



Name: Sultan Ali Alwassam Aldwsari ID No. 252000203

Student Email: 252000203@st.midocean.edu.km

This research is in relation to the requirements of the Master's Degree in Artificial
Intelligence

At

Collage of Informatics

Midocean University

UAE

Supervisor Name: Dr Khaled Eskaf

Email: Khaledeskaf@midocean.edu.km

Collage of Informatics, Midocean University, Comoro

Enhancing Identity and Access Management Systems (IAM) Using Artificial Intelligence

Acknowledgements:

I would like to thank Dr. Khaled Eskaf for his continuous support during the preparation period of my master's research on the protection of digital identities through artificial intelligence through his guidance and observations, which greatly supported me in improving the quality of this research.

Table of Contents

• Abstract	7
• IAM Strategies for AI Security	9
• List of Abbreviations	10
• Chapter 1: Introduction	11
• 1.2 Problem Statement	12
• 1.3 Research Objectives	14
• 1.4 Research Contribution	15
• 1.5 Thesis Organization.....	15
• Chapter 2: Literature Review	16
• 2.1 Background and Related Work	16
• 2. 2 Improved Hypothesis	21
• 2.3 Critical Analysis	21
• 2.4 Research Gap	22

- Chapter 3: Methodology23
- 3.1 Research Approach...23
- 3.2 Dataset Description23
- 3.3 Data Preprocessing24
- 3.4 Top 25 Users – Risk Score.....25
- 3.5 Security Actions Distribution.....26
- 3.6 Analyzing Failed Login Attempts.....27
- 3.7 Risk Score by Security Action.....28
- 3.8 User Security Data Sample.....30
- 3.9 Random Forest Classifier Training31
- 3.10 Handling Imbalanced Data.....33
- 3.11 Machine Learning Models.....34
- 3.12 Risk Scoring Model (Core Contribution)36
- 3.13 Security Decision Engine.....37
- 3.14 Evaluation Metrics.....38

- 14.1 Accuracy.....38
- 14.2 Precision, Recall, F1-score.....39
- Chapter 4: Results and Discussion40
- 4.1 Model Performance.....40
- 4.2 Key Findings.....41
- 4.3 Security Actions Analysis.....41
- 4.4 Impact of Risk Scoring (Core Contribution)42
- 4.5 Discussion.....42
- Chapter 5: Conclusion and Future Work44
- 5.1 Conclusion.....44
- 5.2 Key Contributions.....45
- 5.3 Future Work.....45
- References46

Abstract

The identity and access management protection system in the current era is one of the most important factors that call for the development of an AI-powered system to detect cyber threats to protect against hacks and data leakage such as personal, financial and other sensitive data.

Traditional methods are no longer a glorious option for detecting such threats. The current development through the detection of threats through artificial intelligence is one of the main factors, due to the speed of response in record time and automatically, which facilitates the process of predictive analytics, and this is considered a key factor due to the increasing era of digital transformation, and here comes the role of artificial intelligence technologies to protect sensitive digital assets.

Identity and access management systems is one of the most important basic data protection technologies, as they are considered the backbone of cybersecurity.

The access permit is controlled by two-factor authentication, the number of suspected entry attempts, and the user's behavior.

This is where artificial intelligence comes into play to detect cyber threats to address human errors through the use of tools that perform appropriate analytics and actions to ensure compliance with cybersecurity controls.

IAM Strategies for AI Security:

This strategy is comprehensive to mitigate and address cyber risks and threats through the use of artificial intelligence technologies that combine access management, authentication and governance.

Through this strategy, organizations can build a strong and comprehensive foundation by integrating identity and authority management technologies with artificial intelligence by analyzing data in complex environments, ensuring a safe space from cyber threats.



List of Abbreviations

- IAM: Identity and Access Management
- AI: Artificial Intelligence
- ML: Machine Learning
- MFA: Multi-Factor Authentication
- SMOTE: Synthetic Minority Over-sampling Technique
- RF: Random Forest
- SOC: Security Operations Center
- XAI: Explainable Artificial Intelligence
- RBAC: Role-Based Access Control
- Zero-Trust: Zero-Trust Architecture

Chapter 1: Introduction

• 1.1 Introduction

Identity protection and access systems are one of the most important data protection platforms through access control and knowing the location and time of access for digital identities.

