Reference Report

Sentinel: Intelligent Cyber Threat Defender

By

Mittal Chauhan

Enrolment No: 202126940007

Under the Supervision of

Saurabh Das

A Report Submitted to

Gujarat University

In Partial Fulfilment of the Requirements for

the Degree of **M.Sc. IT in**

Data Management and Visual Insight (5 Years Integrated)

December 2025

Centre for Professional Courses,

Gujarat University, Ahmedabad

CERTIFICATE

This is to certify that research work embodied in this report entitled "**Sentinel: Intelligent Cyber Threat Defender**" was carried out by **Mittal Chauhan (Enrolment No: 202126940007)** at Centre for Professional Course for partial fulfilment of M.Sc. IT degree to be awarded by Gujarat University. This research work has been carried out under my supervision and is to the satisfaction of department.

Date:

Place: CPC, GU

Mr. Saurabh Das
(Guide)
CPC, Gujarat University

Ms. Namita Doshi
Program In-Charge
CPC, Gujarat University

Dr. Paavan Pandit

Director

CPC, Gujarat University

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this Project report and that neither any part of this Project report nor the whole of the Project report has been submitted for a degree to any other University or Institution.

I certify that, to the best of my knowledge, my Project report does not infringe up on any one's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my Project report, published or otherwise, are fully acknowledged in accordance with the standard referencing practices.

I declare that this is a true copy of my Project report, including any final revisions, as approved by my Project report review committee.

Date:

Place: CPC, GU

Mittal Chauhan

Enrolment No: 202126940007

PROJECT REPORT APPROVAL

This is to certify that research work embodied in this Project report entitled "**Sentinel: Intelligent Cyber Threat Defender**" was carried out by **Mittal Chauhan (Enrolment No: 202126940007)** at Centre for Professional Course for partial fulfilment of M.Sc. IT degree in [Data Management and visual insight]to be awarded by Gujarat University.

Date:

Place:

Examiner(s):

_____	_____	_____
(	(	(
)	)	)

ACKNOWLEDGEMENT

We are sincerely thankful to our guide, **Mr. Saurabh Das** for their constant support, stimulating suggestions, and encouragement, which greatly assisted us in successfully completing our project work. Their close supervision over the past few months and helpful insights have been invaluable. Despite their busy schedule, their valuable advice and unwavering support have been an inspiration and a driving force for us. Their experience and knowledge have continuously helped shape our initial ideas into a comprehensive form.

I, hereby, take an opportunity to convey my gratitude for the generous assistance and cooperation, that I received from the **Ms. Namita Doshi** and to all those who helped me directly and indirectly.

We are deeply indebted & thankful to our Department Faculties who helped and rendered their valuable time, knowledge and information and whose suggestion and guidance has enlightened on the subject.

We also thank "**Dr. Paavan Pandit**", Director, CPC, GU for extending all the help and cooperation during our training period.

Finally, I am also indebted to my friends without whose help I would have had a hard time managing everything on my own.

Mittal Chauhan

Enrolment No: 202126940007

ABSTRACT

In the modern digital era, the rapid growth of interconnected systems and cloud-based services has led to a significant increase in cyber threats. Traditional cybersecurity solutions are largely rule-based and reactive, making them ineffective against sophisticated and evolving attacks. To address these challenges, this project presents **"Sentinel: Intelligent Cyber Threat Defender"**, an AI-driven system designed for real-time cyber threat detection and automated response.

The proposed system integrates big data streaming and processing technologies to analyse diverse security data such as network logs, authentication events, email traffic, and CCTV feeds. Machine learning techniques are employed for anomaly detection and threat classification, while computer vision methods enable physical intrusion detection using surveillance data. A reinforcement learning–based mitigation agent is used to recommend appropriate response actions based on real-time risk levels. Additionally, large language models generate concise executive summaries of security incidents, improving decision-making for administrators.

The system is implemented using simulated datasets in a cloud-based environment and demonstrated through an interactive dashboard. The results show improved threat detection efficiency, reduced response time, and enhanced situational awareness, highlighting the effectiveness of intelligent, AI-based cybersecurity solutions.

