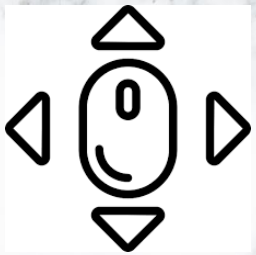
Behavioral Biometrics for Enhanced Authentication Systems

24-25J-073



❑ Our Team

SUPERVISOR

Dr. Harinda Fernando

Assistant Professor

Faculty of Computing

CO-SUPERVISOR

Mr. Tharaniyawarma

Assistant Lecturer

Faculty of Computing

Madhubhashana H. N. D

IT21391668

CYBER SECURITY

Edirisinghe E.M.N

IT21340864

CYBER SECURITY

Anupama K.G. A

IT21345678

CYBER SECURITY

Rajapaksha R. P. K. D

IT21336072

CYBER SECURITY

❑ Introduction to the Project

Goal

- Develop an advanced authentication system using multiple behavioral biometrics.

Focus

- Integrate gait, mouse dynamics, keystroke dynamics, and voice authentication.

Significance

- Enhance security and reliability compared to single-modal systems.

Innovation

- Utilize a hybrid CNN-RNN model for effective data integration and analysis.

❑ Research Gap

Features/ Technologies	Scalability	Use of Online Datasets	Hybrid Model (CNN + RNN)	Specialized Hardware Required
Project X				
Project Y				
Project Z				
OUR PROJECT				

❑ Research Problem

Challenges with Traditional Biometric Authentication

- ➢ Vulnerable to spoofing and privacy concerns.
- ➢ Requires physical contact or proximity.

Limitations of Existing Gait Analysis Methods

- ➢ Often lack robustness and accuracy under diverse conditions.
- ➢ Need for improved feature extraction and modeling techniques.

Need for Robust and Accurate Behavioral Biometric Systems

- ➢ Behavioral biometrics offer non-intrusive and unique patterns.
- ➢ Potential to significantly enhance user authentication security.

❑ Research Objectives

✓ Primary Objective

Develop an integrated user authentication system leveraging behavioral biometrics.

Enhance user authentication accuracy and security by combining multiple behavioral biometric modalities.

❑ Research Objectives

✓ Secondary Objective

Gait Analysis

- Develop a hybrid CNN-RNN model.
- Improve feature extraction techniques.

Mouse Dynamics

- Implement an RNN model.

Keystroke Dynamics

- Develop a hybrid CNN-RNN model.

Voice Biometric Authentication

- Develop a combined CNN-RNN model.

Data Security

- Ensure data security and privacy with advanced encryption and storage techniques.

Model Performance

- Validate through extensive testing and real-world data collection.

❑ Research Question

How can a hybrid CNN-RNN model be optimized to integrate gait, mouse dynamics, keystroke dynamics, and voice biometrics for better authentication?

❑ Research Solution

Integration Strategy

- Design multi-input model.