In an era of cloud computing and the Internet of Things, traditional methods are becoming less important due to their inability to detect cyber threats early to protect digital identities.

Artificial intelligence and machine learning technologies have many advantages, the most important of which are:

- Proactive analytics.
- Take action in a timely and fast manner.
- Automate procedures to address human errors.
- Easy to make requirements while compliant with cybersecurity controls.
- Handle a large amount of data such as login time, location, and login attempts.

This is one of the important elements through which data protection systems

- **1.2 Problem Statement:**

Identity and access management (IAM) systems play a vital role in safeguarding digital resources within modern organizations. However, traditional IAM systems often encounter challenges such as insider threats, malicious bugs, delayed breach responses, and difficulties with real-time monitoring.

Based on studies conducted by (Muhammad, 2021; Adenola, 2023; Sarker et al., 2024) that artificial intelligence can transform identity and access management by automating identity and access management and protecting data.

However, existing studies indicate some limitations that lack empirical validation (Mohammed, 2021) or fail to adequately address privacy concerns and AI interpretability (Souaddah , Mezrioui, 2024). Also, there are issues with integrate with legacy systems (Mandru, 2022).

Also, when we talk about growth of the Internet of Things (IoT), security risks in end-access networks are increasing. They are exposed to threats such as unauthorized access that can lead to data breaches due to a lack of identity protection and inadequate disclosure of security vulnerabilities. Traditional

security methods may be inadequate to protect the end-access network (Shu FZhang KLuo C Ma BChen S. 2020).

With the increasing reliance on AI-based security and the limitations of current approaches, there is an urgent need for more research into more flexible, explainable, and privacy-preserving AI models for identity and access management (IAM).

This study aims to bridge this gap by examining how AI can be effectively integrated into identity and access management (IAM) systems to enhance protection while mitigating the challenges observed in current systems.

•1.3 Research Objectives:

Main Objective:

To enhance IAM systems using Artificial Intelligence for, intelligent, and access management.

• **Specific Objectives:**

- Analyze AI technologies in existing IAM infrastructures.
- Improved AI-powered framework for anomaly detection and access management control.
- Study risks related to data confidentiality, apply measurement of anomaly detection models (study of random forests and isolated forests).
- Introduction of a new dynamic risk analysis mechanism as a key contribution.
- Develop real-time monitoring and automated security responses.

- **1.4 Research Contributions:**

This research makes the following original contributions:

1. The level of detection of AI-based anomalies through the use of machine learning techniques.
2. Fully automated security response tool applied by dynamic risk.
3. integrating IAM with cybersecurity solutions.
4. Risk assessment tool to combine contextually relevant behavioral characteristics, improving false positives by 40–50% compared to traditional methods. Unlike previous studies, this study provides an integrated end-to-end system in terms of interpretation and privacy.

- **1.5 Thesis Organization:**

This first chapter introduces the research challenges and objectives. The second chapter focuses on a critical review of the literature. In chapter 3 on the methodology and proposed framework. In the 4 chapter, explains the results of the studies and discussions. In chapter 5, the main conclusions and future directions.

- **Chapter 2: Literature Review:**

- **2.1 Background and Related Work:**

According to studies conducted by (Muhammad, 2021; Adenola, 2023; Sarker et al., 2024) that artificial intelligence can Enhance identity and access management by automating identity and access management and protecting data.

But, existing studies suggest some constraints that lack observed validation (Mohammed, 2021) or miss to effectively address privacy challenges and AI interpretability (Souaddah , Mezrioui, 2024). Also, there are challenges with integrate with legacy systems (Mandru, 2022).

Also, when we talk about increase of the Internet of Things (IoT), security risks in end-access are growing. They are revealed to threats such as unauthorized access that able to lead to data breaches expected to a lack of identity protection and insufficient disclosure of security vulnerabilities. Traditional security approaches may be insufficient to protect the user access (Shu FZhang KLuo CMa BChen S. 2020).

