



Sri Lanka Institute of Information Technology

## Mouse Dynamics Authentication Systems

**Project ID – 24-25J-073**

### **Individual Project Proposal Report**

Submitted by:


<b>Student Registration Number</b>	<b>Student Name</b>
IT21345678	Anupama K.G A

**Department of Computer System Engineering**


Date of submission

## Declaration

I declare that this is my own work, and this proposal does not incorporate without acknowledgement of any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
Anupama K.G A	IT21345678	

The above candidate is carrying out research for the undergraduate Dissertation under my supervision.

  
.....  
Signature of the supervisor

22/8/24  
.....  
Date

## **Abstract**

Biometric systems such as fingerprint and facial recognition in the current world of cyber threats have always been challenged by issues to do with privacy, precision, and versatility. This paper provides a new way of authenticating users, which involves using the mouse movements as the mode of identifying the user. In this technique we take advantages of features of moving mouse such as speed, rhythm and its moving trajectory. The goal of this project is to design an advanced level of authentication that would be sufficiently secure and at the same time privacy preserving.

To analyze and understand complex patterns of rodent movements, the study uses state of art machine learning algorithms such as CNN, RNN, LSTM and so on. Hence, the goal of the system is to identify such patterns accurately so as to either recognize a user or reject a likely imposter without much compromise on the variability in user behavior.

However, more than enhancing the levels of authentication accuracy, the project is concerned with the other quality attributes or non-functional requirements, which include the system incorporates highly secure encryption, the protection of users and their information, and it addresses most regulations. It is also designed for possible scaling up of the number of users to accommodate increased numbers as well as designed to be usable by more people as possible so as to popularize it. Therefore, this research provides a solution to the security, privacy, and usability issues while implementing biometric authentication in organizations, making a way for more effective biometric authentication innovations.

## Table of Contents

1. Introduction .....	5
2. Background and Literature Review .....	6
Overview of Keystroke Dynamics .....	6
Research Gap .....	7
Research Problem .....	8
3. Objectives .....	10
4. Methodology .....	11
5. Requirements .....	15
User Requirements .....	15
Functional Requirements .....	16
Non-Functional Requirements .....	16
6. Feasability Study .....	18
Technical Feasibility Assessment .....	18
Economic Feasibility .....	19
Schedule Feasibility .....	20
7. Budget .....	22
8. Commercialization .....	24
9. References .....	26
10. Plagiarism Report .....	27

## Table of Figures

Figure 1- System Architecture .....	12
Figure 2- Bidirectional Recurrent Neural Network .....	14
Figure 3- Gantt Chart .....	21
Figure 4- Plagiarism Report .....	27

## 1. Introduction

In the evolving landscape of digital security, the need for robust and reliable user authentication systems is more critical than ever. Traditional biometric systems, such as fingerprint and facial recognition, have established benchmarks in security, but they come with limitations related to privacy, accuracy, and adaptability. This research explores a novel approach to authentication by leveraging mouse movement patterns as a biometric identifier. Mouse movement biometrics offers a unique opportunity to enhance security without relying on intrusive technologies, thus addressing some of the privacy concerns associated with conventional methods. [1]

Mouse movement dynamics such as the speed, rhythm, and patterns of cursor movements are inherently unique to everyone, making them a valuable feature for authentication purposes. The project aims to develop an advanced authentication system that uses sophisticated machine learning models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks, to analyze these dynamics. By capturing and analyzing the subtle nuances of mouse movements, the system will provide a high level of accuracy in distinguishing between genuine users and potential impostors. [1]

Moreover, this research addresses critical non-functional requirements such as security, scalability, and usability. The system will be designed with strong encryption and data protection measures to safeguard user information, ensuring compliance with privacy standards. It will also be optimized for scalability, allowing it to handle increasing user loads efficiently, and designed for ease of use to promote widespread adoption. Through these objectives, the project aims to contribute to the advancement of biometric authentication technology by offering a solution that balances security, privacy, and practical implementation in everyday scenarios. [1]

## **2. Background and Literature Review**

### **Overview of Keystroke Dynamics**

#### **Existing Research**

This research in mouse movement has been realized as having huge potential as a behavioural biometric in improving user authentication systems. This form of biometrics involves capturing unique patterns for every user in mouse movement cursor speed, trajectory, frequency of clicks, acceleration which helps in uniquely identifying the user.

#### **Convolutional Neural Networks (CNNs) in Mouse Movement Analysis**

The recent development applied CNNs in automating the extraction of spatial features from mouse movement data. CNNs efficiently identify complex patterns in trajectory data of two dimensions, which are hard to detect using traditional approaches. Studies have proved that, with high accuracy and robustness, authentication systems could capture and analyse path curvature, click distribution, and movement directionality by CNN. CNN reduces the dimensionality in mouse movement data by applying several convolutional layers, maximum pooling layers, and flattening layers, preserving all critical characteristics in the process to define user behaviour. [2]

#### **Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNNs)**

LSTM networks, belonging to the RNN class, have been increasingly investigated in modelling the temporal sequences specific to data representing the movement of mouse. LSTMs are powerful in scenarios where recognitions rely on timing and the order of movements to separate the genuine user from unauthorized users. The study has further shown that the LSTMs can capture the existing temporal dependencies of the mouse dynamics, including intervals and changes both in direction and magnitude, to better improve the recognition of the genuine users and detect the unauthorized users. Moreover, bidirectional RNNs have further supported analysing the sequence of mouse movements because its data processing goes in both forward and reverse directions, therefore enabling extraction of a more extended temporal pattern set. [2]

#### **Hybrid CNN-RNN Models**

Models that integrate CNNs for spatial features extraction and RNNs, especially LSTMs and GRUs for temporal sequence analysis, have high results in mouse movement-based biometrics. In other words, these hybrid models combine the strength of CNN and RNN to enable more robust user authentication. For example, a CNN can pre-process the raw mouse movement data to extract spatial features such as paths and clicking patterns, which are then inputted into an LSTM for temporal analysis. Doing so has been proven to enhance the system's capability in distinguishing between two users having similar movement behaviour.