TABLE OF CONTENTS

CONTENTS

Chapter 1: Introduction

1.1 Introduction

With the increasing dependence on digital platforms, cloud services, and networked systems, organizations are facing a growing number of cyber threats such as unauthorized access, phishing, malware, and data breaches. Traditional cybersecurity systems rely heavily on predefined rules and signatures, making them ineffective against new and evolving attacks. Hence, there is a need for intelligent systems capable of real-time threat detection and response.

Sentinel: Intelligent Cyber Threat Defender is an AI-based cybersecurity system designed to monitor, analyse, and respond to cyber threats in real time. The system integrates data streaming, machine learning, reinforcement learning, and visualization techniques to provide a comprehensive security framework.

1.2 Motivation

The motivation behind this project arises from the limitations of conventional security systems that require manual intervention and lack adaptability. Security analysts often face challenges in handling large volumes of security data and responding quickly to incidents. Automation and intelligence are essential to reduce response time and improve accuracy.

This project aims to demonstrate how artificial intelligence techniques can be used to automate threat detection and mitigation, making cybersecurity systems more proactive and efficient.

1.3 Objectives

The primary objectives of the project are:

- To design a real-time cyber threat detection system
- To apply machine learning for anomaly detection
- To use reinforcement learning for response recommendation
- To generate understandable summaries of security incidents
- To visualize threats and alerts through an interactive dashboard

1.4 Organization of the Report

This report presents the design and implementation of **Sentinel: Intelligent Cyber Threat Defender**, an AI-driven cybersecurity system for real-time threat detection and response. The report begins with an introduction that outlines the problem domain, motivation, and objectives of the project. It then discusses the essential theoretical concepts and technologies that form the foundation of the system, focusing only on those techniques directly used in the implementation.

The proposed system design and architecture are explained next, highlighting the overall workflow and the role of each module within the solution. The implementation environment, including hardware specifications, software tools, and programming platforms, is then described. Subsequent sections detail the dataset used, data preprocessing steps, and model implementation process.

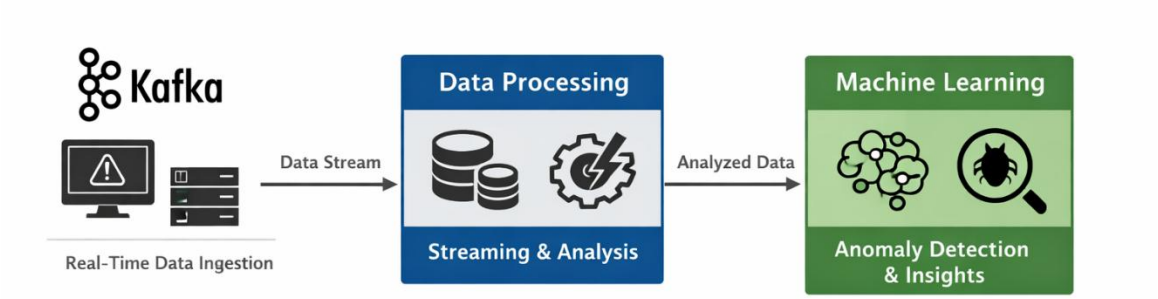
Finally, the report presents the experimental results, evaluation metrics, and dashboard outputs to demonstrate system performance, followed by a conclusion that summarizes the outcomes of the project and discusses possible future enhancements.

Chapter 2: Background Theory

2.1 Streaming and Data Processing Technologies

Real-time cybersecurity systems require continuous monitoring of large-scale data. In this project, Kafka is used conceptually to simulate real-time data ingestion, while Spark-like processing is used to process and analyse streaming data efficiently. These technologies enable scalable and fast handling of security logs and events.