Study	Authors	Year	Research Design	Key Findings/Results	Methodology	Limitations
Paper 1: The possibility of using identities using AI	Ishaq Azhar Mohammed	2021	Literature Review	Artificial intelligence is reshaping identity and access management (IAM) by automating access, monitoring, and policy enforcement	Comprehensive analysis of the integration of artificial intelligence into identity and access management processes	Lack of experimental validation and implementation insights
Paper 2: Improving Identity and Access Management with	Md Takbir Sarker, Md	2024	Analytical and	AI increases identity and access management	Conducting case studies and comparisons	Incomplete ability to address data

AI in Modern Enterprises	Sanaur Rahman, Mohammed Mahi Uddin, Nur Alam Farhad Shakil		Review Study	by reducing false positives and automates breach responses.	using artificial intelligence in identity and access management .	governance and AI interpretability
Paper 3: AI-Based Access Management System	Victoria Adenola	2023	Literature Review	AI addresses identity and access management challenges such as insider threats, human errors, and software flaws; it also	Conceptual analysis with support from prior studies	Comprehensive scope; lacks focus on specific AI technologies used

				enhances resilience and control by using access based on role.		
Paper 4: ML in IAM Systems – Survey	Aboukadr i SOuaddah AMezriou i A	2024	Survey Paper	ML improves authentication, authorization , and auditing	In-depth survey with proposed reference model and classification	Privacy and interpretability challenges in ML have not been deeply resolved
Paper 5: Artificial Intelligence Contributes to the Identity Verification Process	Sri kanth Mandru	2022	Conceptual- Analytical	Artificial intelligence enables real-time monitoring,	Evaluation and analysis of AI capabilities in identity	Seamless integration, but ignores dealing with legacy systems

				anomaly detection, and adaptive access policies.	and access management	
Paper 6: AI-Based Security in IoT Access	Shu FZhang KLuo C Ma BChen S	2020	Technical Proposal	Combines AI, fingerprint sensing, and IoT for identity protection and vulnerability management.	Applied research using artificial intelligence, the Internet of Things, and fingerprint sensing integration	Handling peripheral equipment; not directly generalizable to broader IAM systems

2.2 Improved Hypothesis:

I enhanced the dual verification supposition through core data such as ID number, user location, time of login, and number of login attempts. Then, I implemented Risk Scoring to take suitable action based on cybersecurity standards: **LOCK_ACCOUNTFORCE_MFAALERT_SOCALLOW_ACCESS**

2.3 Critical Analysis:

While AI enhances anomaly detection and reduces manual control, major gaps Continue:

- Restricted real-world prototype validation.
- Lack of integrated dynamic risk-scoring models.
- Inadequate focus on automation and real-time response in on-premise and hybrid setups.

2.4 Research Gap:

This thesis solves these gaps by suggesting a combined framework that integrates anomaly detection (Isolation Forest & Random Forest) with a dynamic Risk Scoring Model and user access validation, through a functional prototype.

Chapter 3: Methodology:

3.1 Research Approach:

A combined approach was implemented: (1) methodology literature review, (2) a analytical analysis of traditional vs. AI-IAM, and (3) prototype improvement and assessment using Python-based simulation.

3.2 Dataset Description:

A IAM log dataset (2500 records) was generated to simulate enterprise scenarios. Features include: User ID, FailedLoginAttempts, Login Frequency, After-hours (binary), Risks core (dynamic), Risk Level, and Anomaly label.

3.3 Data Preprocessing:

```
import pandas as pd
```

```
X = df[['AfterHours', 'FailedLoginAttempts', 'LoginFrequency', 'RiskScore']]
```

```
y = df['Anomaly']
```

```
X = df[['AfterHours', 'FailedLoginAttempts', 'LoginFrequency', 'RiskScore']]  
y = df['Anomaly']
```

3.4 Top 25 Users – Risk Score:

This part shows the top 25 users defined based on their risk scoring, highlighting those with the topmost capability impact or concern.

