

Modelling and Predicting User Behaviour

A major project report submitted in partial fulfilment of the requirement
for the award of degree of

Bachelor of Technology

in

Computer Science & Engineering / Information Technology

Submitted by

Aastha Verma (201432)

Vipul Arora (201151)

Under the guidance & supervision of

Dr. Rakesh Kanji



**Department of Computer Science & Engineering and
Information Technology**

Jaypee University of Information Technology,

Waknaghat, Solan - 173234 (India)

CERTIFICATE

This is to certify that the work which is being presented in the project report titled ‘**Modelling and Predicting User Behavior**’ in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology in Computer Science & Engineering / Information Technology** submitted in the Department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by **Vipul Arora (201151)** and **Aastha Verma(201432)** during the period from August 2023 to May 2024 under the supervision of **Dr. Rakesh Kanji** (Assistant Professor(SG), Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat.


Aastha Verma

Aastha Verma

201432


Vipul Arora

Vipul Arora

201151

This is to certify that the above statement made by the candidate is true to the best of my knowledge.



Dr. Rakesh Kanji

Assistant Professor (SG)

Department of CSE & IT

Jaypee University of Information Technology

CANDIDATE'S DECLARATION

I hereby declare that the work presented in this report entitled '**Modelling and Predicting User Behavior**' in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology in Computer Science & Engineering / Information Technology** submitted in the Department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat is an authentic record of my own work carried out over a period from August 2023 to May 2024 under the supervision of **Dr. Rakesh Kanji** (Assistant Professor(SG), Department of Computer Science & Engineering and Information Technology).

The matter embodied in the report has not been submitted for the award of any other degree or diploma.


Aastha Verma

Aastha Verma
201432


Vipul Arora

Vipul Arora
201151

This is to certify that the above statement made by the candidate is true to the best of my knowledge.



Dr. Rakesh Kanji
Assistant Professor (SG)
Department of CSE & IT
Jaypee University of Information Technology

ACKNOWLEDGEMENT

Firstly, we express our heartiest thanks and gratefulness to almighty God for His divine blessing which made it possible to complete the project work successfully.

We are grateful and wish our profound gratitude to Supervisor **Dr. Rakesh Kanji**, Associate Professor, Department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat. Deep Knowledge & keen interest of our supervisor in the field of “**Modelling and Predicting User Behavior**” to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, and valuable advice have made it possible to complete this project.

The in-time facilities provided by the Computer Science department throughout the project development are also equally acknowledgeable.

Last but not the least, our sincere thanks to all our teachers, lab assistants and friends who have helped us straightforwardly or in a roundabout way in making this project a win.

Finally, we acknowledge with due respect the constant support and patience of our parents.

TABLE OF CONTENT

<u>TITLE</u>	<u>PAGE NO.</u>
CERTIFICATE	i
DECLARATION	ii
ACKNOWLEDGEMENT	iii
CONTENTS	iv- v
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS	vii
ABSTRACT	viii
CHAPTER 1: INTRODUCTION	1-12
1.1 TYPES OF AUTHENTICATIONS	1
1.1.1 USER IDENTIFICATION	3
1.2 PROBLEM STATEMENT	4
1.3 OBJECTIVE	8
1.4 SIGNIFICANCE AND MOTIVATION	10
1.5 ORGANIZATION	12
CHAPTER 2: LITERATURE SURVEY	13-26
2.1 OVERVIEW	
2.1.1 INTRODUCTION	13-14
2.1.2 SUMMARY OF RELEVANT PAPERS	15
CHAPTER 3: SYSTEM DEVELOPMENT	27-44
3.1 REQUIREMENTS AND ANALYSIS	27
3.2 PROJECT DESIGN AND ARCHITECTURE	29
3.2.1 PROJECT DESIGN	36
3.2.2 PROJECT ARCHITECTURE	37
3.3 DATA PREPERATION	38
3.4 IMPLEMENTATION	41
3.5 ALGORITHMS USED	43
3.6 KEY CHALLENGES	44
CHAPTER 4: TESTING	45-51
4.1 TESTING STRATEGY	45