#### **Privacy and Security Concerns**

While mouse movement biometrics presents a non-intrusive way of establishing the identity of users, it institutes some big concerns in terms of privacy and security. Easy collection of this data in a surreptitious way, as can be done for mouse movement data, may result in potential

violations of privacy. Researchers have proposed various measures to protect this data by considering methods to encrypt and anonymize it for storage in a database against possible misuse. It is also important to protect these systems against spoofing and replay attacks to avoid any doubt about reliability and security regarding mouse movement-based authentication.

### **Challenges and Future Directions**

Although a great deal of work concerning mouse movement biometrics with CNNs, RNNs, and their hybrids exists, it remains challenging. Various factors relating to the accuracy of such a system may include the type of device used, contextual factors, or even simple user fatigue. Real-time authentication with low latency along with high accuracy is still an open challenge. Future research will have to focus on flexibility in different situations and be more resistant to changes in users' behaviour.

### **Innovations in Data Collection and Preprocessing**

Some research in mouse movement biometrics has focused on novel ways of data collection and preprocessing. Models could be trained for different user behaviours and tested using publicly available datasets and custom datasets collected in controlled environments. Normalization, noise reduction, and feature extraction are some of the preprocessing techniques in use to ensure data quality and consistency. These techniques are therefore useful in standardizing raw mouse movement data and making it more appropriate for analysis, especially where there are diverse datasets

## **Research Gap**

### **Scalability and Real-Time Processing Challenges**

CNN and RNN (hybrid) models have proven their worth in the analysis of mouse movement-related data, the computational complexity behind these models is a challenge for real-time processing of tasks, especially in large-scale systems where the number of users is large. In future, efficient algorithms could be developed or possibly more efficient hybrid models designed, which would give the same level of accuracy but save on computational overhead for real-time authentication.

### **Data Privacy and Ethical Concerns**

Mouse movement data collection and analysis are surrounded by a number of privacy and ethical concerns, since it is possible to capture such data without the explicit consent of the user. Scarce research is dedicated to the development of appropriate frameworks that meet the competitive demands of accurate user authentication while preserving user privacy. Furthermore, there is a lack of effective ethical guidelines aimed at collecting, storing, and utilizing behavioural biometric data.

### **Integration with Multimodal Biometric Systems**

Most of the existing research has used mouse movement as a unimodal biometric. Integrating mouse movement into other behavioural or physical biometrics (e.g., keystroke dynamics,

voice recognition, gait analysis) would increase the security and accuracy of the system. Further studies must be conducted on how to best design and implement multimodal biometric systems where mouse movement would be combined with other biometric indicators.

### **Lack of Preprocess Datasets**

One of the major reasons is the lack of large, publicly available, and standardized datasets for mouse movement to be benchmarked and compared with the models and techniques used throughout the studies. Future research will be focused on creating and sharing comprehensive datasets capturing a wide range of user behaviours and scenarios to facilitate further robust model development and evaluation.

<b>Research</b>	<b>Individual Variability</b>	<b>Real-Time Processing</b>	<b>Scalability</b>	<b>Privacy</b>	<b>Integration</b>	<b>Adaptation</b>	<b>Generalization</b>
<b>Research A</b>	Yes	No	No	Yes	No	Yes	No
<b>Research B</b>	Yes	Yes	No	Yes	No	Yes	Yes
<b>Research C</b>	No	No	Yes	No	Yes	No	Yes
<b>Research D</b>	No	Yes	No	No	Yes	Yes	No
<b>Proposed Project</b>	Yes	Yes	Yes	Yes	Yes	Yes	Yes

### **Research Problem**

The problem of research in the domain of mouse movement behavioural biometrics is to develop an accurate, adaptive user-authentication system that could effectively distinguish between the legitimate user and impostors, utilizing subtle variations in mouse movement patterns. Current methods are restricted by variability in user behaviour across different devices and contexts, computational complexity required for real-time processing, and the need for system adaptation to changes in behaviour over time. Furthermore, the lack of standardization of data sets and the ethical concern with respect to privacy in using data are some reasons these systems do not diffuse massively and effectively. Therefore, dealing with these issues, this research investigates novel methods that aim to contribute toward increasing the reliability and scalability and at the same time develop a few ethical application scenarios of mouse movement behavioural biometrics for user authentication.

### **Integration of Mouse movement Analysis with Multi-Modal Biometrics**

Current biometric systems typically operate with only one modality and so these may limit the usefulness and reliability of the system. Mouse movement, though high in potential, has never been successfully integrated into other biometric modalities, including voice recognition, gait analysis, or keyboard patterns. The problem of the study is to design and develop an analysis of an integrated biometric system using mouse movement with many modalities to enhance



global accuracy and robustness. The challenge, therefore, is to implant a smooth integration of both approaches, allowing better performance in identification while keeping the system user-friendly and efficient.