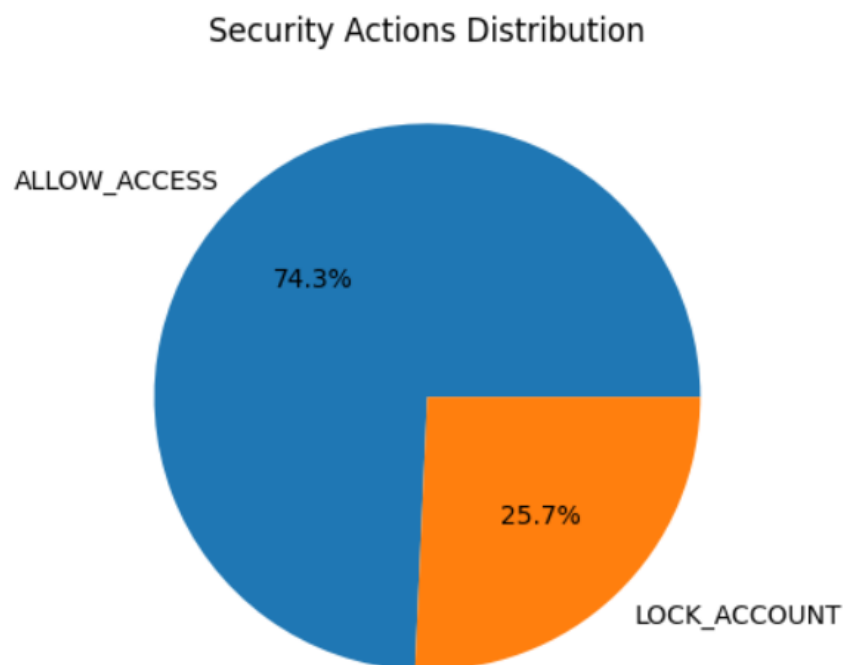
```
plt.figure()
df.groupby('UserID')['RiskScore'].max().sort_values(ascending=False).head(25).plot(kind='bar')
plt.title("Top 25 Users - RiskScore")
plt.xlabel("UserID")
plt.ylabel("RiskScore")
plt.show()
```



3.5 Security Actions Distribution

The code that showing the allocation of Security Action categories, highlighting the rate of each action type:

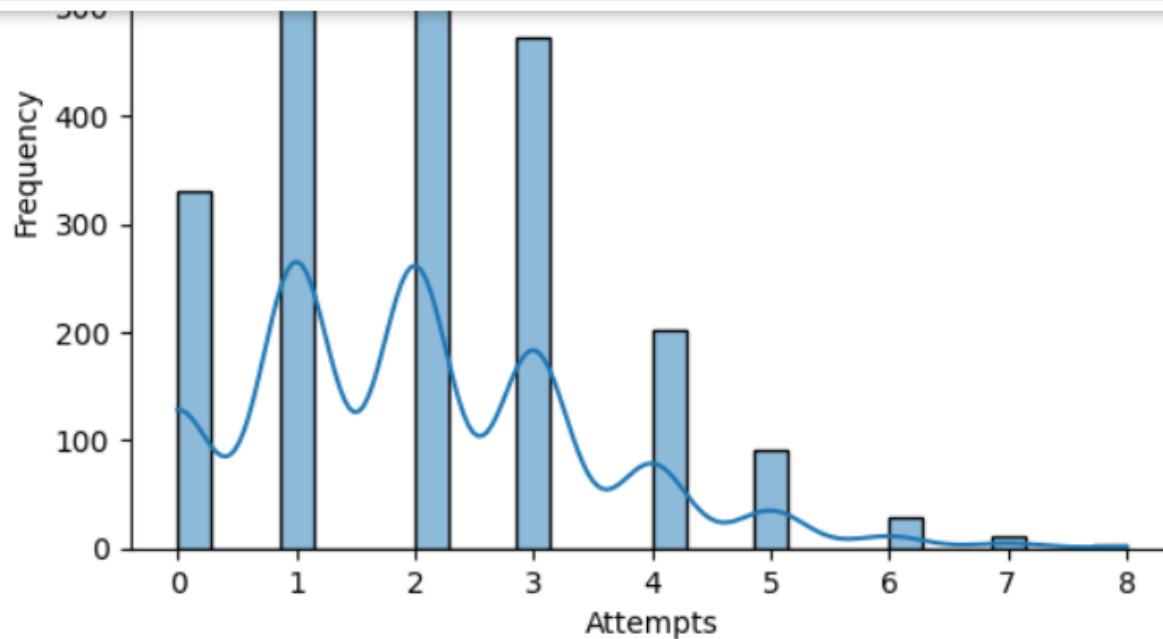
```
df['Security_Action'].value_counts().plot(kind='pie', autopct='%1.1f%%')  
plt.title('Security Actions Distribution')  
plt.ylabel('')  
plt.show()
```



