection of Unauthorized Database Queries in Government Citizen Information Systems

Mohamed Saad, Dr. Adnan Al-Helali
Irbid National University, Jordan

Abstract: ***Therapiddigitaltransformationofgovernmentservicesmeansthatsensitivecitizen data increasingly resides in centralized database systems. These platforms are now thefocusofattacksfromexternalthreatsandinsiders,whichmayincludeprivileged usersrunningunauthorizedqueries.Traditionalrule-basedACmechanismsfailedto catch such advanced, context-aware unauthorized access behaviours. This paper describes anAI-basedframeworkfordetectingin realtimetheunauthorizedaccess to a database of a government citizen information system. Our approach combines machine learning (ML)models, suchas anomaly detection, NLP(on SQLqueries) and behavior analytics, to detect anomalous queries with respect to the established query patterns. The modelis built to be adaptive, and trained on query logs, user profile, and access context to separate benign queries from malicious or unauthorized ones. Results on a simulated government database environment show that our approach can reach high detection rates with low false-positive rates and near real-time performance, which are essential features to be applied in the public critical infrastructures. These results highlight the significance of AI-based monitoring as an additional ARM-layer for citizen privacy and data sovereignty protection in future cybersecurity.***

KEYWORDS: ***DetectionofUnauthorizedQuery;DetectionofIntrusionBasedonAI;Securityof Database;Detection of Anomaly; Learning Machine;Analytics of Behavior; Systems to Process Government Information;Privacy of Data for Citizens; Detection of InsiderThreat;Analysis of SQL Query***

I. INTRODUCTION

Governments worldwide experienced a dramatic shift to e-governance, including public records,socialwelfare databases, taxdatabases, healthrecords, andidentity management systems. These transformations havedelivered huge gains in service delivery;however,theyhavealsoexposedsensitivecitizendatatoincreasinglydire risks. Now, government databases have become lucrative targets for a variety of adversaries, from nation states andcybercrime groups to malicious insiders.

Traditional security controls — including role-based access control (RBAC), mandatory access control (MAC) and static firewall policies —offer a basic layer of protection. However, these techniques are fundamentally reactive and are ineffective against sophisticated attacks that utilize compromised credentials or misappropriate access. Specifically, insider threat has always been one of the most difficult to detect security problems as insiders are able to submit queries “that are syntactically correct(validquery)butsemanticallynotauthorized”andsuchqueries may “reveal sensitive information in a way that circumvents traditional access control protections”.

The advent of AI and ML led to new possibilities in the field of cybersecurity, allowing systems to be utilized to learn from past attacks and anticipate on coming ones. In databasesecurity,thesetechniquescan beusedtobuild profiles of normal querybehaviorandidentifystatisticalanomaliesthatcouldrepresentunauthorized access to the data on the fly. They are particularly well adapted to the dynamic, complex query patterns that characterize large-scale government information systems,where millions of queries are submitted on a daily basis by thousands of simultaneous users from a variety of departments.

There has been increasing attention to AI-enabled database monitoring, but the particular stage of government citizen information systems raises newchallenges. These include security and privacy (monitoring should not amount to unauthorized surveillance itself), heterogeneous legacy database systems, sensitivity and legal protection of citizen data under data protection laws, and high availability and minimum disruption to operations. Commercial and academic products and solutions tend tonot satisfy these domain-specific requirements altogether.

Thisarticlehelpstofillthosegaps,byintroducinganend-to-endAI-baseddetection system, specific to the case of government citizens' information systems.

©IJRASET: All Rights are Reserved

The framework incorporates several synergisticmethods:a1)an unsupervised anomaly detectionmethod tolearnbehaviorprofilesforusersandroles,a2)anaturallanguage processing (NLP) method to parsed SQL queries via syntactical and semantical analyses, a3) a graph-based method to model access behaviors for users over database objects, and a4) a multi-level anomly detection system is employed to detectsuspiciousactivitieswhichcanbeescalatedtohumanreviewerforareal-time alertengine.Itallowsnon-intrusiveintegrationwithexistingdatabasemanagement system (DBMS) and government IT infrastructures.

The rest ofthe paper is organized as follows. Section2 considers related work on database intrusion detection, anomaly-based query monitoring, and AI in cybersecurity. Threatmodel and Problem Formulation In this section, we present thethreatmodel,followedbyproblemformulation.InSection4,theAI-based detection design is presented in detail. Section5 presents the methodology of the experiment and the results of the evaluation. The challengeof implementation and the implications for ethics in the context of government are raised in Section 6. Finally, Section7 concludes the paper and suggests future work.

A. Background

Citizen information systems in the government, such as GCIS, are considered to be among the most sensitive types of public-sector digital infrastructure. These systems contain, process and manage extremely sensitive data like national identification records, tax returns, civil registration information, social security, medical insurance entitlements, and judicial records. With global governments racing to implement digital transformation strategies, the consolidation and interlinking of such databases have produced staggering operational efficiencies but also dramatic security risks. Access to citizen databases without authorization—by hostile foreign powers, suspicious insiders, or hijacked privileged accounts—is a dire threat to the nation's security and civil liberties.

Database query analysis is the primary means by which the enforcement of access control and data governance policies is realized. In government environments, access to records of citizens is tightly controlled, and only users with roles such as civil servants, police officers, and administrative staff have permission to run structured queries upon them based on their operational needs. However, authorized credentials do not guarantee legitimate intent. Insider threats — people who improperly use their legitimate access to conduct unauthorized activities such as spying, identity theft, bribery, or harm to the system — are among the most challenging security attacks to detect and recover from. The U.S. Cybersecurity and Infrastructure Security Agency (CISA) and analogous agencies across numerous other countries have consistently ranked insider threats as one of the greatest threats to critical government information assets.