Simple Kafka → Processing → ML flow diagram



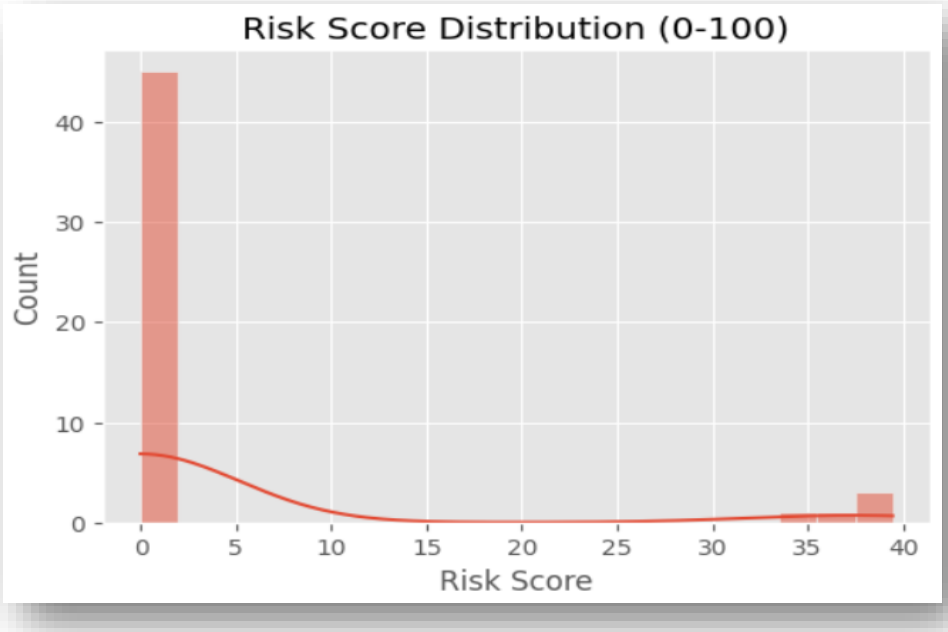
2.2 Machine Learning for Cyber Threat Detection

Machine learning techniques are used to identify anomalies and classify threats in security data. Models are trained using synthetic datasets to learn normal and abnormal behaviour patterns. ML helps in detecting unknown attacks that traditional rule-based systems fail to identify. The models continuously analyse incoming data streams and assign risk scores to suspicious activities. This enables faster detection and proactive response to potential cyber threats.

model training results:

```
===== Part 4: ML Anomaly Detection Demo =====
[INFO] Anomaly detector trained and saved.
[INFO] Incident classifier trained and saved.

[Confusion Matrix]
[[45  0]
 [ 0  5]]

[Classification Report]
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00        45
         1.0       1.00      1.00      1.00         5

    accuracy                           1.00        50
   macro avg       1.00      1.00      1.00        50
weighted avg       1.00      1.00      1.00        50


[Sample Anomalies / Risk Scores]
Index: 1, Risk Score: 33.93, Label: 1.0, Anomaly Score: 0.06
Index: 2, Risk Score: 39.02, Label: 1.0, Anomaly Score: 0.13
Index: 15, Risk Score: 36.02, Label: 1.0, Anomaly Score: 0.09
Index: 17, Risk Score: 37.82, Label: 1.0, Anomaly Score: 0.11
Index: 27, Risk Score: 39.44, Label: 1.0, Anomaly Score: 0.13
```

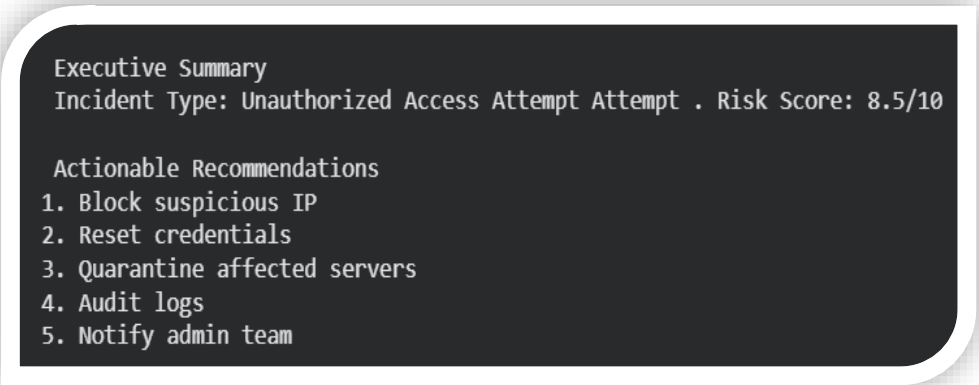
Risk score:



2.3 Reinforcement Learning and LLM Overview

Reinforcement learning is used to recommend appropriate mitigation actions such as alerting administrators or blocking suspicious activities. Large Language Models (LLMs) are used to convert technical security alerts into readable summaries, helping decision-makers understand incidents quickly.