### **Mitigating Privacy Concerns**

The behavioural biometric systems are very private, especially in the event of the enrolment of sensitive data from the user, such as mouse movement patterns. Mouse movement biometrics offer a less invasive technique compared to other more intrusive means of biometrics; however, they still raise concerns related to security of data and user consent. In this research, the concern of preserving privacy is alleviated through the deployment of robust techniques of data anonymization and encryption so that the system meets the stringent standards of privacy while maintaining anonymity to the very end. This research is about forming a mouse movement authentication system wherein the preservation of the privacy of the user does not come at the cost of accuracy and reliability of the biometric process.

### **Improving Mouse Movement Recognition Accuracy and Handling Variability**

Among the key challenges in mouse movement biometrics is to realize high recognition accuracy across different user environments and with multiple input devices. Variability in user behaviour, engendered by different devices or interfaces, or even user fatigue, will grossly deteriorate the performance of such a system. In this paper, advanced machine learning models such as Convolutional Neural Networks and Recurrent Neural Networks, mainly Long Short-Term Memory network, were used to enhance recognition accuracy. In this respect, the goal will be to develop resilient algorithms that can process and recognize patterns of mouse movement with accuracies despite these variabilities, hence improving the system's overall effectiveness.

### **Enhancing Computational Efficiency**

Most computationally demanding are systems using mouse movement biometrics, especially those utilizing complex neural networks. These will thus have to be optimized in regard to guaranteeing efficient performance in real-time applications without compromising the accuracy of such systems. This kind of research should target designing models that are more computationally efficient and able to process reams of data from mouse movements fast enough to perform desirable tasks. This will make the system feasible for deployment in many varied settings, more so in low-computational-resource environments, such that it remains practical for general use.

### 3. Objectives

The primary goal of this research will be to design an effective and accurate mouse behavioral authentication system with the mouse movement as the key biomarker. Implemented with deep and complex algorithms such as CNNs, RNNs, and LSTM networks, the system's goal is to identify and analyze small deviations of the users' mouse movements. This will ensure a very fine line can be drawn between real users and fake users even where users' behavior may change over time. [3]

Another important goal is to address the insecurity of the authentication process as critically. It is nominally intended to strong encryption techniques and other secure methods of managing biometric data to try to prevent breaches. This way, the system is to take into consideration both the problem of correctly identifying the user and the problem of protecting his and her personal information, to provide a completely safe solution meeting the high level of security and people's trust to protect their data. [3]

Also, it will focus on the complexity of the system and its usability, so that, at any time, it will be possible to add more features to the project. To this end it should be possible to expand the number of users to the desired level without degrading the performance of the system. This includes enhancing the data processing mechanisms to incorporate scale-up plans that will not interfere with the user experience. The structure of the system will be simplistic and user-friendly, thereby, reducing complexity, the demand for training to a minimum and enable wide applicability in the actual world setting. [3]

## 4. Methodology

A system overview of the mouse dynamic model shows how robustness and security for user authentication are enhanced by precision. It demonstrates how to create a hybrid model that blends Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). The analysis of the spatial data is done with CNN, and the analysis of the temporal data is done with RNN. In addition, this covers steps like feature extraction, sequence analysis, data preprocessing, data collecting, and model integration for authentication choices. The way these elements flow is depicted in the diagram.

### System Architecture

The system architecture for this project is designed to efficiently capture, process, and analyze mouse movement data for user authentication. At the core of the architecture is a client-server model, where the client-side application collects raw mouse movement data from users during their interaction with the system. This data, which includes metrics such as movement speed, direction, click patterns, and pauses, is then securely transmitted to the server for processing.

On the server side, the data undergoes a preprocessing phase where noise is filtered out, and relevant features are extracted. These features are then fed into a suite of machine learning models, including Convolutional Neural Networks (CNNs) for spatial pattern recognition and Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, for analyzing temporal sequences. These models work in tandem to identify and authenticate users based on the unique characteristics of their mouse movement patterns.

The architecture also incorporates robust security measures, including encryption protocols for data in transit and at rest, ensuring that sensitive user information is protected against potential breaches. Additionally, the system is designed for scalability, with the ability to handle increasing user loads through distributed processing and load balancing techniques. The architecture is modular, allowing for easy updates and maintenance, and is optimized for performance to ensure real-time authentication with minimal latency, making it both effective and practical for deployment in various real-world scenarios.

[1]

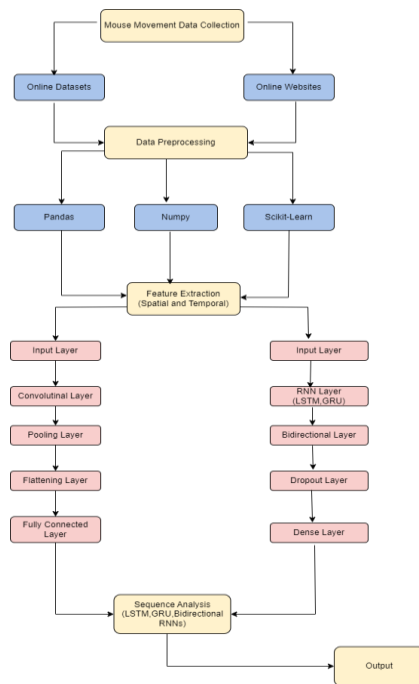


Figure 1- System Architecture

## Data collection

In this data collection phase, collect the mouse movement data sets from the online resources and online websites. In this data sets contain timestamp, mouse speed, scroll speed, velocity, mouse movement patterns and more. Additionally, can use python scripts and online web sites to capture the data resources.