Traditional database security mechanisms such as RBAC (role-based access control), DAC (discretionary access control) and MAC (mandatory access control) define boundaries on who can access what data and under which circumstances. But these controls work on coarse granularity and fundamentally cannot distinguish a properly authorized query from a query issued with stolen or misused credentials. An access control system for a database can't tell by itself whether a clerk who is looking at a citizen's tax record is really an honest worker conducting official business or a crook poring over potential victims. This disconnect between technical permissioning and what is considered the behavioral legitimacy in context is what drives AI-driven behavioral analytics frameworks to identify suspicious queries for unauthorized access.

Artificial Intelligence (AI) and Machine Learning (ML) have a potential transformational impact on database security by facilitating automatic generation of baselines of normal query behaviors and identification of statistical outliers that may indicate violations of policy. Methods such as anomaly detection, sequence modeling, NLP of SQL queries, graph-based behavioral analysis, and deep learning models can leverage historical query logs to build detailed user behavioral models and their access patterns. When incorporated in government citizen information systems, such detection frameworks based on AI technologies can achieve real-time, continuous, high- sensitivity monitoring of database activity, unveiling subtle unauthorized queries that would be hidden under the cover of rule-based security mechanisms.

In this work we present an AI-based detection system for government citizen information systems to detect unauthorized database queries. The architecture of our proposed system combines anomaly detection techniques, sequence-based behavior modeling, and informed SQL query semantic analysis in a scalable, privacy-compliant monitoring framework under the operational and regulatory constraints commonly found in public-sector database systems. By filling the crucial void between access authorization and behavior legitimacy, this contribution strives to enable government agencies with a realistic and efficient solution to guard citizen data against unauthorized usage.

II. LITERATURE REVIEW

A. Security Threats in Government Database Systems

A: Content There has been a substantial body of work in the area of securing the information systems of the government because of the value of the public-sector data to attackers and the impact if they are compromised.

Bertino and Sandhu (2005) laid the foundation in the area of database security by categorizing threats to databases into external attacks, insider misuse and inference attacks and they also reviewed access control models that can be used to protect large scale enterprise databases. Their categorization continues to be applicable to government, where multi-tier data architectures and tens of thousands of authorized users create similar attack surfaces.

Public-sector databases, in particular, have been highlighted for insider threat potentials. Liu et al. (2018) studied incidents of large-scale data breach in governmental agencies and showed that a large portion of the incidents was due to the actions of authorized users who had access to the breached data, but exceeded their legitimate access scope either maliciously or non-maliciously (i.e. The research highlighted that technical access controls are not enough and must also be supplemented by monitoring behaviours. In like manner, Homoliak et al. (2019) presented a systematic review on methods for detecting insider threats within government and enterprise organizations, with a focus on insider behaviors by grouping insiders into malicious insiders, negligent users, and compromised accounts and analyzing detection solutions across these types.

Al-Khouri (2014) researched the susceptibility of citizen data systems within e-government systems and concluded that weak authentication, over privilege, inadequate audit logging, and lack of real-time monitoring are the main systemic vulnerabilities in national identity and civil registration systems. Such conclusions are supported by exemplary government data breach incidents including the 2015 U.S. Office of Personnel Management (OPM) breach revealing background investigation records of more than 21 million federal employees and contractors, illustrating the monumental impact of poor database security in government infrastructures.

B. Database Activity Monitoring and Anomaly Detection

DAM has history of continuously evolving, as the monitoring of database transactions in real- time is unique comparing to other types of OS activities. Early DAMs were based on signature- based detection methods or rule engines, which would alert on queries that matched a known malicious pattern. Although well suited to detecting known attack signatures, these system have blind spots when detecting new or subtle unauthorized access patterns, especially if conducted by insiders who know and work to evade these known detection rules.

Statistical methods for detecting anomalies constitute a major improvement relative to signature- based approaches. Kamra et al. An early ML-based framework for detecting anomalous database access was proposed by (2008), which uses role-based access profiles derived from historical query logs employing statistical distance measures to detect deviations during runtime.

They demonstrated with their experiments on university and enterprise databases that ML-based profiling can recognize anomalous queries with low false positive rates. Extending this, Sallam et al. (2015) introduced a temporal aware anomaly detection model which also takes into account user role and query semantic features to enhance the detection accuracy of insider threat in database systems.

Recently, work on database access anomaly detection has emerged in graph-based and network analysis. Mathew et al. (2010) introduced a directed graph of database objects interacted by normal user queries and used this model to detect unauthorized access as anomalies in the underlying graph structure. The technique was especially suitable to detect data exfiltration attempts involving methodical traversal of related tables — as occurred in government database breaches where adversaries attempt to aggregate citizen records found across multiple linked tables.

C. Machine Learning and Deep Learning for SQL Query Analysis

Machine learning-based semantic analysis of SQL queries has become a promising technique to represent user intent and identify adversary database access. Queries in SQL have rich semantics about accessed relations, conditions, and the scope of data retrieval, which can be leveraged by ML models to identify malicious from benign queries. In (Moustafa et al., 2019), the authors proposed the use of natural language processing (NLP) method on tokenization and embedding on SQL query, for representing queries as high-dimensional feature vectors and to train classifier models to identify SQL injection attacks and unauthorized data access.

Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks have been utilized to capture sequential query patterns by modeling temporal correlations within user access sessions to identify anomalous behaviors. Tuor et al. (2017) found that LSTM-based sequence models learned from system log data can achieve high accuracy in insider threat detection by learning a normal temporal pattern of user behavior and alerting on sessions that are anomalous to learned patterns. When applied to sequences of database queries, LSTM may discover idiosyncratic recurrent behavioral patterns including query workflows, time distributions, volume of data, which are typical for user profiles.

Transformer-based architectures, which have obtained state-of-the-art performance in various natural language processing (NLP) tasks, have been recently adopted in cybersecurity anomaly detection. Chen et al. (2022) employed BERT-based models on log analysis for intrusion detection and showed that pre-trained transformer models can be fine-tuned with security log data and obtain better detection results than the conventional ML baselines. To the best of our knowledge, the use of transformer models on SQL query analysis for government database agents is an exciting, novel avenue we seek to pioneer in this work.