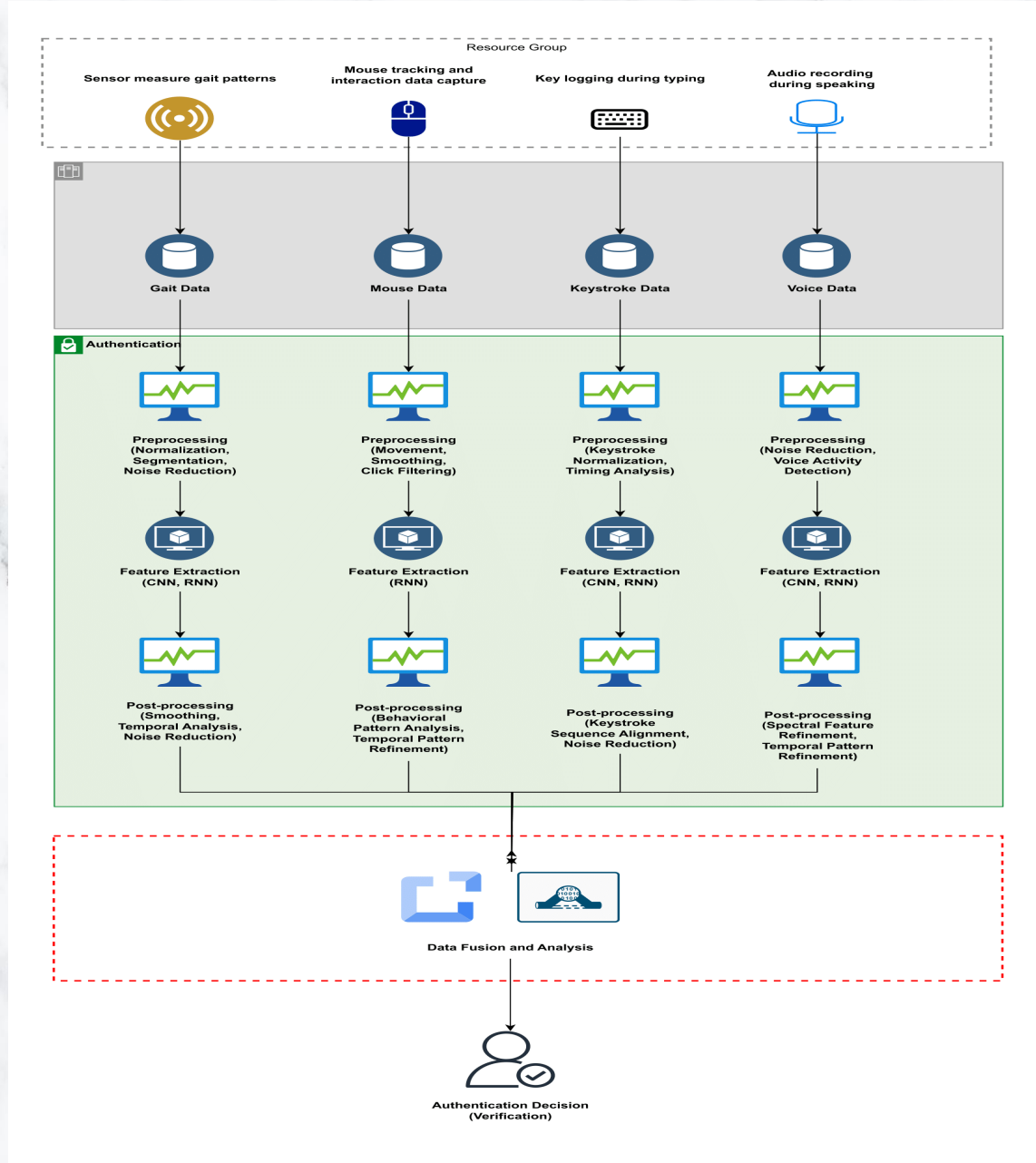
Evaluation

- Metrics: Accuracy, precision, recall, F1-score.
- Perform cross-validation.
- Test with real-world data.

Continuous Improvement

- Refine model based on feedback and new data.