The datasets used for the following project:

- Dataset from Kaggle - [Mouse dynamics](#)
- Dataset from GitHub - [BALABIT MOUSE CHALLENGE DATA SET](#)
- Dataset from datacite - [ReMouse - mouse dynamic dataset](#)

## Preprocessing

In this phase doing the preprocessing the collecting mouse dynamic datasets, remove the duplicate rows, manage missing values, convert the data into suitable way to that use in feature extractions. this phase can use python tools like NumPy, Pandas and scikit learn. This step ensures the accuracy of keystroke dynamics data, quality of the data and more. [4]

## Feature extraction

In this phase, extract the key features from pre-processed data using both RNN and CNN methods. Using CNN can extract the spatial features like Trajectory Images, velocity heatmaps and using RNN can extract the temporal features like average speed, velocity, scroll speed, button clicks changes etc.

The input layer gathers the dynamic mouse movement data that has already been pre-processed before beginning the CNN feature extraction procedure. Next, the convolutional layer applies several filters after identifying the spatial features in the mouse movement data. The data then enters a layer known as pooling. The most valuable characteristics are obtained, and the dimensionality of the feature maps is decreased by this layer. The data proceeds to the Flattening Layer from Pooling. This layer prepares the data for the fully linked layer by converting the 2D feature matrix to a 1D vector. Next, data travels to the fully connected layer, which is the last layer. To incorporate the spatial elements, this layer uses flattened vector data transfer through mouse movement dynamics and more completely connected layers. [5]

The pre-processed dynamic mouse movement data is obtained by the input layer of the RNN feature extraction procedure. The data processing is transferred from the input layer to the RNN layers, which include the Gated Recurrent Units and Long Short-Term Memory. The temporal sequences and patterns in the pre-processed mouse movement data are captured using LSTM and GRU. Data from the RNN layers is sent to the Bidirectional RNN layer, where features are extracted from both forward and backward directions. It a Bidirectional RNN layer to record mouse movement data and capture typing patterns from the past and future. Data then goes to the layer known as Dropout.

Additionally, both RNN and CNN feature extraction processes are parallel processes. In this process extract various features. CNN may extract from mouse movement data such as Trajectory Images, Velocity Heatmaps, Acceleration Patterns, Distance Matrices and more. RNN may extract from mouse dynamic data such as Temporal Sequences, button press duration, average speed, scrolling speed, velocity, acceleration and more. The above process is supported by TensorFlow, Keras, or PyTorch. After the feature extraction process then data moves to the next process called Sequence Analysis.

### **Sequence Analysis**

Sequence analysis uses a combination of CNN and advanced RNN technologies, such as LSTM, GRU, and Bidirectional RNNs, to analyze the temporal patterns in mouse dynamic data. Using CNN, first extract the spatial features from the mouse movement data and then find the crucial pattern. TensorFlow and Keras or Pytorch are used to train and fine-tune the LSTM, GRU, and Bidirectional RNN layers. This discovered feature and temporal features are transferred through the RNN models. Comprehensive robustness analysis for both spatial and temporal mouse movement features is made possible by this hybrid approach. This procedure guarantees the authentication mechanism's increased resilience and correctness.

### **Data pre-Processing tools**

#### **Pandas**

Pandas is a python library that more fast, powerful. This tool is used for the data preprocess tasks such as manager missing values, normalization and cleaning. [6]

#### **NumPy**

NumPy is an open-source Python library that provides support for matrices and large multi-dimensional arrays. which is used for handling and preprocessing keystroke data. [7]

## Data Analysis Tools & Technologies

### TensorFlow/Keras

The main framework for creating and executing deep learning and machine learning models is TensorFlow/Keras. CNN and RNN models may be developed with these tools. Complex neural networks may be developed and optimized with TensorFlow. It makes use of big scale data processing and GPU acceleration to better manage the complex patterns in the dynamic keystroke data. Conversely, TensorFlow's high-level API, known as Keras, facilitates quick prototyping and experimentation and helps streamline the model construction process. Both the CNN and RNN layers were put up in this project using Keras. CNN layers to extract geographical characteristics from dynamic keystroke data. RNN layers for extracting and analysing the temporal properties of keystroke dynamic data utilizing LSTM, GRU, and Bidirectional RNNs. The accuracy and performance of the model are guaranteed by the joint usage of TensorFlow and Keras. [8]

### Bidirectional RNN

Bidirectional RNNs is neural network that allows too process both forward and backward directions. The goal is the analyses both past and future data from input data. Bidirectional RNN has two recurrent hidden layers such as input sequence forward and processes it backward. Collect the output of these hidden layers and input them to prediction making final layer. Additionally, recurrent neural network cell also using the Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU). [9]

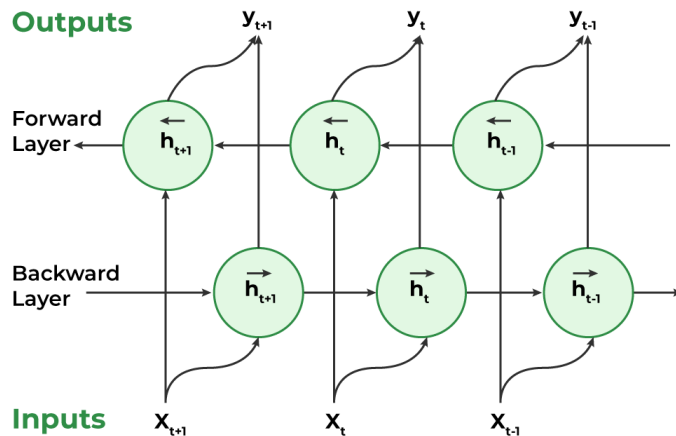


Figure 2- Bidirectional Recurrent Neural Network

## **5. Requirements**

### **User Requirements**

#### **Functionality**

##### **Accurate Authentication using mouse movement**

The system should rightly identify humans through their mouse dynamic movement patterns with very minimal errors, even under varying conditions. The user will expect the system to analyse and detect correct mouse movements in real time or near real-time and provide rapid identification.