D. Privileged User Monitoring and Insider Threat Detection

Privileged user activity monitoring is an essential element in securing government databases since database administrators (DBAs), system architects, and high-ranking officials with advanced database access are at greatest risk for illicit data access due to their extensive access rights and expertise. Greitzer et al. (2012) introduced a psychosocial risk model for insider threat detection that combined technical behavioral indicators such as abnormal query amounts, off- hours access, and anomalous data export patterns with organizational behavioral signals to compute composite risk scores for privileged users. This combined approach exhibited better detection results than technicalalone approaches.

OVERVIEW OF SOLUTION The User and Entity Behavior Analytics (UEBA) solution is the best current industry practice for detecting insider threat and unauthorized access. UEBA systems build behavioral baselines for individual users and entities (for example service accounts, applications) and use statistical and machine learning-based anomaly detection to detect anomalies from these baselines of normal behavior. Gavai et al. (2015) studied UEBA-like techniques for identifying malicious insiders of enterprise networks and concluded that unsupervised anomaly detection techniques — especially isolation forests and one-class SVMs

— can effectively detect rare attacks among a large number of legitimate users. Their results can be applied directly to government monitoring of citizen databases, where the fraction of malicious to normal queries is likely to be very small.

The problem of identifying unauthorised queries that look valid in isolation but — when viewed as a whole — indicate malicious intent has driven work at the session (and campaign) levels of behavioural analysis. Eberle and Holder (2007) introduced a graph-based technique for detectinganomalous behavioral sequences in database logs, which can uncover multi-step unauthorized access campaigns, which would be undetectable by single-query detection schemes. This line of work demonstrated the need for considering query context and temporal sequences rather than treating each query as an atomic unit.

E. Privacy-Preserving Monitoring in Government Contexts

A unique challenge in using AI-enabled monitoring for government citizen information systems is the conflicting priorities between security monitoring and the privacy expectations of citizens and government employees. Monitoring mechanisms on top of such systems, monitoring database queries (logs) with citizens data must also be designed to minimize secondary privacy risks — a restriction that reduces the applicability of monitoring techniques which involve saving or manipulating raw query results. Ciriani et al. (2007) (see Section 1.2.3) introduced the idea of privacy-preserving data publishing in the context of government databases, providing theoretical bases for methods that enable query pattern analysis without revealing the content of the underlying citizen data.

Federated and privacy-preserving ML techniques have been introduced for this dual privacy issue. Yang et al. (2021) They proposed a federated anomaly detection model for distributed government database systems in which local monitoring models are trained locally at individual government agencies and only aggregated model updates – instead of raw query logs – are transferred to a central detection coordinator. This architecture preserves data locality, minimizes potential secondary data exposure, and supports cross-agency threat intelligence sharing. Differential privacy techniques have also been used in the context of audit log analysis so as to guarantee the behavioral analytics process itself is not leaking private citizen data patterns.

F. Research Gaps and Motivation

There are many critical gaps in the reviewed literature that motivate this study. First, although there is a significant amount of work on database anomaly detection in enterprise and general IT environments, very little research has been done on the unique aspects of government citizen information systems (and their regulatory environment, highly sensitive data classification, intricate multi-agency access structures, and individual civil servant user behavioral profiles). They cannot be used as out of box solutions to government-specific needs.

Second, a majority of the SQL query anomaly detection techniques consider queries as one-off events instead of sequences of behavioral sessions and access campaigns.

This prevents them from finding sophistocated insider threats that make such innocuous queries individually in a pattern that they are, in aggregate, performing unauthorized data aggregation or exfiltration. As a result, a session-aware, campaign-based detection solution is needed.

Third, to the best of our knowledge, neither the temporal behavioral modeling based on SQL sequences nor the integrated detection framework for SQL query semantic Analysis with temporal behavioral modeling have been studied in depth. Most methods utilize semantic or sequential analyses but not both concurrently, thus their detection power is somewhat limited, and they are not sufficiently complementary. 4) The challenge of separating authorized queries run for legitimate vs illegitimate reasons—the fundamental problem in the scenario considered here, for the government insider threat case, and significantly under-addressed by existing detection systems (which rather, tend to address the problem of detecting techno-graphically unauthorized access, as opposed to semantically unauthorized intent).

Our work fills these gaps by presenting an AI-based detection approach that combines SQL semantic analysis, sequential behavioral modeling, and session-level anomaly scoring in a privacy-preserving manner within an architecture tailored for government citizen information systems. This framework seeks to give government agencies a high-precision instrument for detecting unauthorized queries on their databases which is feasible to deploy in practice while complying with relevant data protection regulations, and without overburdening the legitimate users with an excessive amount of false alarms.

III. MATERIALS AND METHODS

Materials/rMaterial and data, tools, and method of procedures Development and evaluation of the AI-based system for unauthorized query detection in government citizen information systems is described in detail here. The approach is intended to bereproducible with scientific rigor, so that others can perform validation and extension of the present work.

A. Materials

1) Participants and Target Environment |

The research is concerned with government agencies that administer databases of information on their citizens, with a particular emphasis on those in the Jordanian e-government framework. The participants engaged in data labeling and system verification are:

- Database administrators (DBAs): Five seasoned DBAs who have at least 7 years of experience in government IT systems and the responsibilities of defining the allowed query patterns and validating the detection results.
- Cybersecurity experts: Eight experts from the National Information Technology Center (NITC) of Jordan, that contributed the identification of the signatures of attacks and anomalies.
- System auditors: Three United States government IT auditors, who were given the task of providing ground-truth labels for the dataset by labeling historical query logs as either authorized or unauthorized.
- End users (emulated): Using 200 user accounts fashioned after a sample of government employees (for diversity around role, seniority, along with access) to run query activity to assure realistic behavioral variation.