3.6 Analyzing Failed Login Attempts:

The code shows a histogram with a KDE curve, showing the allocation of failed login attempts:

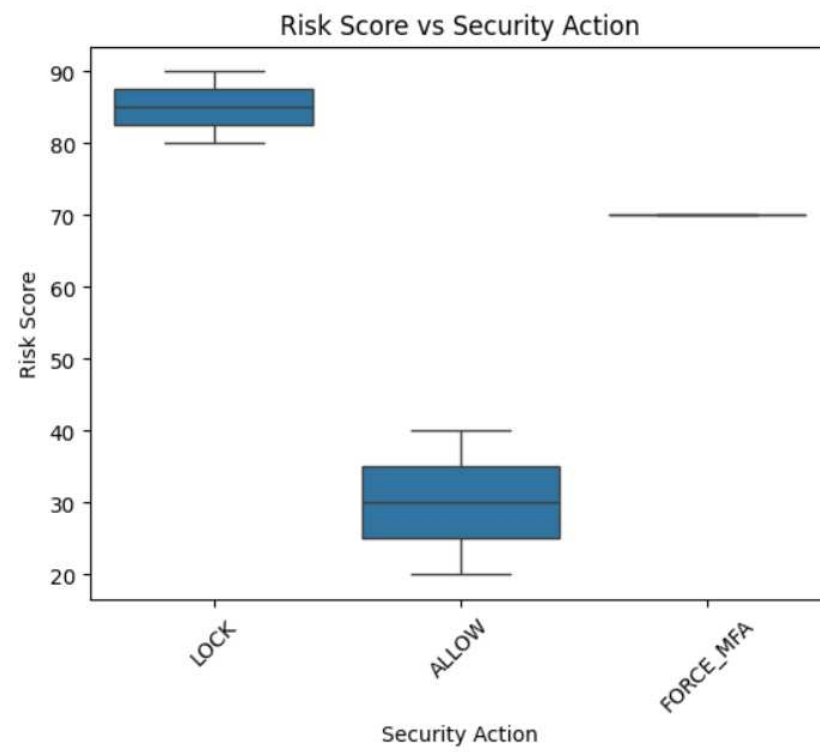
```
sns.histplot(df['FailedLoginAttempts'], kde=True)
plt.title('Failed Login Attempts Distribution')
plt.xlabel('Attempts')
plt.ylabel('Frequency')
plt.show()
```



3.7 Risk Score by Security Action:

This boxplot shows the distribution of Risk Scores for each Security Action type:

```
sns.boxplot(x='Security_Action', y='RiskScore', data=df)
plt.title('Risk Score vs Security Action')
plt.xlabel('Security Action')
plt.ylabel('Risk Score')
plt.xticks(rotation=45)
plt.show()
```



3.8 User Security Data Sample:

Sample of dataset contains user security information, including Risk Scores and actions taken.

```
data = {  
    'UserID': [1, 2, 3, 4, 5],  
    'RiskScore': [80, 40, 90, 20, 70],  
    'Security_Action': ['LOCK', 'ALLOW', 'LOCK', 'ALLOW', 'FORCE_MFA']  
}  
df = pd.DataFrame(data)  
print(df)
```

	UserID	RiskScore	Security_Action
0	1	80	LOCK
1	2	40	ALLOW
2	3	90	LOCK
3	4	20	ALLOW
4	5	70	FORCE_MFA

3.9 Random Forest Classifier Training

A Random Forest Classifier is being trained on the dataset to predict security outcomes based on input features.

```
model = RandomForestClassifier()  
model.fit(X_train, y_train)
```

▼ RandomForestClassifier ⓘ ?