##### **User-Friendly Interface**

The user interface should be very simple and easy to use, that enabling the modification of settings with the view of results and keeping of data without extensive training. Access should be made available on multiple platforms, such as desktop, considering the different needs and situations of users.

##### **Scalability**

The system should be interfaced with other existing security and access control systems so that mouse dynamics enhance the real-world applicability of the system. Exportation of data and results on mouse dynamics analysis in various formats for further analysis or reporting shall be allowed.

##### **Integration Capabilities**

The system interface well with previously installed security and access control systems. Additional features for mouse dynamics analysis include Allowing the user to export analysis data and results in various formats for further analysis or reporting.

##### **Security and Privacy**

The system must provide the secure administration of biometric data, respect user privacy and conform to appropriate data protection standards.

##### **Performance**

##### **High Accuracy and Reliability**

The system should realize high rates of correct authentication by mouse movement, very low false positives, and false negatives. Also, performance should be relatively consistent across different contexts and conditions, maintaining reliability in different settings.

##### **Robustness**

The system must be robust against alterations in mouse movement regarding speed, Mouse floor condition, mouse condition etc. It must be fault-tolerant: must handle errors in data input and processing well, recover from discrepancies and finally not let them affect the overall functioning.

## **Efficient Resource Utilization**

The system should run efficiently, optimizing the utilization of computational resources and energy to maintain good performance while minimizing operational expenses.

## **Functional Requirements**

### **Accurate authentication using mouse movement**

The system identifies the high system accuracy in recognizing people through their mouse movement patterns, targeting an accuracy of 95% or higher. This will ensure that the system correctly identifies various persons based on their unique Mouse Movements. The system should maintain a low mistake rate, minimizing false positives mistakenly recognizing someone and false negatives failing to recognize someone. This is very critical in ensuring that the system is reliable and trustworthy.

### **Real-Time Analysis**

The system should be able to evaluate dynamic mouse data and come up with recognition results in real time. This should mean that the processing latency shall be kept at the lowest, ideally within 10 seconds, to facilitate fast and efficient recognition.

### **Integration with Authentication Systems**

The system should be designed in such a manner that the interfaces are sustained with the existing authentication and security systems, such as access control methods. This should make mouse dynamic analysis a biometric security dimension for most applications. The system has to support the safe and efficient of data with other systems and be compatible with the many data formats while following security measures in the protection of sensitive biometric information.

### **Result Reporting**

The system shall generate and display the results regarding recognition, inclusive of identification outcomes and confidence scores. This feature gives customers clear and actionable feedback regarding how the recognition process is going. The system will facilitate the export of results and data in various forms, such as CSV or JSON. This functionality will support further analysis, reporting, and integration with other tools.

## **Non-Functional Requirements**

### **Scalability**

The need is, therefore, for a system that scales efficiently in terms of users and volume of mouse movement data. Scalability means that the system is able to cope with growing demands, such as an increasing number of authentication requests, without degradation in performance. This could involve refactoring the system architecture, applying load balancing techniques, or



ensuring the underlying algorithms and infrastructure can be expanded or adjusted to accommodate additional resources.

### **Reliability**

This means that the biometric authentication system performs its required function continuously without failure under specified conditions. The reliability ensures correct authentication and detection of unauthorized access, consequently keeping off any errors or downtimes in the operation. All this has to do with fault-tolerant mechanisms that gracefully handle unexpected situations, such as network failures or power outages, in a manner that assures quick bounce-backs with data integrity during such events.

### **Security**

In the development of a biometric authentication system, security will be optimized. The system will ensure good security through measures such as data encryption, secure channels of communication, and access controls that any sensitive information of users from being accessed by unauthorized persons, data breaches, or cyber-attacks. The system shall also conform to industry best practices concerning security to protect the product from any potential vulnerabilities against replay attacks, spoofing, and access to unencrypted data.

### **Usability**

If the system interface is too complicated or difficult to use, usability can readily become an issue that will hinder full adoption of the system. An easy authentication process, with clear instructions and feedback, is needed so that the user may easily comprehend it and continue through the system without special knowledge or large amounts of training.

### **Maintainability**

This will allow easier maintenance and upgrades to the system without affecting the overall functioning of the system. This will include user guides, technical instructions, and problem-solving materials. This in continuous maintenance and development. The support structure should be enabled in the system for proper management of issues and modification implementations. Regular updating and patches are required to remove the vulnerabilities and enhance the functionality; a procedure must be clear to make the updates known and installing the same.