RL action recommendation and summary:



Chapter 3: Proposed Work

3.1 Problem Statement

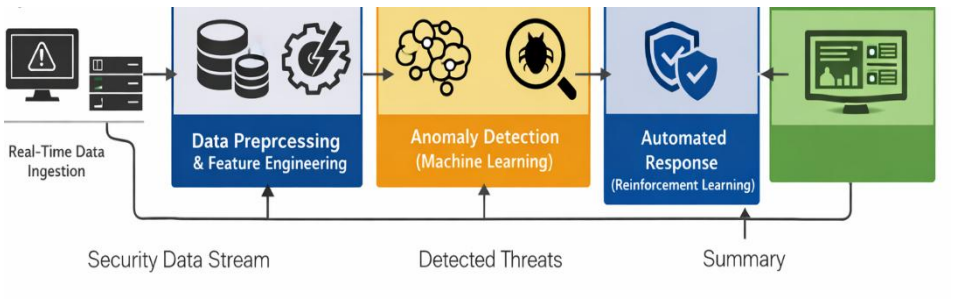
Existing cybersecurity solutions are reactive, fragmented, and heavily dependent on manual analysis. They lack integration between detection, response, and reporting mechanisms, leading to delayed mitigation and higher risk exposure.

3.2 System Architecture

The proposed system consists of multiple modules including data ingestion, preprocessing, machine learning detection, reinforcement learning-based response, LLM summarization, and dashboard visualization. All modules work together to provide end-to-end cyber threat defence.

System Architecture Diagram:

(Data → ML → RL → LLM → Dashboard)



3.3 Module Description

The system is divided into the following modules:

- Data ingestion and simulation
- Data preprocessing and feature extraction
- Anomaly detection using ML
- Response decision using RL
- Incident summarization using LLM
- Dashboard visualization

3.4 Expected Outcome

The expected outcome of the project is an intelligent system capable of detecting cyber threats in real time, recommending mitigation actions, and presenting results in an easy-to-understand format for administrators. The system supports automated response recommendations such as alert generation, threat prioritization, and basic mitigation actions, thereby reducing manual intervention. From a business perspective, the solution helps minimize potential financial and operational losses caused by cyber incidents by reducing response time, improving decision-making efficiency, and strengthening the overall security posture of the organization.

Chapter 4: Implementation Environment

4.1 System Specifications

The project is implemented using a cloud-based environment with the following specifications:

- Platform: Google Colab
- Operating System: Linux (Cloud)
- Processor: Virtual CPU
- RAM: Sufficient for ML workloads

4.2 Software and Tools Used

The project is implemented using **Python** for data processing, machine learning, and visualization. Libraries **NumPy** and **Pandas** are used for handling and structuring security event data. **Scikit-learn** is employed for anomaly detection and classification models. **deep learning and reinforcement learning libraries**, including PyTorch-based frameworks and **Stable-Baselines3**, to support intelligent threat mitigation strategies. **Apache Kafka** and **Apache Spark** are simulated to demonstrate real-time data ingestion and streaming-based processing.