❑ System Diagram



IT21391668 | H.N.D. MADHUBHASHANA

BSc (Hons) in Information Technology Specialising in Cyber Security

Introduction to Gait Component

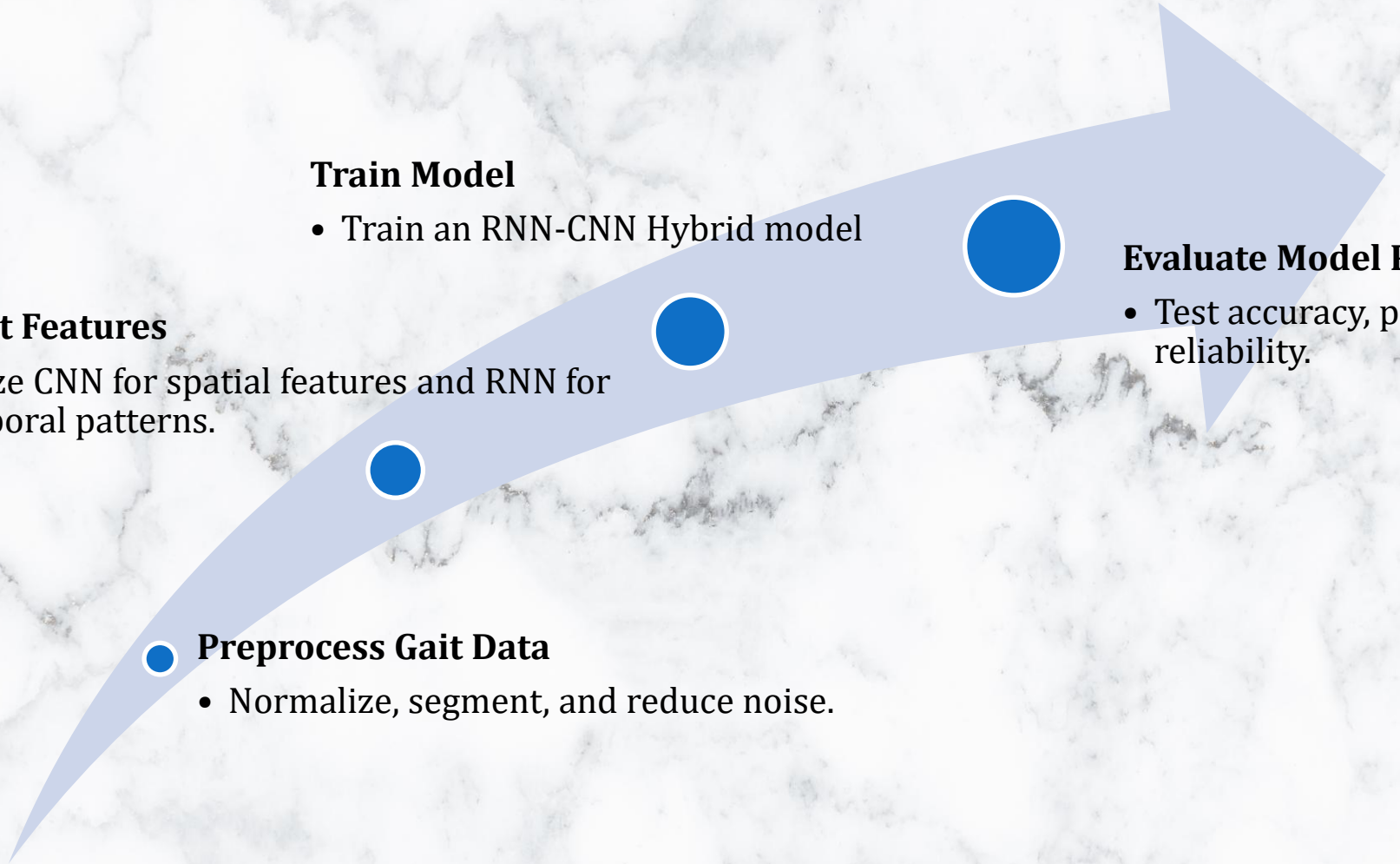
Goals

- **Analyze Gait Patterns:** Understand the key features and characteristics of gait data.
- **Enhance Model Integration:** Improve how gait data is processed and fused with other biometric modalities

Focus

- **Objective:** Integrate gait data into a hybrid CNN-RNN model for enhanced multi-modal authentication.
- **Relevance:** Gait patterns are unique and can provide an additional layer of security in authentication systems.

Research Sub-Objectives



Preprocess Gait Data

- Normalize, segment, and reduce noise.

Extract Features

- Utilize CNN for spatial features and RNN for temporal patterns.

Train Model

- Train an RNN-CNN Hybrid model

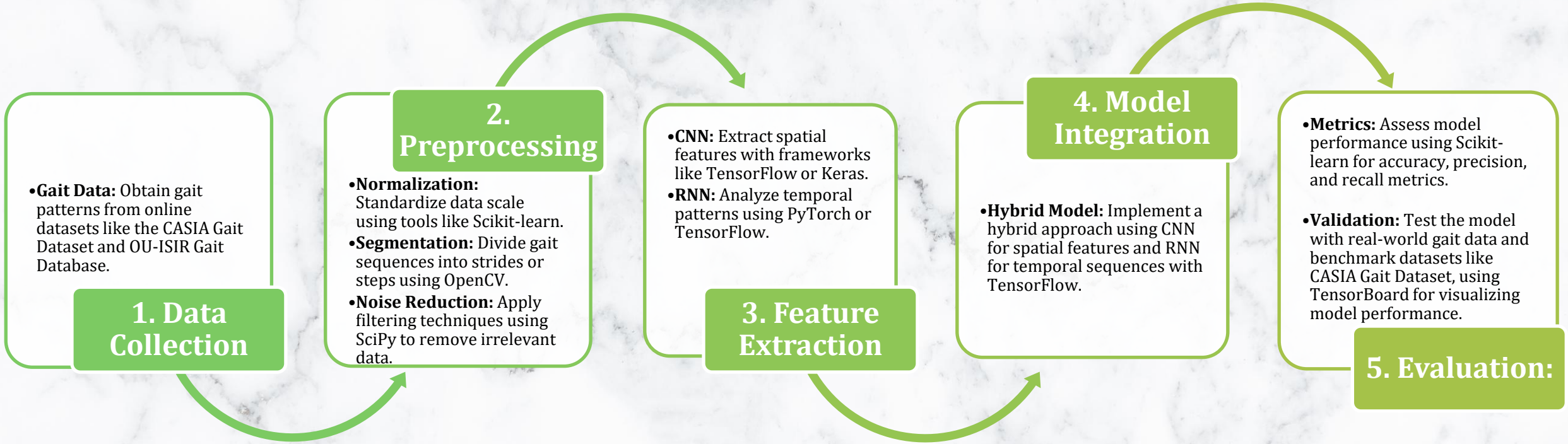
Evaluate Model Performance

- Test accuracy, precision, and real-world reliability.

Research Question

How can gait data be effectively integrated into a hybrid CNN-RNN model to enhance multi-modal authentication systems?