## **6. Feasibility Study**

### **Technical Feasibility Assessment**

#### **1. Technology Suitability**

##### **Machine Learning Models**

The hybrid CNN-RNN model is the best method to analyse the mouse movements. It can capture spatial information (using CNNs) and temporal sequences (using RNNs). CNN can be used for extracting specific spatial variables from mouse movement data such as image-like grids showing mouse activities. While RNNs, especially LSTM or GRU variants, excel in understanding the temporal dynamics of mouse movement data. Integrating mouse movement data with other behavioural biometric systems (e.g., keyboard dynamics, gait analysis, voice biometrics). [10]

##### **Data Collection and Processing**

Utilizing online mouse movement datasets has in online resources. However, the quality and representativeness of these datasets are vital. Datasets should cover a diverse range of mouse movement patterns and variations to enable model robustness and generalizability. Techniques such as normalization, noise reduction, and feature extraction are necessary for processing raw mouse movement data for model training. These preparation activities must be efficient and capable of processing massive volumes of data and There are online websites can create own data sets. [10]

##### **Integration with Existing Systems**

##### **Behavioral Biometric Integration**

Integrating mouse movement analysis with other biometric systems involves building a feasible data fusion framework. This entails aligning different biometric outputs and establishing a decision-making framework that accurately depicts the combined authentication results. Developing APIs or interfaces for integrating mouse movement analysis with other systems will enhance seamless interaction between components. These interfaces should support common data formats and communication protocols [10]

##### **Data Privacy and Security**

##### **Encryption and Anonymization**

This includes the imposition of strong encryption measures for storage and transit, anonymization that prevents personal identity, and enhancement of privacy. The system respect data protection regulations, such as GDPR or CCPA, to be legally and ethically correct for treating biometric data.

## **Access Controls**

Ensure proper access control, whereby only authorized persons can view or make changes to sensitive data. This provides role-based access controls and multifactor authentication to system administrators. **Usability and Performance User Interface**

The user interface for end-users and administrators should be user-friendly, clean, and easy to navigate. It should be constructed based on usability testing and feedback from users. Ensuring that users clearly obtain feedback about authentication status and any action required on their part—that is, being prompted for re-authentication shall help make the user experience as seamless as possible.

## **Development and Maintenance**

### **Development Tools**

Next, leverage very deep machine learning frameworks like TensorFlow or PyTorch, along with libraries, for mouse movement analysis model development. These technologies provide support for training, evaluation, and deployment of the model.

### **Maintenance**

Maintenance and regular up-gradation are necessary for resolving the vulnerabilities and enhancing system. A framework of support in the diagnosis and overcoming of the challenges must be in place.

## **Economic Feasibility**

### **Development Costs**

- **Software Tools and Libraries:** There exist a few free libraries related to machine learning, data analysis, and biometric systems. Consider making use of tools like TensorFlow, PyTorch, or scikit-learn.
- **Paid Software Tools:** If any tools or libraries require licenses or subscriptions (e.g., MATLAB, particular data analysis tools), identify these charges.

### **Hardware Requirements**

- **Computational Resources:** Any specific hardware for development (e.g., GPUs for deep learning models). Costs for high-performance computing resources or cloud-based services (e.g., AWS, Google Cloud) should be examined.

### **Development Environment Costs**

- Integrated Development Environments (IDEs): Some IDEs or development tools could have related costs (e.g., PyCharm Professional, Visual Studio).

### **Implementation Costs**

- Software Integration: Costs for integrating varied software tools or libraries, including any additional software essential for integration.
- Testing Tools: If you need specific software for testing and validation of the mouse movement analysis system, add these charges.

### **Maintenance Costs**

- Software Updates: Ongoing charges for updating software tools or libraries. Some tools could require periodic renewals or upgrades.
- Technical Support: If any of the tools or libraries come with support plans, consider in these charges.

### **Contingency Costs**

- Unexpected Expenses: Allocate Budget for incidental expenses, such as for software and tools that might be needed during the project.

## **Schedule Feasibility**

The project schedule ranges from July 2024 to May 2025 and is strategically planned to ensure smooth development and timely completion of the mouse movement analysis system. In the timeline, **several key phases:**

### **Data Collection (Weeks 1-9)**

First, mouse movement data are collected from some online datasets. Emphasis should be paid on getting comprehensive and diverse data for the analysis. Further, websites will be used to capture the data online.

### **Preprocessing (Weeks 10-15)**

The recorded data on mouse movement will be normalized into a standard form. The data will then be further segmented into strides and steps for detailed analysis, while applying noise reduction techniques to improve the quality and reliability of the data.

### Feature Extraction (Weeks 16-26)

A Convolutional Neural Network (CNN) will be developed to extract spatial features from the mouse movements data. A Recurrent Neural Network (RNN) like speeds, velocity etc. will be employed to handle temporal features. Features will be extracted and visualized to prepare the data for model integration.

### Model Integration (Weeks 27-34)

This will finish the mouse movement features integration. A hybrid model combining CNN and The RNNs architectures shall be developed, and the integrated model will be tested to gauge performance and make the necessary adjustments.

### Evaluation (Weeks 35-40)

Performance metrics will be defined to measure the model's effectiveness. Initial testing will be carried out on the model to estimate its accuracy and reliability. Model validation by using real-world data will be done to ensure that it is of practical use.

### Report Preparation (Weeks 41-43)

A final report will be written that includes the findings, methodologies, and results of the project. This timeline provides a systematic development of tasks in a clear sequence of milestones. The dependencies of the tasks are given to take care of the workflows, maintaining efficiency to complete the project on time. The structured schedule allows for flexibility in case any unforeseen challenges arise, ensuring successful development of the mouse movement.