4.2 TEST CASES AND OUTCOMES	48
CHAPTER 5: RESULTS AND EVALUATION	52-56
5.1 RESULT	52
5.2 PERFORMANCE EVALUATION	56
CHAPTER 6: CONCLUSION AND FUTURE SCOPE	57-62
6.1 CONCLUSION	57
6.1 FUTURE SCOPE	60
REFERENCES	63-65
APPENDIX	66

LIST OF FIGURES

<u>FIGURE</u>	<u>PAGE NO</u>
1-Raw Data Processing Pipeline	30
2-Website Design	31
3-Registration Form	32
4-Authentication stage 1	33
5-Authentication stage 2	34
6-Training your biometrics	35
7-Biometric System Architecture	36
8-Dwell time, Flight time	36
9-Inter key latency	37
10-Training Dataset Image	38
11-User Behavioural Biometrics Model	39
12-User Screen Movement	39
13-Training Performance	40
14-Event Transition	40
15-User's Screen Movement	52
16-Different users line plot	53
17- Dataset of ten samples for one user	53
18-Typing pattern of a user vs an imposter.	54
19 - User's screen movement	55
20- Different users line plot	55

LIST OF ABBREVIATIONS

KD	Keystrokes Dynamics
SVM	Support Vector Machine
CA	Continuous Authentication
FAR	False Acceptance rate
FRR	False Rejection rate
EER	Equal error rate
UDKL	Up Down Key Latency
DUKL	Down Up Key Latency
UUKL	Up Up Key Latency
DDKL	Down Down Key Latency
DT	Dwell Time
FT	Flight Time
KNN	K-Nearest Neighbors

ABSTRACT

In the last couple of years, the merging of keystroke dynamics with mouse activity has been receiving significant attention due to the increasing realization that this indeed is a powerful method of behaviour modelling and consequently predicting in numerous digital spheres. This article focuses on the feasible two different modalities of information analysis and how they will help in analysing user behaviour.

Significance of keystroke dynamics and mouse activity is not possible to diminish since they can not only reveal the behaviour but also so much more. Keystrokes dynamics track the spectrum of the way people type like speed, rhythm, and errors, and mouse activity contains data about the motion of mouse cursor, clicking rate and scrolling pattern. The combination of the components offers researchers an opportunity to develop a more holistic idea about user conditions and digital design.

This article is an overview of how combined keystroke dynamics and keyboard activity is used in different areas of cybersecurity applications, including user authentication, behaviour-based anomaly detection, and user interface improvement. Utilizing fitted-together-feature, ensemble-learning and deep neural network, predictive models is used so that user actions can be predicted with a high degree of accuracy.

Additionally, the article discloses the difficulties and factors that analysis of keystroke and mouse activity encompass, e.g. means of data synchronization, feature selection and model interpretability. To take advantage of the combined power of this integrated way, the key is to overcome these difficulties involved.

From a final word, it follows that a combination of keyboard movements and mouse activity will big up the modelling of keystroke dynamics as well as the anticipating of user behaviour. Utilizing their different strengths, researchers, and technical people in using this blend can develop the more reliable methods for providing good user experience, security, and stability of the system in the digital environment.

CHAPTER 1 - INTRODUCTION

1.1 INTRODUCTION

Fraud and impersonation are the two most common reasons that are the threats to the data, digital network and computer system security. A lot of systems of authentication through the Web have been developed to protect the commercial transactions and to secure the data. Among the concepts of account username and password, IP address filtering, message digest authentication, etc. are the most popular ones. It is obvious that these systems will not be all perfect, thus, it is only possible to make them better with more and more security. One case of it is that, if a user chooses a weak password, then it can easily be cracked. This has led to many research studies being done to find ways of processing user input data so that it can be used as a form of authentication. Among the many new approaches, one of those has been Keystroke biometrics which is the habitual patterns or rhythms a person exhibits while writing on a keyboard device.