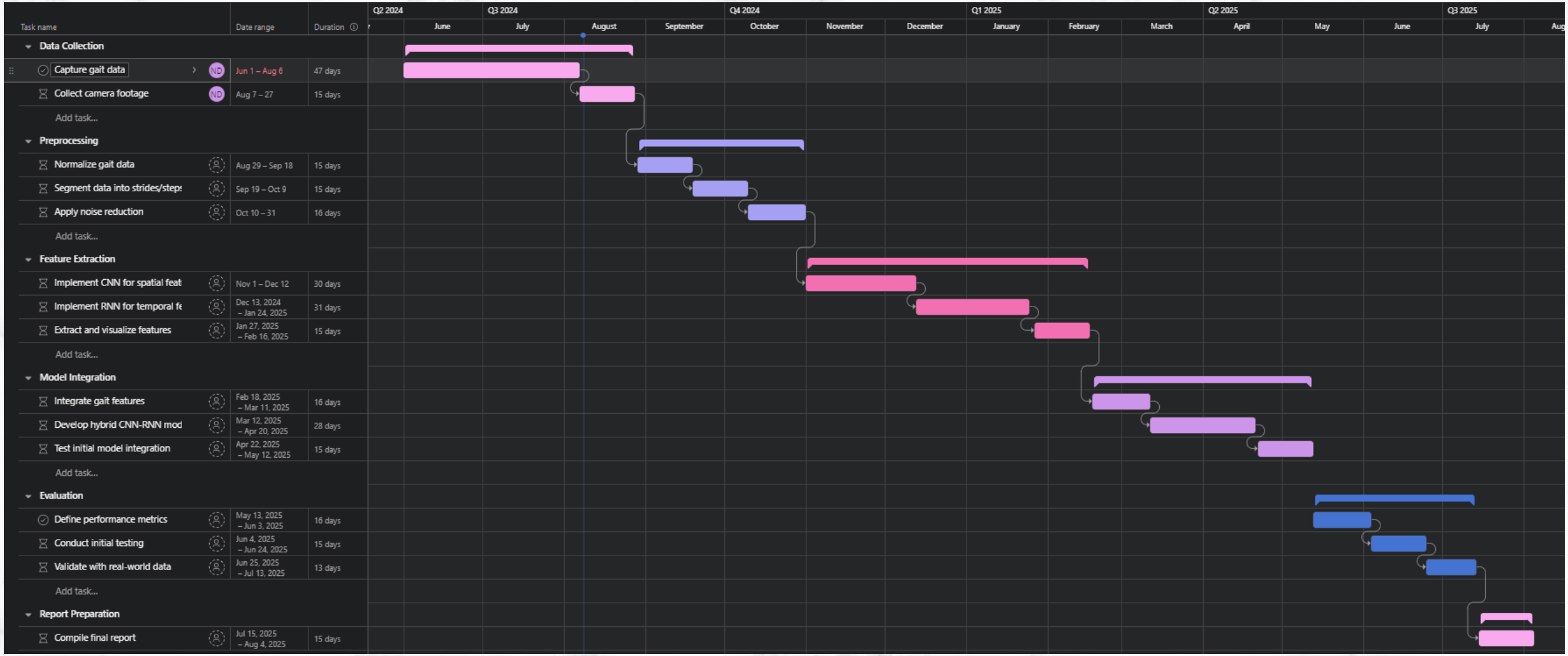
Methodology



Evidence for Completion

Data Collection	Preprocessing	Feature Extraction	Model Integration	Evaluation
• Downloaded Datasets	• Documentation of methods. • Example preprocessed data.	• Code and visualizations of features.	• Integration documentation. • Results of fused features.	• Performance metrics. Validation results.

Project Timeline: Gantt Chart



REFERENCES

G. Giorgi, F. Martinelli, A. Saracino, and M. Sheikhalishahi, "Walking Through the Deep: Gait Analysis for User Authentication Through Deep Learning," *Inria*, [Online]. Available: https://inria.hal.science/hal-02023725/document. Accessed: Aug. 4, 2024.

I. Stylios, "Behavioral Biometrics for Continuous Authentication: Security and Privacy Issues," *ResearchGate*, Jan. 2023. [Online]. Available: https://www.researchgate.net/publication/369142299_Behavioral_Biometrics_for_Continuous_Authentication_Security_and_Privacy_Issues. Accessed: Aug. 4, 2024.

IT21340864 | E.M.N. EDIRISINGHE

BSc (Hons) in Information Technology Specialising in Cyber Security

Introduction to Keystroke Dynamics Component

Goals

- **Analyze Typing Patterns:** Understand key features of typing behavior, including timing between key presses and pressure applied.
- **Enhance Model Integration:** Improve how keystroke dynamics are processed and combined with other biometric modalities to strengthen authentication.

Focus

- **Objective**: Integrate keystroke dynamics into a hybrid RNN model to enhance multi-modal authentication systems.
- **Relevance**: Typing patterns are unique to each individual and can provide a valuable layer of security in authentication systems, making them harder to replicate and bypass.

Research Sub-Objectives

Develop a Keystroke Dynamics Dataset:Create a comprehensive dataset of keystroke dynamics for training and evaluation purposes.

Extract and Analyze Keystroke Features:Identify and extract key features from typing patterns, such as dwell time and flight time, to distinguish between users.

Train and Validate RNN Models:Develop and validate Recurrent Neural Network (RNN) models to accurately capture the temporal dependencies in typing behavior.