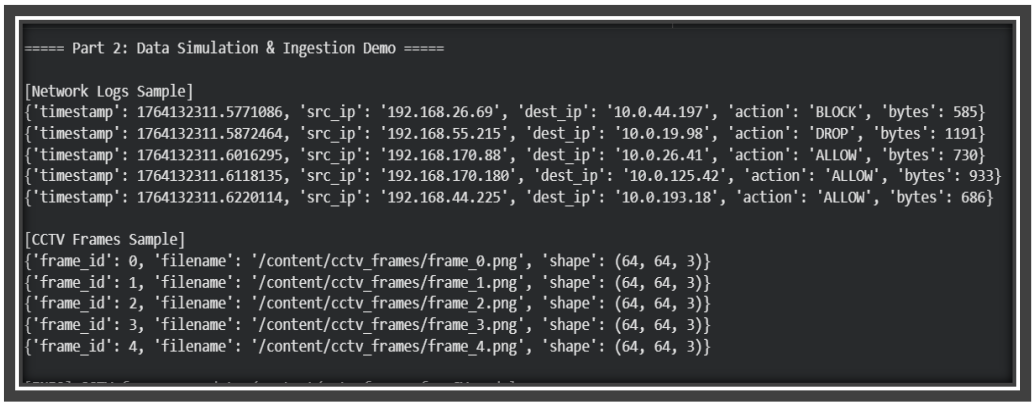
For computer vision–based intrusion detection, **OpenCV** is used along with lightweight pretrained models such as **YOLOv8** to analyse synthetic CCTV frames. **Transformer-based models** are used to generate incident summaries, representing Large Language Model (LLM) integration. An interactive dashboard is developed using **Streamlit**, and **pyngrok** is used to expose the dashboard during demonstration. The complete system is executed in a **Google Colab environment**, ensuring reproducibility and ease of experimentation using simulated data sources.

Chapter 5: Dataset and Implementation

5.1 Dataset Description

Synthetic datasets are generated to simulate network traffic, authentication events, and security logs. These datasets are used to train and test the models without relying on real sensitive data.

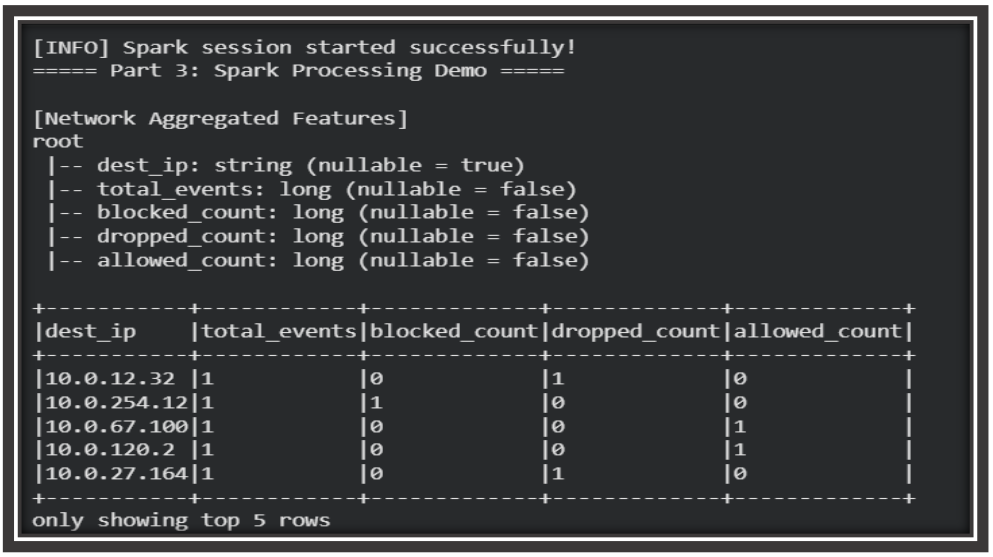
Data simulation and Ingestion:



5.2 Data Preprocessing

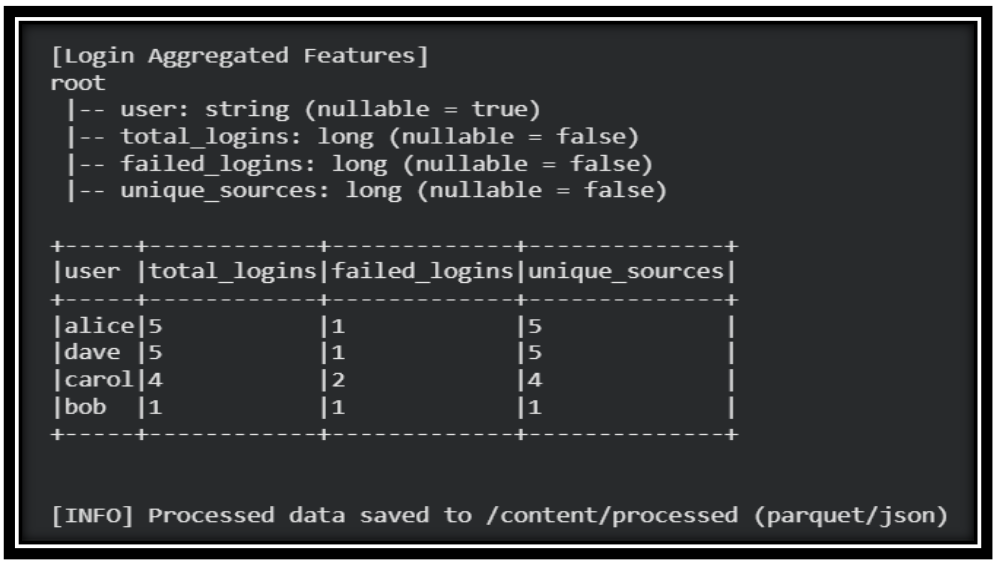
Preprocessing in this project is performed using **Apache Spark–based processing** on **synthetically generated security data**. Since the data is generated in a controlled format, extensive cleaning is not required. Spark is used to structure incoming events, standardize feature formats, and organize data into micro-batches suitable for analysis. Basic transformations such as timestamp alignment and feature selection are applied to prepare the data for machine learning models. This approach ensures efficient handling of simulated real-time data while maintaining consistency across processing stages.

Pre-processed dataset (network features):



```
[INFO] Spark session started successfully!
===== Part 3: Spark Processing Demo =====

[Network Aggregated Features]
root
 |-- dest_ip: string (nullable = true)
 |-- total_events: long (nullable = false)
 |-- blocked_count: long (nullable = false)
 |-- dropped_count: long (nullable = false)
 |-- allowed_count: long (nullable = false)

+-----------+------------+-------------+-------------+-------------+
|dest_ip    |total_events|blocked_count|dropped_count|allowed_count|
+-----------+------------+-------------+-------------+-------------+
|10.0.12.32 |1           |0            |1            |0            |
|10.0.254.12|1           |1            |0            |0            |
|10.0.67.100|1           |0            |0            |1            |
|10.0.120.2 |1           |0            |0            |1            |
|10.0.27.164|1           |0            |1            |0            |
+-----------+------------+-------------+-------------+-------------+
only showing top 5 rows
```

Pre-processed dataset (login features):



```
[Login Aggregated Features]
root
 |-- user: string (nullable = true)
 |-- total_logins: long (nullable = false)
 |-- failed_logins: long (nullable = false)
 |-- unique_sources: long (nullable = false)

+-----+------------+-------------+--------------+
|user |total_logins|failed_logins|unique_sources|
+-----+------------+-------------+--------------+
|alice|5           |1            |5             |
|dave |5           |1            |5             |
|carol|4           |2            |4             |
|bob  |1           |1            |1             |
+-----+------------+-------------+--------------+


[INFO] Processed data saved to /content/processed (parquet/json)
```

5.3 Model Training

Machine learning models are trained using the processed synthetic security data generated within the system. The training process focuses on learning normal system behaviour patterns so that deviations can be identified as potential threats. The models are trained using features extracted from streaming data processed through the Spark-based pipeline. This training enables the system to detect anomalies and suspicious activities in real time without relying on predefined rules.