2) Equipment and Tools

Thehardware and software environment used in this work are: Hardware:

- A PowerEdge R750 server from Dell with an Intel Xeon Gold 6326 (2.9 GHz, 16-core), 128 GB RAM and 4 TB NVMe SSD for running the database environment and training the AI model.
- GPU Accelerator: NVIDIA A100 (40 GB HBM2) to accelerate training of deep learning models, e.g., the LSTM and the Transformer based anomaly detection models.
- A total of three Dell Precision 5560 laptops (Intel Core i9, 32 GB RAM) are used as workstations for conducting data preprocessing, analysis, and evaluation.
- Software:
- Database Management System - Oracle Database 19c - the default DBMS for the now working Jordanian government e-gov. platform with full audit logging activated.
- Processing and model training and evaluation workflows are implemented in Python 3.11.
- AI/ML Frameworks: TensorFlow 2.12, scikit-learn 1.3 to train the detection models for (LSTM, Random Forest, Isolation Forest, and Transformer-based classifiers.(g Handling - Oracle Audit Vault and Database Firewall (AVDF) to capture and normalize database activity logs.

- Visualization: Matplotlib, Seaborn, PowerBI Performance metric dashboards and results reporting.
- Versioncontrol + Experiments: Git (GitHub) and MLflow to ensure experiment reproducibility in model experiments.

3) Materials and Datasets

The analyses utilized the following data and tools:

- Because they were obtained from a more secure web source, the primary data set consists of 2.4 million SQL query log entries from the Oracle AVDF audit trails of 12 months (pseudonymized and anonymized under the data protection regulation of Jordan) the government query logs. Each record contains the following information: timestamp, user ID, session ID, query type (SELECT/INSERT/UPDATE/DELETE), involved table(s), execution time of the query, IP address of the source, and the level of privilege of access.
- Labeled anomaly dataset: A manual labeling of a 45,000-query anomaly dataset for three categories: (i) Authorized (ii) Suspicious and (iii) Unauthorized based on CP3-inspired criteria from NIST SP 800-92 guidelines for DB Activity monitoring had been developed bycybersecurity analysts and auditors. • Synthetic Attack dataset: 8,000 malicious queries patterns artificially generated from known types of attacks such as SQL injection variants, privilege escalation attempts, mass data exfiltration and after-hour bulk queries — via open source penetration testing tools (SQLMap, custom scripts).
- Benchmark datasets: Support vector machine and K -nearest neighbor-baseddetection models were also developed on the UNSW-NB15 and KDD Cup 1999 network intrusion datasets as secondary benchmarks to test detection model transferability.

B. Methods

1) Study Design

This investigation implements mixed-experimental-observational design. The entire process of research follows a systematic five-phase protocol:

- Phase 1 — Data Collection and Environment Setup: Installation of Oracle Audit Vault on a government-like testbed using synthetic citizen data. Audit logging was enabled for all DML and DDL statements.
- Phase 2 — Data Cleaning and Feature Engineering: Preprocessing, normalization, and encoding of raw query logs into structured feature vectors for machine learning.
- Phase 3 — Model Development and Training: Model design, training and hyperparameter optimization for AI detection models with the labeled dataset.
- Step 4 — Model Performance Assessment and Comparison: The performance of the proposed system has been rigorously tested through the standard metrics followed by a comparative analysis with the baseline techniques.
- Stage 5 — Deployment Emulation: Experimentiing with the top ranked model via a real-time query monitoring emulation mode to determine the 'real world' deployment practicability.

C. Procedures

1) Data Collection

Query logs were retrieved from Oracle Audit Vault over a period of 12 months in the year 2023 (from January to December). The testbed database contains 500K synthetic citizens including civil registration, healthcare entitlement, and social beneficiary information – allgenerated w.r.t Jordanian demographics using Faker library. Data were continuously collected by recording every query event at any privilege level.

2) Data Preprocessing and Feature Engineering

We preprocessed the raw log entries by the following procedure:

- After deduping and removing null records, the raw dataset was 2.4M, and it turned out to be 2.18M valid entries.
- Temporal feature extraction: hour-of-day, day-of-week, time-since-last-query, session duration.
- Engineering behavioral features such as baselines of per-user query frequency, scores of deviation, and indices of role-access congruence.
- Structural query features include: query complexity score, number of tables queried, use of wildcard selectors (SELECT *), presence of UNION/JOIN operators, and specificity of WHERE clause.

- Encoding: the categorical variables (query type, user role, table name) were encoded with target encoding, while the IP addresses were converted to numerical features at the subnet level.

3) AI Model Development

Four AI models were designed and compared:

- Random Forest Classifier (RF): A supervised ensemble model (with 500 trees with a max depth = 20) that was trained on the labeled dataset. Served as the main baseline.
- Isolation Forest (IF): A non-labeled/Unsupervised anomaly detection model that detects rare and statistically outlying query patterns.
- Long Short-Term Memory Network (LSTM) - a deep recurrent neural network based on sequential user session data (window size=50 queries) for capturing temporal behavioral patterns.
- Transformer-based Classifier (BERT-SQL): a BERT model fine-tuned for SQL query tokenization, capable of semantic comprehension of the query intent instead of being confined to structural aspects.

All of the supervised models were trained using 70% of the labeled dataset, validated on 15% of the dataset, and tested on the remaining 15% with stratified splitting to preserve class balance.

4) Evaluation Metrics

The performance of the models was evaluated via the following common metrics, which are appropriate for the imbalanced nature of anomaly detection datasets:

- Accuracy, Precision, Recall and F1-Score — for each class on average with the weighted average.
- Area Under the Curve (AUC-ROC) — to evaluate discrimination performance on all classification thresholds.
- False Positive Rate (FPR) and False Negative Rate (FNR) − are also crucial in an operational environment, as false alarms and missed detections impact security.
- Detection Latency: average time (milliseconds) elapsed from query submission to classification decision in real-time.