Enhance Multi-Modal Integration:Improve the integration of keystroke dynamics with other biometric modalities, such as gait, mouse dynamics, and voice, within a hybrid CNN-RNN model.

Evaluate Authentication Performance:Assess the performance of the keystroke dynamics component in terms of accuracy, precision, recall, and overall robustness in authentication scenarios.

Research Question

How can a hybrid model combining spatial pattern recognition and sequential behaviour analysis improve the accuracy and robustness of keystroke dynamics for user authentication?

Methodology

1

Data Collection

Use keyloggers and online datasets to gather keystroke data.

2

Data Preprocessing

Clean and normalize the collected data to ensure consistency and quality.

3

Feature Extraction

Use CNNs to extract spatial features from the keystroke data.

Extract key features from the keystroke data using RNN models

4

Sequence Analysis

Analyze the sequences of keystroke events using advanced RNN architectures, including Bidirectional RNNs.

5

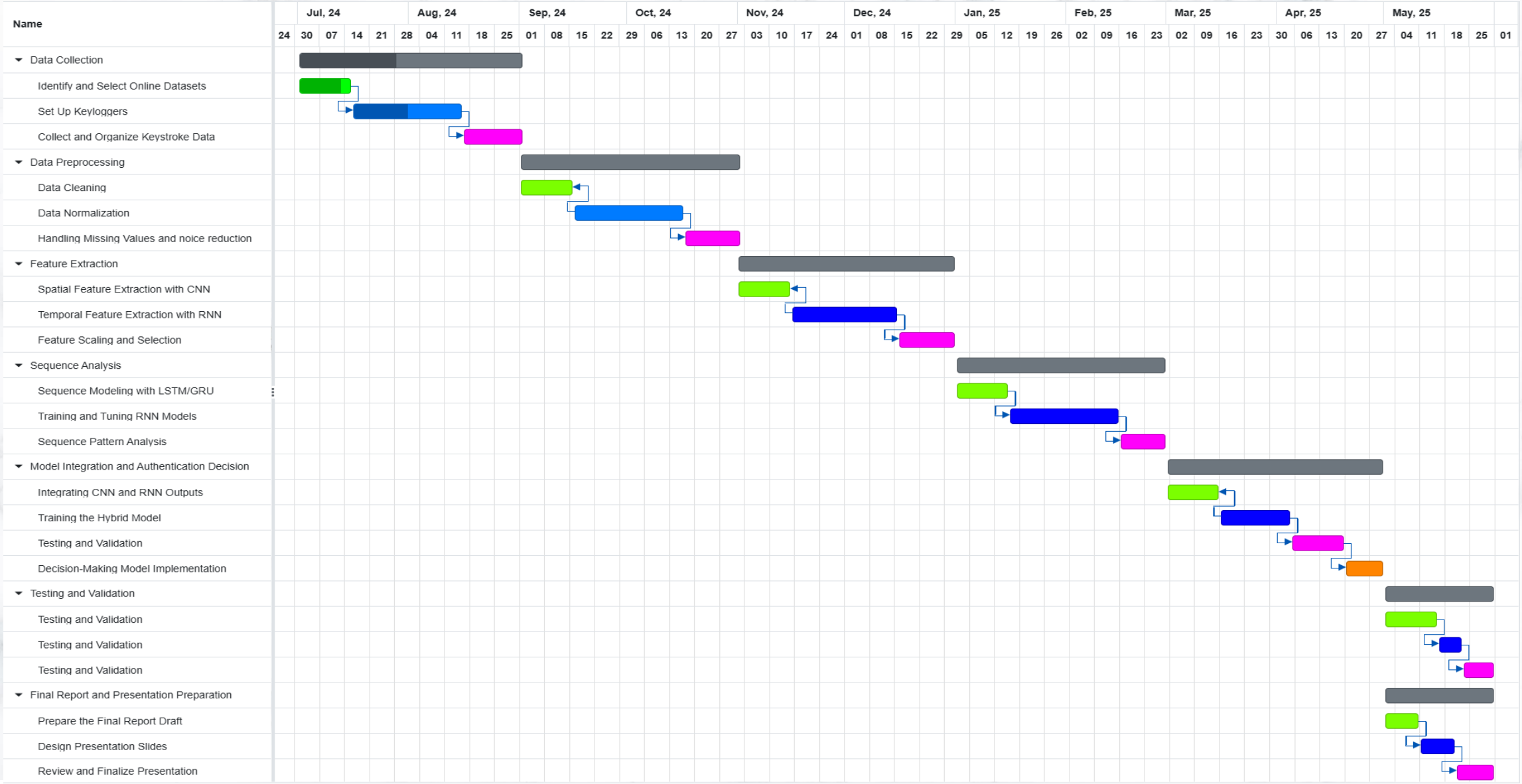
Model Integration and Authentication Decision

Integrate the outputs from sequence analysis into a training model, and apply decision-making techniques such as thresholding, SVM, or Random Forest.