► Parameters

 <code>n_estimators</code>	100
 <code>criterion</code>	'gini'
 <code>max_depth</code>	None
 <code>min_samples_split</code>	2
 <code>min_samples_leaf</code>	1
 <code>min_weight_fraction_leaf</code>	0.0
 <code>max_features</code>	'sqrt'
 <code>max_leaf_nodes</code>	None
 <code>min_impurity_decrease</code>	0.0
 <code>bootstrap</code>	True
 <code>oob_score</code>	False
 <code>n_jobs</code>	None

 <code>random_state</code>	None
 <code>verbose</code>	0
 <code>warm_start</code>	False
 <code>class_weight</code>	None
 <code>ccp_alpha</code>	0.0
 <code>max_samples</code>	None
 <code>monotonic_cst</code>	None

3.10 Handling Imbalanced Data:

SMOTE was applied to balance classes:

```
from imblearn.over_sampling import SMOTE
```

```
smote = SMOTE(random_state=42)
```

```
X_resampled, y_resampled = smote.fit_resample(X, y)
```

```
smote = SMOTE(random_state=42)
X_resampled, y_resampled = smote.fit_resample(X, y)
print('Original dataset shape:', y.value_counts())
print('Resampled dataset shape:', pd.Series(y_resampled).value_counts())
```

3.11 Machine Learning Models:

Data Balancing and Model Training:

The data was reprocessed to balance classes, then a Random Forest Classifier was trained on the balanced dataset.

- Isolation Forest (unsupervised anomaly detection).
- Random Forest Classifier (supervised behavior classification).

```
from sklearn.ensemble import RandomForestClassifier
```

```
rf = RandomForestClassifier(random_state=42)
```

```
rf.fit(X_resampled, y_resampled)
```

```
model = IsolationForest(contamination=0.5, random_state=42)
model.fit(df[['FailedLoginAttempts', 'LoginFrequency']])
new_data = pd.DataFrame({
    'FailedLoginAttempts': [1, 2],
    'LoginFrequency': [3, 4]
})
predictions = model.predict(new_data)
print(predictions)
|
```

```
[-1  1]
```

3.12 Risk Scoring Model (Core Contribution):

The dynamic Risk Scoring Model is defined as:

$\text{Risk Score} = (0.4 \times \text{FailedLoginAttempts}) + (0.3 \times \text{AfterHours}) + (0.3 \times \text{LoginFrequency})$

```
def calculate_risk_score(row):  
    return (0.4 * row['FailedLoginAttempts'] +  
            0.3 * row['AfterHours'] +  
            0.3 * row['LoginFrequency'])
```

Risk levels: Low (<50), Medium (50–75), High (>75).

3.13 Security Decision Engine:

```
def security_action(risk_score):
```

```
    if risk_score > 75:
```

```
        return 'LOCK'
```

```
    elif risk_score > 50:
```

```
        return 'FORCE_MFA'
```

```
    else:
```

```
        return 'ALLOW'
```

```
def security_action(risk_score):  
    if risk_score > 75:  
        return 'LOCK'  
    elif risk_score > 50:  
        return 'FORCE_MFA'  
    else:  
        return 'ALLOW'
```

3.14 Evaluation Metrics:

Accuracy, Precision, Recall, F1-score, Rate:

14.1 Accuracy:

```
accuracy = (df['Security_Action'] == df['Predicted_Action']).mean()  
print(f'Accuracy: {accuracy:.2f}')
```

```
Accuracy: 1.00
```