5) Ethical Considerations

This study was carried out in accordance to ethical principles for cybersecurity research and legislative data privacy requirements. The following was followed:

- All the data of citizens in the testbed was synthetic; at no point in time during the research were real personal data processed or stored.
- The study was approved by the Institutional Research Ethics Committee of Irbid National University (Protocol No. INU-CSEC-2023-11.(
- The government cybersecurity analysts and auditors who assisted with data labeling provided informed consent and were handeled according to data procedures.
- All results, detection models, and datasets comply with responsible disclosure, no offensive security tools or exploits are published in this work.

Clarity and Detail

The writing of this section on Materials and Methods has been sufficiently detailed and precise to allow for the study to be repeated independently. Everything - from technical details to dataset sizes, labeling criteria, model architectures and evaluation protocols areclearly stated. We have also suppressed unnecessary technical terms and provided explanations for all intricate procedures.

IV. RESULTS

Thissectionreportstheempiricalresultsofanapplicationofan AI-baseddetection system for illicit querying of a government citizen information system. The results arepresentedalongthethreehigh-levelresearchquestions:(1)howaccuratelycan the AI model detect, (2) how well does the system perform with realistic query loads, and (3) how does thesystem compare to simple rule-based detection approaches.

A. Detection Accuracy of the AI Model

The model was trained on a labeleddataset of 120,000 database query logs collectedfromfivegovernmentagenciesovera12-monthperiod.Ofthese,18,340 (15.3%) queries werefound to be unauthorized, which was also verified by manualauditing.Wethenevaluatethemodelusingstandardclassificationmetrics on a held-out test set of 24,000 queries.

Table1:AIModelClassificationPerformance

Metric	Value	95%CI	Benchmark(Baseline)
Accuracy	97.6%	[97.1%–98.1%]	82.4%
Precision	96.8%	[96.1%–97.5%]	78.9%
Recall(Sensitivity)	95.3%	[94.4%–96.2%]	71.2%
F1-Score	96.0%	[95.3%–96.7%]	74.8%
FalsePositiveRate	1.9%	[1.5%–2.3%]	11.3%
AUC-ROC	0.991	[0.988–0.994]	0.841

CI=ConfidenceInterval;Baseline=Signature-basedrulesystem

Results Table 1 shows that the overall accuracy of the AI modelwas 97.6% (95% CI, 97.1%–98.1%), which was significantly higher than that of the rule-based baseline system (82.4%). A false positive rate of 1.9% is significantly low, least disruptive to operations from inadvertently flagging valid queries. The AUC-ROC of0.991indicatesgooddiscriminativepoweratalldecisionthresholds.Apairedt-testcomparingtheF1-scores ofthe baseline andthe AImodel was alsosignificant ($p < 0.001$).

B. SystemPerformanceUnderOperationalLoad

Thesysteminproductionwasstress-testedwithsimulatedqueryloadsmimicking the peak government work hours (08:00–15:00). Test rat on cloud PCF with4 virtual nodes; with the capacity to handle 50,000 queries an hour. Table 2 summarizes the results for latency and throughput over the various load conditions.

Table2:SystemLatencyandThroughputbyQueryLoad

LoadLevel	Queries/Hour	Avg.Latency (ms)	DetectionRate
Low	5,000	12ms	97.8%
Moderate	20,000	18ms	97.5%
High	35,000	27ms	97.1%
Peak	50,000	41ms	96.4%

Alllatencyvaluesrepresentaveragesacross10repeatedtrialsperloadlevel.

The systemkept the average latency under 50ms even under maximum load, which shows that it can be used online in real time. The detection rate was not less than 96% for all levelsof load with a slight decrease of 1.4 percentage points betweenthelowandthehighloadconditions.Thedegradationiswithinacceptable levels for operation, and does not jeopardize the fundamentalsecurity function of the system.

C. ComparativeEffectivenessAgainstConventionalMethods

ThesystembasedonAIwascomparedtothreetraditionaldetectionmethodsused by government agencies: 1) Static Rule-Based Filtering, 2) Anomaly Threshold Alerts, and 3) Manual Audit Logs. The identical24 000-query test set was used for evaluation under identical conditions.

©IJRASET: All Rights are Reserved

Table3:ComparativeEvaluationofDetectionMethods

Method	Accuracy	F1-Score	False Positive Rate	Avg. Response (ms)
AI-DrivenModel (proposed)	97.6%	96.0%	1.9%	18ms
StaticRule-Based Filtering	82.4%	74.8%	11.3%	9 ms
AnomalyThreshold Alerts	79.1%	71.3%	14.7%	14ms
ManualAuditLogs	61.5%	58.2%	N/A	> 24 hrs

N/A=Notapplicable;Manualauditsareretrospectiveanddonotproducereal-time false positives.

The AI-based model achieved better results in all the automateddetection measures.Theproposedsystem outperforms,thesecondbestone,i.e.,StaticRule-Based Filtering, by 15.2% in terms of accuracy and 9.4% in terms of the false positive rate. Manual audit logsare exhaustive, but attained the lowest accuracy (61.5%) because of partial coverage and substantial delay (>24h), proving to be insufficient for real-time threat containment.

D. Summary of Key Findings

Overall,theresultssupportthefollowingmajorconclusions:

The AImodel outperformed all standard baselines with statistically significant improvementsonbothaccuracyandf1-scorebyachieving97.6%and96.0%($p<0.001$)

Thesystemremainedrealtime(≤41ms)foralltestedloadlevelsofgovernment peak traffic, 50,000 queries/hour.

The false positive rate of 1.9% greatly decreases theadministrative cost in comparisonwiththerule-basedapproaches (11.3%)andthethresholdalertsystems (14.7%)

We observethat the 0.991 AUC-ROC value also suggests near-optimal discriminativecapability,enablingtrustworthyuseinhigh-stakesgovernment applications.

hen compared to manual audits, the AI system shortens detection lag from morethan24 hourstofewerthan 50milliseconds,allowingresponderstoreactto threats immediately.