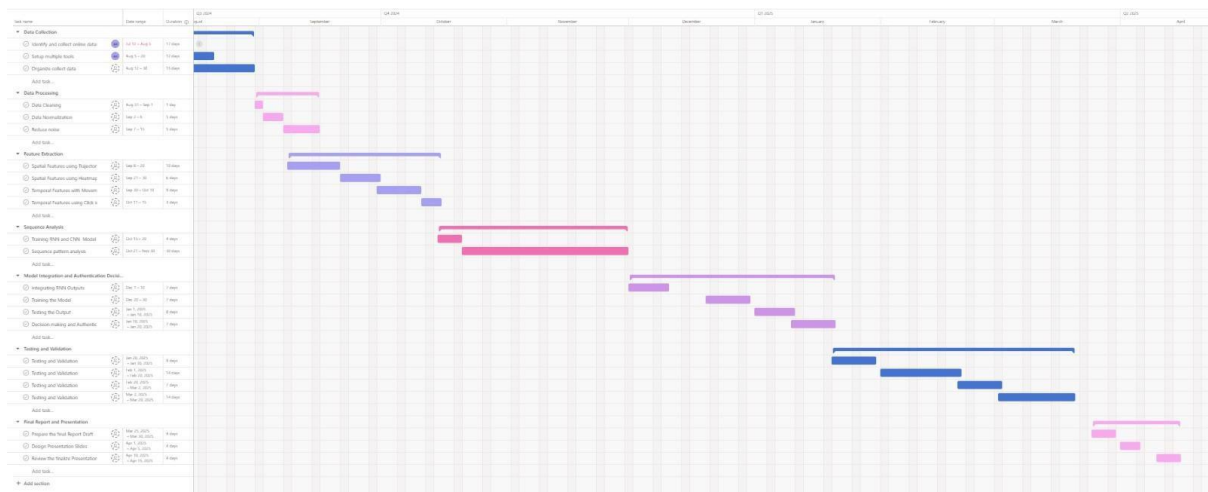


Figure 3- Gantt Chart

## **7. Budget**

### **1. Data collection Online**

#### **Datasets**

Cost of purchasing or accessing mouse dynamic datasets from online repositories or academic sources.

Estimated Cost: \$40

### **2. Software Tools**

#### **Development Tools and Libraries**

Machine Learning Libraries: Licensing fees or subscription costs for libraries such as TensorFlow,

PyTorch, and scikit-learn.

Data Analysis Tools: Software for data cleaning, preprocessing, and visualization (e.g. Python libraries) Estimated Cost: \$50

#### **Integrated Development Environment (IDE)**

Tools and plugins needed for coding and development.

Estimated Cost: \$20

#### **Version Control and Collaboration Tools**

Costs associated with platforms like GitHub or Bitbucket for code management and collaboration.

Estimated Cost: \$0

### **3. Hardware**

#### **Computing Resources**

Cloud computing services for model training and evaluation.

Estimated Cost: \$150

### **4. Miscellaneous Expenses**

#### **Documentation and Reporting**

Costs related to printing, binding, or creating professional reports and presentations.

Estimated Cost: \$20

### Contingency Fund

Reserved for unforeseen expenses or additional requirements that may arise during the project.

Estimated Cost: \$30

Expense Category	Estimated Cost
Online Datasets	\$40
Development Tools and Libraries	\$50
Integrated Development Environment	\$20
Version Control and Collaboration Tools	\$0
Computing Resources	\$150
Documentation and Reporting	\$20
Contingency Fund	\$30
<b>Total Cost</b>	<b>\$290</b>

## **8. Commercialization**

The mouse movement analytics system introduces a new impetus to authentication technology when combined with other behavioural biometric models. Several possible approaches and business models can be developed to commercialize such a system. This will be implemented throughout the whole project from creation to hitting the market thus offering a deliberate road to success.

### **1. Market Opportunities Cybersecurity Sector**

Organizations are seeking innovative ways to face the challenge of these compliance requirements with rising rules set forth regarding data protection. The mouse movement analysis solution can be marketed as a powerful solution to meet the demands of tight security. [11]

### **2. Business Models**

#### **Licensing Model**

License mouse movement analysis technology to cybersecurity firms, robotics manufacturers, and healthcare. Providers will make out a continuous stream of revenue while partners are allowed to embed the technology into their solutions. Permit access to the system's functionality via API to let third-party developers integrate the mouse movement analysis within their applications. [11]

#### **Subscription Model**

License mouse movement analysis technology to cybersecurity firms, robotics manufacturers, and healthcare. Providers will make out a continuous stream of revenue while partners are allowed to embed the technology into their solutions. Permit access to the system's functionality via API to let third-party developers integrate the mouse movement analysis within their applications. [11]

#### **Partnerships and Joint Ventures**

Partner with firms producing biometric hardware or robotics and integrate mouse movement analysis to help such companies provide better value bundle offers to customers. Ally with academic/research institutions to further develop and fine-tune the technology, utilize their expertise and network in the process of market entry. [11]

#### **Direct Sales**

Market and sell the mouse movement analysis system directly to the largest corporations, security companies, technology companies, and individual solutions for sector needs with possibilities for customization. Further development and sale of final consumer goods for mouse movement analysis, such as personal secure access solutions [11]



### **3. Implementation and Growth Strategy**

#### **Market Entry Strategy**

Provide pilot projects run with key companies in the industry to demonstrate the effectiveness of the system and gather real-world feedback. Show off the technology at trade shows to attract potential Partners and customers.

#### **Marketing and Promotion**

Run focused advertising campaigns that highlight the unique benefits and advantages of mouse movements. It would be an analytics system for different sectors. Case studies and success stories should be published to create trust and to have a view of its worth in real-world situations.