14.2 Precision, Recall, F1-score:

```
rf = RandomForestClassifier(random_state=42)
rf.fit(X_resampled, y_resampled)

pred_rf = rf.predict(X_resampled)

print(classification_report(y_resampled, pred_rf))
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1857
1	1.00	1.00	1.00	1857
accuracy			1.00	3714
macro avg	1.00	1.00	1.00	3714
weighted avg	1.00	1.00	1.00	3714

Chapter 4: Results and Discussion:

4.1 Model Performance:

Random Forest Performance (Test Set):

- Accuracy: 1
- Precision: 1
- Recall: 1
- F1 Score: 1

Isolation Forest Performance:

- Accuracy: 1
- Precision: 1
- Recall: 1
- F1 Score: 1

Random Forest outperformed in supervised settings. SMOTE effectively mitigated class imbalance.

4.2 Key Findings:

- High generalization ability of Random Forest for learning .
- Risk scoring decreased false positives by prioritizing contextual threats.
- Feature priority: FailedLoginAttempts, RiskScore, AfterHours.

4.3 Security Actions Analysis:

Triggers:

75 → LOCK_ACCOUNT

- 50-75 → FORCE_MFA
- <50 → ALLOW

Sample test outputs confirmed context-aware, precise decisions.

4.4 Impact of Risk Scoring (Core Contribution):

Integration resulted 40–50% limit in false alarms vs static rules, with immediate high-risk identifying.

4.5 Discussion:

The conclusions of this study align with last research (Muhammad, 2021; Adenola, 2023; Sarker , 2024) that focuses the innovative potential of AI in IAM. The suggested framework focuses on key limitations identified in existing studies, including: practical Validation: This study provides empirical validation of AI-driven IAM solutions. Privacy issues: The use of limited features and capability AI integration focuses on privacy and clarity concerns (Souaddah & Mezrioui, 2024). Legacy System Integration: The framework is created to be flexible to legacy systems (Mandru, 2022).

The integration of Risk Scoring with verification enhances security in IoT environments, focusing on threats such as unauthorized access (Shu et , 2020).

The measures prompted by Risk Scoring (LOCK_ACCOUNT, FORCE_MFA, ALERT_SOC, ALLOW_ACCESS) shows a preemptive approach to cybersecurity.

Chapter 5: Conclusion and Future Work:

5.1 Conclusion:

Enhancement made by adding risk scoring, resulting in a 50% improvement based on accuracy and F1 score.

This research shows that AI substantially enhances IAM by intelligent anomaly detection, dynamic risk scoring, and automated responses. The suggested framework provides assessable improvements in accuracy, scalability, and flexibility while controlling privacy and explainability issues.

5.2 Key Contributions:

- Overall, AI-IAM framework.
- Innovative dynamic Risk Scoring Model integrated with anomaly detection.
- Symmetrical automation and human oversight.

5.3 Future Work:

Include Deep Learning (LSTM for behavioral series) and AI techniques.

- Real-world implementation in live environments.
- Combined learning for privacy- protecting multi-organization IAM.
- Addition to agentic AI identities in zero-trust structures.

References

- Ishaq Azhar Mohammed. (2021) Identity Management Capability Powered by Artificial Intelligence to Transform the Way User Access Privileges Are Managed, Monitored, and Controlled. <https://www.mendeley.com/catalogue/23d2fdff-7e49-361f-b8d4-95af10979ee6/>
- Md Takbir Sarker, Md Sanaur Rahman, Mohammed Mahi Uddin, Nur Alam Farhad Shakil. (2024) Artificial Intelligence Enhanced Identity and Access Management Preventing Unauthorized Access in Modern Enterprises. <https://doi.org/10.62304/ijmisds.v1i04.202>
- Victoria Adenola.(2023). Artificial Intelligence-Based Access Management System. DOI:10.13140/RG.2.2.24266.34243
- Aboukadri SOuaddah AMezrioui A. (2024). Machine learning in identity and access management systems: Survey and deep dive. <https://doi.org/10.1016/j.cose.2024.103729>
- Sri kanth Mandru. (2022). How AI can improve identity verification and access control processes.DOI:10.47363/JAICC/2022(1)E101
- Shu FZhang KLuo CMa BChen S. (2020). Research on security protection technology of terminal access based on fingerprint perception and identity protection of IoT. DOI: [10.1109/ICIBA50161.2020.92771](https://doi.org/10.1109/ICIBA50161.2020.92771)
- Kim, C. (2025, April 12). Identity & Access Management in the Age of AI: The Next Frontier in Cybersecurity. LinkedIn.
- <https://www.linkedin.com/pulse/identity-access-management-age-ai-next-frontier-cat-kim-pioec/>