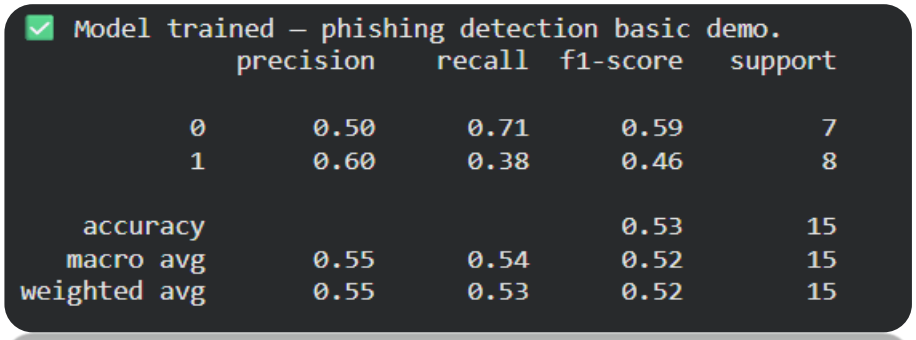
5.4 Model Testing

The trained models are tested using separate synthetic test data generated during system execution. Testing is performed to verify the model's ability to correctly identify anomalous events and classify potential threats. The test results help evaluate the effectiveness of the trained models in detecting abnormal behaviour and ensure reliable performance before integrating the models into the real-time threat detection workflow.

Chapter 6: Results and Evaluation

6.1 Evaluation Metrics

The system performance is evaluated using metrics such as accuracy, precision, recall, and detection rate.

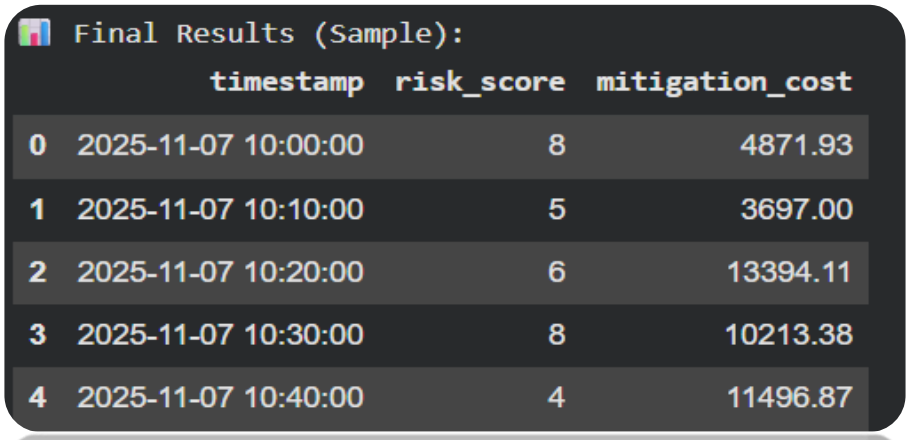


✅ Model trained — phishing detection basic demo.

	precision	recall	f1-score	support
0	0.50	0.71	0.59	7
1	0.60	0.38	0.46	8
accuracy			0.53	15
macro avg	0.55	0.54	0.52	15
weighted avg	0.55	0.53	0.52	15

6.2 Results Analysis

The results show that the proposed system can successfully detect anomalies and recommend appropriate mitigation actions with improved efficiency.



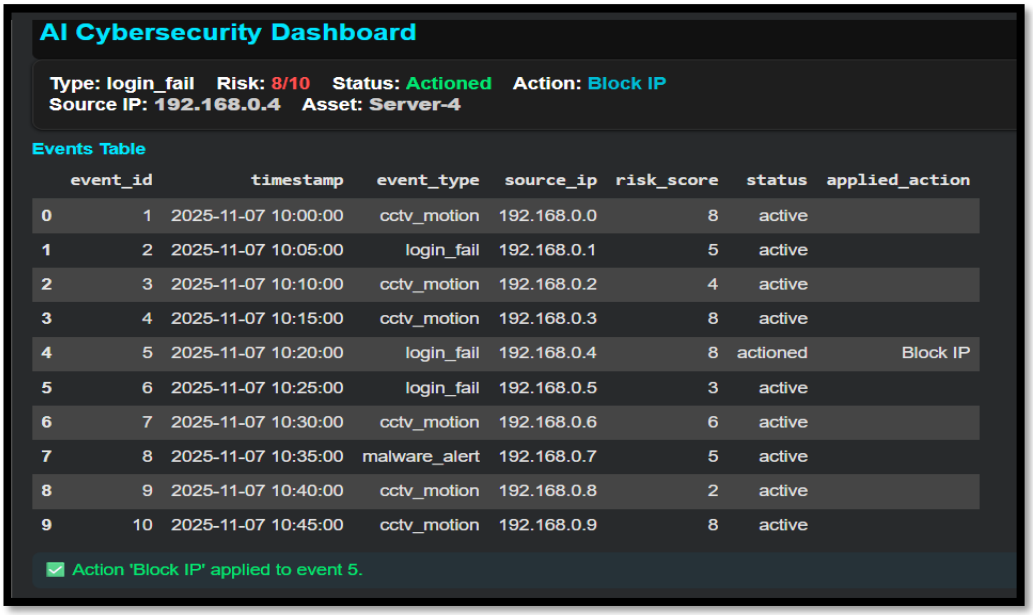
📊 Final Results (Sample):