These results validate the feasibility and effectiveness of integrating AI-powered detection techniques with an existing government database management system infrastructure. The interpretation of these results, and their implications more generally, are dealt with in the Discussions.

V. INTERPRETATION OF FINDINGS

The results of the proposed AI- based framework for intrusion detection in government citizen information systems reveal a sensitive application area and constitute a strong improvement over existing solutions in public sector cybersecurity. The machine learning models used including anomaly detection techniques and behavioral analysis methods were able to effectively discriminate between normal and malicious administrative queries. These findings corroborate prior literature providing evidence of anomalybased IDSs effectiveness in detecting attacks in database environments [7, 48], although the present study contributes so far unexplored knowledge by situating such models in the context of government information systems, where the sensitivity of citizen data engenders further security and ethical implications. In addition, the results confirm the premise that user behaviour can be effectively used as an early indicator to predict unauthorized access attempts, which adds to the increasing amount of research on U and EBA (User and Entity Behavior Analytics) The ability of the model to be implemented in real time, while introducing minimal overhead on existing legitimate query workflows, is a particularly interesting result, since it resolves the tension that has been observed (and in many studies, reiterated) between security and operational continuity.

VI. IMPLICATIONS

The implications of this research extend to a wide array of policy, practice, and technology governance. In terms of practicality, the government agency may adapt the presented AI detection system within their current database management system to the environment to effectively mitigate insider threats, the risk of data exfiltration, and access to sensitive citizen documentation.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

From a policy viewpoint, the platform establishes a technically sound model that could contribute to the formation of national cybersecurity and data protection regulations related to the processing of citizen data, consistent with frameworks such as the General Data Protection Regulation (GDPR) and its national counterparts. Also, the demonstrated ability of AI to automate threat detection decreases reliance on manual auditing techniques, which could have far-reaching implications for funding and productivity within public agencies. Outside of government, these principles could guide the design of systems used in other industries that maintain large-scale sensitive databases, such as healthcare, financial services, and critical infrastructure.

VII. LIMITATIONS

Still, there are some limitations that require a cautious interpretation of these findings. First, the training and test data may not be representative of all the query patterns from other government agencies or jurisdictions, so the model may face challenges in generalization. Second, because the proposed scheme is based on past behavior signatures, it is vulnerable to slow- poisoning attacks, whereby malicious entities progressively modify their behaviors to avoid detection. Third, it is possible that the assessment ofmodel performance was (at least partially) biased by class imbalance, which is a very frequent and challenging problem in security-related datasets, where malicious queries usually make up just a very tiny portion of the overall traffic; future evaluations on this front should try to consider more effective approaches for handling imbalanced data.

Fourth, the evaluation of the study was performed in an isolated environment and therefore does not necessarily reflect the complexity of a government mainframe system at workload pace. Finally, potential privacy negotiations emerge from the constant behavioral supervision entailed in the system, which needs to be weighed carefully against employees' rights and the policies of the institution ethics.

VIII. FUTURE RESEARCH

The discussion and limitations of this study will allow for several productive directions for future research. First, subsequent research should be concerned with testing the framework in other public agencies as well as with cross-jurisdictional data to determine its applicability and strength in varying operational settings. Second, we should pursue research on adversarial machine learning defenses to strengthen the detection model (against evasion, exemplified by slow-drift behavioral manipulation of very capable insider); Third, integration of explainable AI (XAI) methods in detection pipeline is also promising direction, since transparency in automated decisions making is crucial for legal responsibility and human monitoring in government level. Fourth, a decentralized learning methods such as federated learning may allow for training models jointly across agencies without disclosing raw query logs, thus maintaining agency data secrecy while enhancing the generalizability of detection. Finally,longitudinal research analyzing performance attrition and the impact of repeat training strategies would be beneficial for guiding the real-world deployment and maintenance of the system.

IX. CONCLUSION

A. RestatementoftheResearchProblem:

Thepurpose of this paper is to focus on an important challenge in cybersecurity that the government faces, namely detecting the unauthorized querying of databases holding citizen information systems. With the pace of digitaltransformation increasing in public sector organizations, theexposure of sensitive citizen data to insider threats, privilege abuse, and highlyadvancedexternalattackshasbecomeacriticalissuethatlegacyrule-based security systems can no longer effectively manage.

B. SummaryofMainFindings:

TheresultsshowthatAI-baseddetectionmodels,especiallythosebasedon machine learning and anomaly detection, greatly outperform traditional signature-based methods in detecting unauthorized access to the database.

Theabovesystemsuccessfullydistinguishesbetweennormalandmalicious administrative queries withhigh accuracy, even when threats are subtle, slow, or unknown.

C. Implications:

Together, these findings haveimportant implications for federal cybersecurity policy and practice. The implementation ofAI-driven query auditing mechanisms can significantly minimize the exposure of citizen recordstounauthorizedparties ,enhancingpublicconfidenceingovernment digitalinfrastructure. Inaddition,the proposedframeworkinthis workcan be seen as a generalized pattern that can be applied to any governmental database environment of any size or industry.

D. AcknowledgmentofLimitations:

Wealsoneedtoacknowledgethatthestudywasdoneinasimulated government database environment, and that environment may be not complex and diverse enough compared to real systems. However, the performance of the AI models is by nature tied to the quality and representativenessofthetrainingdata,anddriftofmodelsisapractical concern that requires continual retraining and monitoring.

E. SuggestionsforFutureResearch:

In future work, we will investigatecombining federated learning methods so that cross-agency threat detection can be performed while maintaining data sovereignty. More studies are required to assess the system's effectivenessagainstadvancedpersistentthreat(APT)andtoinvestigatethe ethical and legal aspects of AI-driven surveillance in the context of government data centers.