Evidence for Completion

Data Collection	Data Preprocessing	Feature Extraction	Sequence Analysis	Model Integration and Authentication Decision
• Online Datasets • Logs from keyloggers	• Scripts used for data cleaning and normalization. • Raw keystroke data and the cleaned, pre-processed data	• Code and visualizations of features. • Performance Metrics	• Integration documentation. • Results such as accuracy, loss curves, and performance metrics from sequence analysis models	• Performance metrics. Validation results. • Scripts used for training the final model. • Summarize the final authentication process

Project Timeline: Gantt Chart



REFERENCES

- Aditya Arsh, Nirmalya Kar , and Subhrajyoti Deb , "Multiple Approaches Towards Authentication Using Keystroke Dynamics," 2024.

- Rashik Shadman, Ahmed Anu Wahab, Michael Manno, Matthew Lukaszewski, Daqing Hou, Faraz Hussain, “Keystroke Dynamics: Concepts, Techniques, and Applications” ,2024.

- Yutong Shi, Xiujuan Wang, Kangfeng Zheng, “User authentication method based on keystroke dynamics and mouse dynamics using HDA”, 2022.

IT21345678 |ANUPAMA K. G. A

BSc (Hons) in Information Technology Specialising in Cyber Security

Introduction to Mouse Dynamics Component

Goals

- **Analyze Mouse Movement Patterns:** Understand key features Mouse speed behaviors, including timing between pressing left and right buttons, scroll speed, direction changes
- **Enhance Model Integration:** Improve how mouse dynamics are processed and combined with other biometric modalities to strengthen authentication.

Focus

- **Objective**: Integrate mouse dynamics into a RNN model to enhance multi-modal authentication systems.
- **Relevance**: Mouse movement patterns, Scrolling pattens , Buttons pressing speeds are unique to each individual and can provide a valuable layer of security in authentication systems, making them harder to replicate and bypass.

Research Sub-Objectives

Collect a Mouse Dynamics Dataset: Collect a comprehensive dataset of Mouse dynamics for training and evaluation purposes.

Extract and Analyze Mouse dynamic Features: Identify and extract mouse dynamic features like mouse movement patten and speed , scrolling speeds , clicking pattens, direction changes .

Train and Validate RNN Models: Develop and validate Recurrent Neural Network (RNN) models to accurately capture the temporal dependencies in typing behavior.

Enhance Multi-Modal Integration: Improve the integration of mouse dynamics with other biometric modalities, such as gait, keystroke dynamics, and voice, within a hybrid CNN-RNN model.

Evaluate Authentication Performance: Assess the performance of the mouse dynamics component in terms of accuracy, precision, recall, and overall robustness in authentication scenarios.

Research Question

How does the implementation of Recurrent Neural Networks (RNNs) in mouse movement analysis enhance the capture and interpretation of nuanced temporal patterns, thereby improving the accuracy of user authentication systems?

Methodology

1

Data Collection

Use multiple tools and online datasets to gather mouse movement data.

2

Data Preprocessing

Clean and normalize the collected data to ensure consistency and quality.