	timestamp	risk_score	mitigation_cost
0	2025-11-07 10:00:00	8	4871.93
1	2025-11-07 10:10:00	5	3697.00
2	2025-11-07 10:20:00	6	13394.11
3	2025-11-07 10:30:00	8	10213.38
4	2025-11-07 10:40:00	4	11496.87

6.3 Dashboard Output

An interactive dashboard is developed to visualize threats, alerts, and recommended actions in real time.

Dashboard UI output:



	event_id	timestamp	event_type	source_ip	risk_score	status	applied_action
0	1	2025-11-07 10:00:00	cctv_motion	192.168.0.0	8	active	
1	2	2025-11-07 10:05:00	login_fail	192.168.0.1	5	active	
2	3	2025-11-07 10:10:00	cctv_motion	192.168.0.2	4	active	
3	4	2025-11-07 10:15:00	cctv_motion	192.168.0.3	8	active	
4	5	2025-11-07 10:20:00	login_fail	192.168.0.4	8	actioned	Block IP
5	6	2025-11-07 10:25:00	login_fail	192.168.0.5	3	active	
6	7	2025-11-07 10:30:00	cctv_motion	192.168.0.6	6	active	
7	8	2025-11-07 10:35:00	malware_alert	192.168.0.7	5	active	
8	9	2025-11-07 10:40:00	cctv_motion	192.168.0.8	2	active	
9	10	2025-11-07 10:45:00	cctv_motion	192.168.0.9	8	active	

Chapter 7: Conclusion

7.1 Conclusion

This project successfully demonstrates the application of artificial intelligence techniques in cybersecurity through the development of **Sentinel: Intelligent Cyber Threat Defender**. The system integrates real-time data ingestion, anomaly detection, automated mitigation recommendations, and interactive visualization into a unified framework. By combining machine learning, reinforcement learning, and dashboard-based monitoring, the solution reduces manual intervention and enhances the efficiency of threat detection and response. From a business perspective, the system supports faster decision-making, reduces potential financial losses caused by delayed responses, and improves the overall security posture of an organization.

.

7.2 Future Scope

Future enhancements of the system may include **deployment in a real-time production environment** using cloud platforms such as AWS or Google Cloud. Integration with **live network traffic, real email systems, and CCTV camera feeds** can further improve system accuracy and applicability. Advanced deep learning models can be incorporated to enhance detection capabilities, while real-time video analytics can strengthen physical security monitoring. From a business standpoint, the system can be extended to provide cost estimation of cyber incidents, compliance reporting, and organization-wide risk assessment, making it a complete enterprise-grade cybersecurity solution.

References

1. Apache Software Foundation, *Apache Kafka Documentation*, Available: https://kafka.apache.org/documentation/

2. Apache Software Foundation, *Apache Spark Structured Streaming Guide*, Available: https://spark.apache.org/docs/latest/structured-streaming-programming-guide.html

3. Scikit-learn Developers, *Scikit-learn: Machine Learning in Python*, Available: https://scikit-learn.org/stable/

4. Goodfellow, I., Bengio, Y., and Courville, A., *Deep Learning*, MIT Press, 2016.

5. Sutton, R. S., and Barto, A. G., *Reinforcement Learning: An Introduction*, MIT Press, 2018.

6. Brown, T. et al., "Language Models are Few-Shot Learners," *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

7. Streamlit Inc., *Streamlit Documentation*, Available: https://docs.streamlit.io/