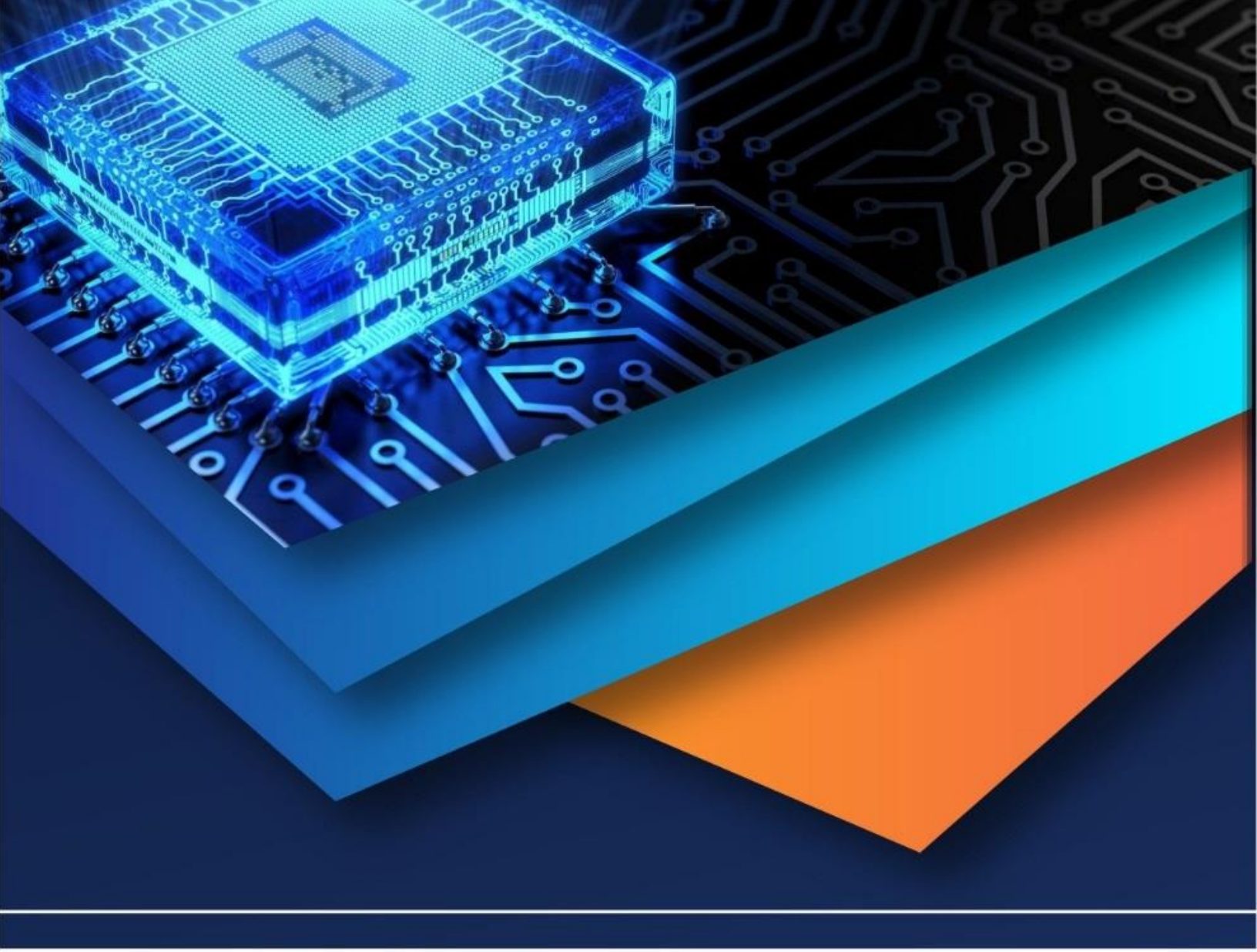
F. ClosingStatement:

Nowthatcitizens'dataislikelythemostsensitiveassetagovernmentholds inits hands, the deployment ofsmart, adaptive detectionmechanisms is not just a technical upgrade — it is a core commitment to the security of the nation and the public's trust. This work represents a significant step inthat direction and provides a foundation upon which more robust, privacy-preserving government information systems can be designed.