3

Feature Extraction

Spatial Features

1. Trajectory Analysis
2. Heatmaps

Temporal Features

1. Movement Speed
2. Click Intervals

4

Sequence Analysis

Analyze the sequences of mouse movement events using advanced RNN architectures

5

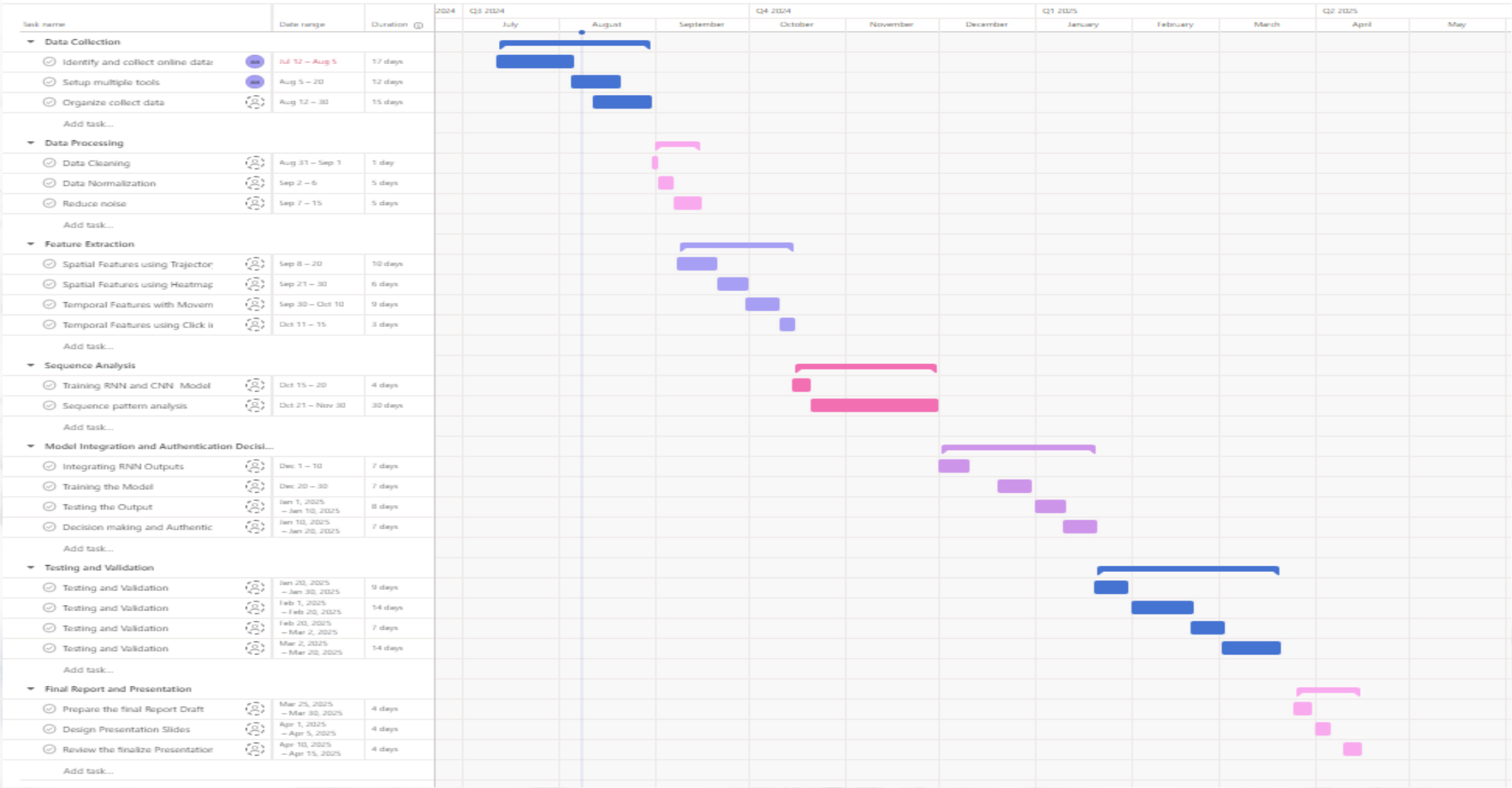
Model Integration and Authentication Decision

Integrate the outputs from sequence analysis into a training model, and apply decision-making techniques such as thresholding, SVM, or Random Forest.

Evidence for Completion

Data Collection	Data Preprocessing	Feature Extraction	Sequence Analysis	Model Integration and Authentication Decision
• Online Datasets • Multiple tools Ex : Python tools , Browser extensions.	• Scripts used for data cleaning and normalization. • Raw mouse movement data and the cleaned, pre-processed data	• Code and visualizations of features. • Performance Metrics	• Integration documentation. • Results such as accuracy, loss curves, and performance metrics from sequence analysis models	• Performance metrics. Validation results. • Scripts used for training the final model. • Summarize the final authentication process

Project Timeline: Gantt Chart



REFERENCES

- Ahmed, A. A. E., & Traore, I. (2007). A new biometric technology based on mouse dynamics. *IEEE Transactions on Dependable and Secure Computing*, 4(3), 165-179.
- Zach Jorgensen, & Ting Yu .(March 22-24, 2011) On mouse dynamics as a behavioral biometric for authentication
- P. Bours and C. J. Fullu. A login system using mouse dynamics. In Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pages 1072–1077, 2009.

IT21336072 | R.P.K.D RAJAPAKSHA

BSc (Hons) in Information Technology Specializing in Cyber Security

Introduction to Voice Biometrics Component

Goals

- **Analyze Voice Patterns:** Understand the key features and characteristics of voice data.
- **Enhance Model Integration:** Improve how voice data is processed and fused with other biometric modalities.

Focus

- **Objective**: Integrate voice data into a hybrid CNN-RNN model for enhanced multi-modal authentication.
- **Relevance**: Voice patterns are unique and provide an additional layer of security in authentication systems

Research Sub-Objectives

Data Preprocessing

- Preprocess the collected voice data using tools Like Librosa or PyDub to enchance quality and consistency

Extract Features

- Preprocessed voice data using Convolutional Neural Networks(CNN) for spectrogram analysis

Sequence Analysis

- Analyze sequence of extracted features using Recurrent Neural Networks (RNN)

Hybrid Model Development

- Develop a hybrid CNN-RNN model that processes combined features from multiple biometric components for integrated authentication decision making.

Research Question

How can voice biometrics, when integrated with other behavioral biometric components, enhance the accuracy and security of an authentication system?

Methodology

Data Collection

Collect a dataset of a minimum 3000 user voice samples

1

Data Preprocessing

Utilize Librosa or PyDub for noise reduction, normalization, and segmentation of voice data.