#### **Scalability**

Ensure that this system is designed for scalability, so that at times when adaptation and integration are needed with current technologies and infrastructures, this system will allow such with ease. Look for global opportunities to expand. Identify countries or markets in high demand of advanced biometric solutions.

Description of personnel and facilities

## 9. References

- [1] Y. & L. H. Chen, “A deep learning approach to mouse movement biometrics,” 2018.
- [2] Y. W. X. & W. Li, “A novel mouse movement-based authentication system,” 2017.
- [3] Y. & L. Z. Zhang, “A robust mouse movement-based authentication system using deep learning,” 2016.
- [4] J. L. a. X. Wan, “Mouse Movement Biometrics for User Authentication: A Survey,” 2015.
- [5] Y. C. a. H. Li, “A Deep Learning Approach to Mouse Movement Biometrics,” 2018.
- [6] pandas.pydata, “pandas,” pandas.pydata, [Online]. Available: <https://pandas.pydata.org/>.
- [7] numpy, “numpy,” [Online]. Available: <https://numpy.org/>.
- [8] keras, “About Keras 3,” [Online]. Available: <https://keras.io/about/>.
- [9] geeksforgeeks, “Bidirectional Recurrent Neural Network,” 2023. [Online]. Available: <https://www.geeksforgeeks.org/bidirectional-recurrent-neural-network/>.
- [10] I. B. Y. & C. A. Goodfellow, “Deep learning.,” 2016.
- [11] A. & P. Osterwalder, “Business model generation: A handbook for visionaries, strategists, and entrepreneurs,” 2010.

## 10. Plagiarism Report

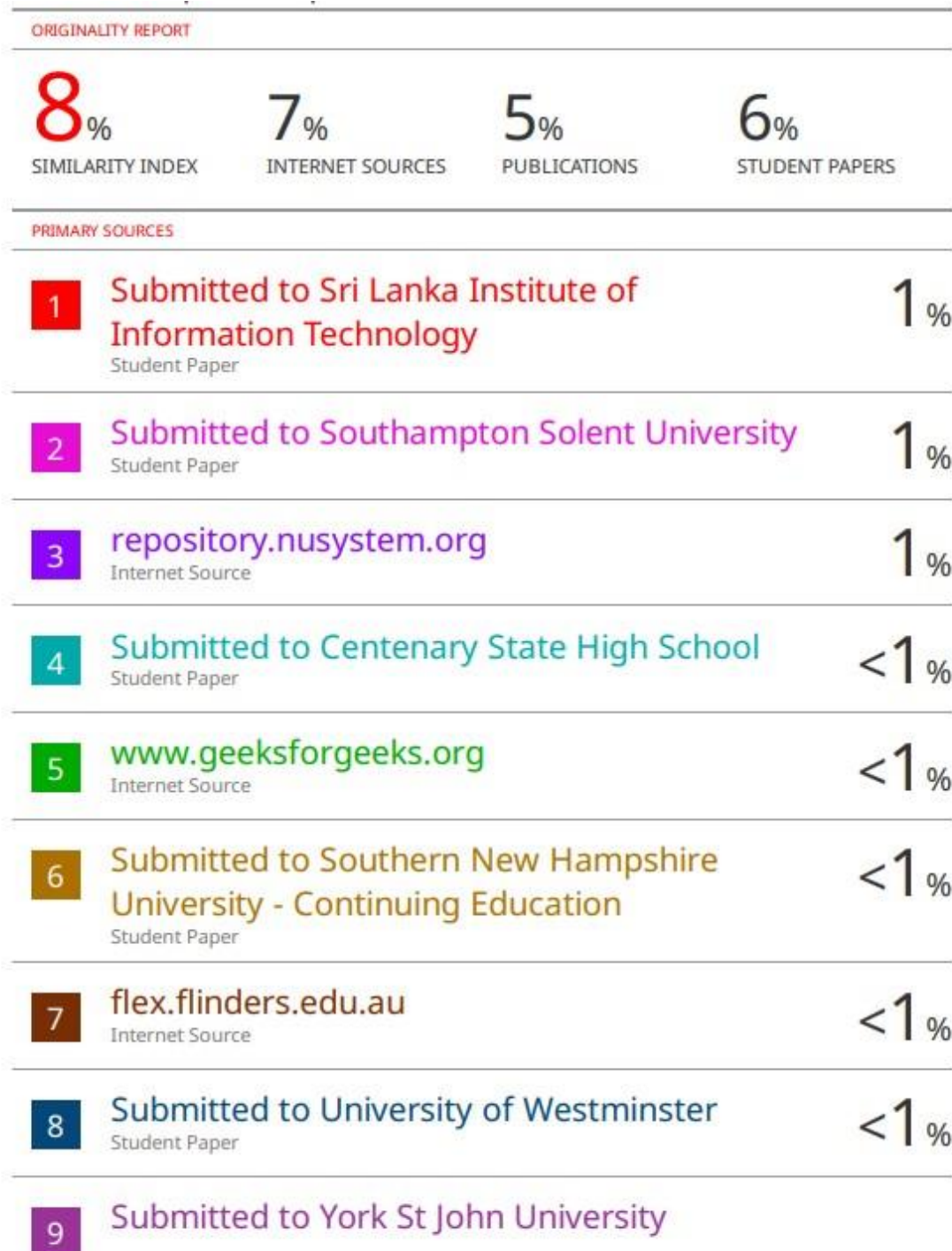


Figure 4- Plagiarism Report