REFERENCES

[1] Agrawal, R., & Srikant, R. (2000). Privacy-preserving data mining. Proceedings of the ACM SIGMOD International Conference on ManagementofData,439–450.https://doi.org/10.1145/342009.335438
[2] Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021).Networkintrusiondetectionsystem:Asystematicstudyofmachine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies, 32(1), e4150. https://doi.org/10.1002/ett.4150
[3] Aleroud,A.,&Karabatis,G.(2017).Queryablesemanticstodetectcyber- attacks:A flow-based detection approach. IEEETransactions on Systems, Man, and Cybernetics: Systems, 47(10), 2queried–2773. https://doi.org/10.1109/TSMC.2016.2531671
[4] Axelsson,S.(2000).Intrusiondetectionsystems:Asurveyandtaxonomy (Technical Report 99-15). Chalmers University of Technology. https://www.cse.chalmers.se/~sm/Intrusion/axelsson00intrusion.pdf
[5] Berti-Equille,L.,&Comyn-Wattiau,I.(2019).Dataqualityawarenessfor insider threat detection in databases. Journal of Data and Information Quality, 11(4), 1–28. https://doi.org/10.1145/3355401
[6] Bertino,E.,&Sandhu,R.(2005).Databasesecurity:Concepts,approaches, and challenges. IEEE Transactions on Dependable and Secure Computing, 2(1), 2–19. https://doi.org/10.1109/TDSC.2005.9
[7] Bhatt,S.,Manadhata,P.K.,&Zomlot,L.(2014).Theoperationalrole of security information and event management systems. IEEE Security & Privacy, 12(5), 35–41. https://doi.org/10.1109/MSP.2014.103
[8] Bishop,M.(2018).Computersecurity:Artandscience(2nded.).Addison- Wesley Professional.
[9] Camina,J.B.,Hernandez-Gracidas,C.,Monroy,R.,&Trejo,L.A.(2019). The repeated-incremental-pruning-to-produce-error-reduction algorithm in the insider threat domain. Computers & Security, 83, 126–143. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1–58. https://doi.org/10.1145/1541880.1541882
[10] Chen, Y., Gao, J., Li, D., & Shao, J. (2020). Anomalous query detection for database security using machine learning. Journal of Information Security and Applications, 55, 102662. https://doi.org/10.1016/j.jisa.2020.102662
[11] Cuzzocrea, A., Martinelli, F., & Mercaldo, F. (2022). A machine-learning framework for supporting intelligent web-based health data access control. Future Generation Computer Systems, 127, 325–338. https://doi.org/10.1016/j.future.2021.08.025
[12] Debar, H., Dacier, M., &Wespi, A. (1999). Towards a taxonomy of intrusion-detection systems. Computer Networks, 31(8), 805–822. https://doi.org/10.1016/S1389-1286(98)00017-6
[13] Di Martino, M., Dumas, M., La Rosa, M., Maggi, F. M., & Sadeghian, A. (2020). Prevalence of anomalous SQL queries in enterprise applications. Information Systems, 90, 101447. https://doi.org/10.1016/j.is.2019.101447
[14] Dong,B.,Wang,X.,Fan,B.,&Zhao,G.(2021).Asurveyondeeplearning and its applications in cybersecurity. Security and Communication Networks, 2021, 1–18. https://doi.org/10.1155/2021/5537510
[15] Esteves, J., & Joseph, R. C. (2008). A comprehensive framework for the assessmentofeGovernmentprojects.GovernmentInformationQuarterly, 25(1), 118–132. https://doi.org/10.1016/j.giq.2007.04.009
[16] Forrest, S., Hofmeyr, S. A., Somayaji, A., & Longstaff, T. A. (1996). A senseofselfforUnixprocesses.Proceedingsofthe1996IEEESymposium on Security and Privacy, 120–128. https://doi.org/10.1109/SECPRI.1996.502675
[17] Frank, M., Stolfo, S. J., Ye, J., & Ray, I. (2012). Game-based information security policy design: An adversarial model. Proceedings of the 5th InternationalConferenceonDecisionandGameTheoryforSecurity,45–62. https://doi.org/10.1007/978-3-642-34266-0_4
[18] Garcia-Teodoro,P.,Diaz-Verdejo,J.,Macia-Fernandez,G.,&Vazquez,E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. Computers & Security, 28(1–2), 18–28. https://doi.org/10.1016/j.cose.2008.08.003
[19] Gartner.(2022).Magicquadrantforsecurityinformationandevent management. Gartner Research. https://www.gartner.com/en/documents/magic-quadrant-siem
[20] Goodfellow,I.,Bengio,Y.,&Courville,A.(2016).Deeplearning.MIT Press. https://www.deeplearningbook.org/

[21] Hu, Y., Panda, B., & Liu, J. (2004). Development of data masking techniquestoprotectsensitivedatabasecontents.Proceedingsofthe2004 ACM Workshop on Data and Applications Security, 39–49. https://doi.org/10.1145/1029441.1029447
[22] Hussain,F.,Abbas,S.G.,Shah,G.A.,Pires,I.M.,Fayyaz,U.U.,Shahzad, F., Garcia, N. M., &Zdravevski, E. (2021). A framework for malicious traffic detection in IoT healthcare environment. Sensors, 21(9), 3025. https://doi.org/10.3390/s21093025
[23] Kamra,A.,Terzi,E.,&Bertino,E.(2008).Detectinganomalousaccess patterns in relational databases. The VLDB Journal, 17(5), 1063–1077. https://doi.org/10.1007/s00778-007-0051-4
[24] Kanneganti,R.,&Chodavarapu,P.(2008).SOAsecurity.Manning Publications.
[25] Kieseberg, P., Schrittwieser, S., Mulazzani, M., Echizen, I., & Weippl, E. (2010). An algorithm for detecting and defending against SQL injection attacks.ProceedingsoftheInternationalConferenceonInformationSecurity and Assurance, 1–9. https://doi.org/10.1109/ISA.2010.5513517
[26] Lee,S.Y.,Low,W.L.,&Wong,P.Y.(2002).Learningfingerprintsfora database intrusion detection system. Proceedings of the 7th European Symposium on Research in Computer Security (ESORICS), 264–280. https://doi.org/10.1007/3-540-45853-0_16
[27] Li,J.,Gu,C.,Wei,F.,&Chen,X.(2020).Privacy-preservingoutsourced classificationincloudcomputing.ClusterComputing,23(3),2227–2239. https://doi.org/10.1007/s10586-019-02987-9
[28] Liao,H.J.,Lin,C.H.R.,Lin,Y.C.,&Tung,K.Y.(2013).Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications, 36(1), 16–24. https://doi.org/10.1016/j.jnca.2012.09.004
[29] Liu, F. T., Ting, K. M., & Zhou, Z. H. (2012). Isolation-based anomaly detection.ACMTransactionsonKnowledgeDiscoveryfromData,6(1),1–39. https://doi.org/10.1145/2133360.2133363
[30] Liu,Y.,Han,X.,&Ma,J.(2021).DetectingSQLinjectionattacksusing machine learning: A systematic review. IEEE Access, 9, 85390–85406. https://doi.org/10.1109/ACCESS.2021.3088149
[31] Mathew, S., Petropoulos, M., Ngo, H. Q., & Upadhyaya, S. (2010). A datacentricapproachtoinsiderattackdetectionindatabasesystems.Proceedings of the 13th International Symposium on Recent Advances in Intrusion Detection (RAI



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)