2

Feature Extraction

Apply CNNs to the spectrograms to extract key features that represent the unique characteristics of each speaker's voice

3

Sequence Analysis

Use RNN architectures , such as LSTM or GRU, to analyze the sequential nature of the extracted voice features.

4

Model Integration and Authentication Decision

Develop a hybrid CNN-RNN model to process the combined features and make integrated authentication decisions.

Evidence for Completion

Data Collection

- Collected voice samples from at least 3000 Users
- Maintained a comprehensive log of all collected data.

Preprocessing

- Successfully preprocessed all collected voice data using Librosa/PyDub.
- Generated spectrograms for all audio samples.

Feature Extraction

- Trained a CNN model on the spectrograms to extract voice features.

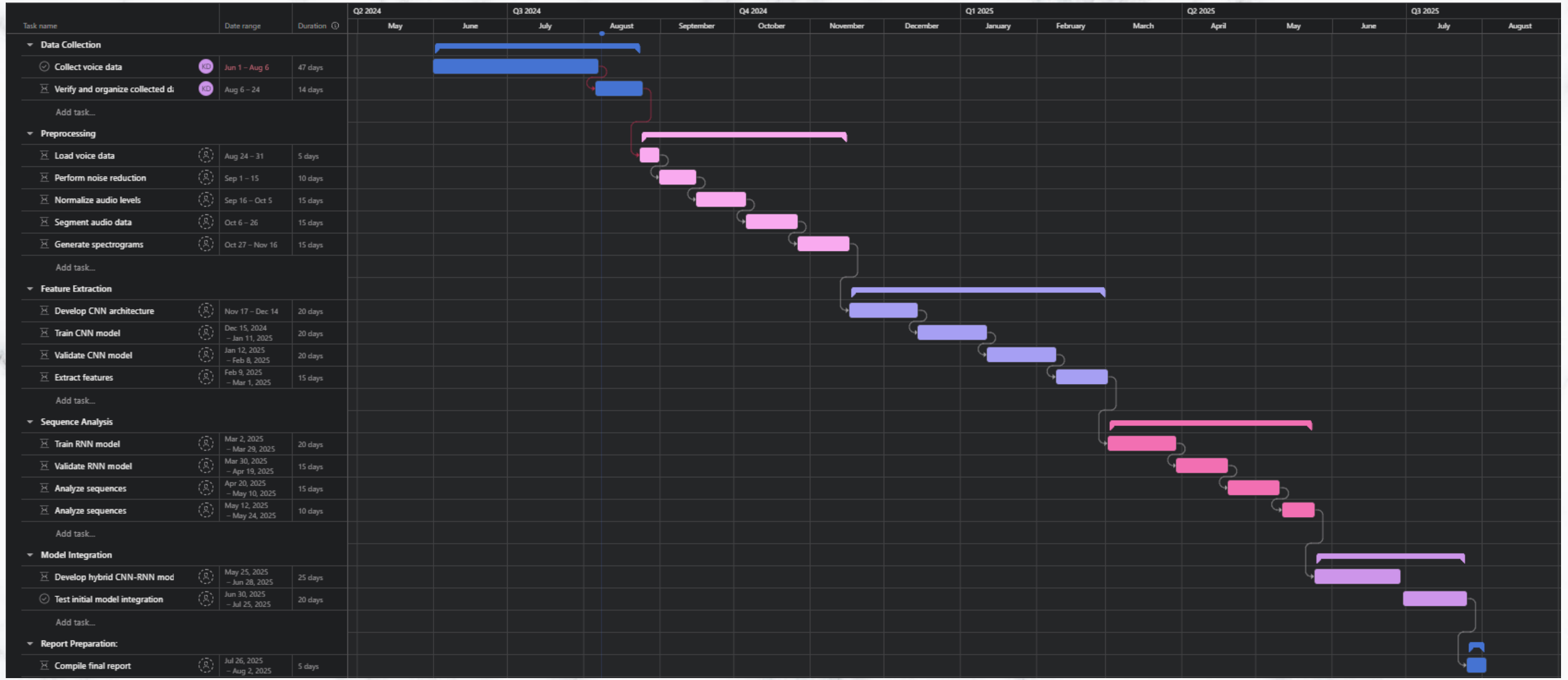
Sequence Analysis

- Trained an RNN model (LSTM/GRU) to analyze the sequence of extracted features.
- Provided detailed analysis and performance metrics of RNN model

Model Integration and Authentication Decision

- Developed and tested the hybrid CNN-RNN model.
- Documented the integrated authentication decision process and evaluated its performance.

Project Timeline: Gantt Chart



REFERENCES

- Kinnunen, T., & Li, H. (2010). "An overview of text-independent speaker recognition: From features to supervectors". Speech Communication, 52(1), 12-40.
- Graves, A., et al. (2013). "Speech recognition with deep recurrent neural networks". IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).
- Sainath, T. N., et al. (2015). "Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks". IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).