Compared to alternative biometric schemas, keystroke has the primary advantages that:

1. There is no external device, for instance a scanner or detector, required. Everyone needs a keyboard.
2. The "rhythm" or the "pattern" of the users is a very reliable indicator.
3. It can be implemented along with the existing authentication systems without any changes.

TYPES OF AUTHENTICATIONS

Knowledge Based Authentication (KBA):

As knowledge-based authentication recalls a set of information which is unique for the user and only known to him to authenticate the user, can serve as a means of verifying the identity.

Common examples include:

Passwords: Users get a certain secret chained string, known by them only, to get in to the system.

Personal Identification Numbers (PINs): Just like the passwords, pin is a numeric code which serves as a mean of identification in a device.

Security Questions: Participants are asked to answer the stated questions (for instance, "What was your mother's original surname before she got married?", etc.) to ascertain their identity. Just as knowledge-based authentication is widely used because of its simplicity and familiarity to the users, biometric identification is less common but more secure sign in method. Consequently, it will still be susceptible to various security threats including passwords guessing, social engineering, and phishing.

Biometric Based Authentication (BBA):

Biometric authentication employs distinctive physiological and behavioral traits of the people which can be used to confirm their identity in the process of identification. Examples of biometric-based authentication methods include: Examples of biometric-based authentication methods include:

Fingerprint Recognition: The users place their fingertips on the fingerprint reader or sensor to make sure that they are allowed to access the data.

Facial Recognition: Face recognition system employs an evaluation based on the facial features of an individual to confirm their identities.

Iris or Retina Scanning: The types of technologies that are crucial for authentication are a scanning of the peculiar patterns present in the eye's iris or retina.

Biometric-based authentication provides a higher level of security as biometric features are seldom duplicated through forgery or any forging technique. Thus, it brings about comfort to users because they do not have to memorize passwords, nor do they have to keep on physical tokens. Nevertheless, the application of biometric systems accompany some barriers, including privacy issues and the diverse accuracy of biometric recognition under different lighting and sensor quality.

Object Based Authentication(OBA):

Compared to token-based or password-based authentication methods, object-based authentication is a way of authentication in which a real physical object or token is used to verify a user's identity for access to an online account. Examples of object-based authentication methods include:

Smart Cards: People inserts smart cards into card readers, which contains chips that has encrypted keys or certificates to allow them access in to the applications.

Security Tokens: Users benefit from these small digital devices as they provide one-time passwords or authentication codes which are to be entered along with a person's credentials.

RFID Tags: As an example, Radio-frequency identification (RFID) tags can be deployed at the proximity-based log in point which is used as a tool for authentication.

The object based of authentication establishes an extra way of security that layer on the top of the passwords or PINs because a user must possess the physical token as well as the user's knowledge. Yet, consumers can find it less convenient relative to traditional methods as they will need to have the cards with them.

1.1.1 USER IDENTIFICATION BASED ON MOUSE ACTIVITY FINGERPRINT

The key to online safety and identity is proper user authentication that increases with the increasing digitalization. Despite that, these traditional ways like username-passwords are still applied but they are full of security holes. However, this raises curiosity regarding new biometric based approaches that may be used to overcome such challenges. The first boundary is a bit difficult to identify on the basis of a mouse movement and this is located in a special area. Individuals use varying manners when dealing with their gadgets. This also shows up in how a person walks or uses a mouse. It is referred to as mouse fingers, an online kind of behavior characterized by repetitive actions, clicking, and scrolling through a digital medium. Traditional identifiers like passwords are quite different from the properties of mouse clicks. They may include biometrics, but we've still never seen someone log on using their fingerprints. The properties ensure the continuity and thus, a behavioral trait that can help in pointing out individual persons.

This new direction is promising in many aspects including enhanced security and improve of use of systems. Nonetheless, complex software techniques enable a "digital signature" of each connection via close analysis of minute peculiarities about how people interact with computer mouse: short mouse taps, fast tempo, clip, and click. Such unique characteristic becomes a trustworthy security means, reducing risk related to other common ways of identification.

However, that it is not the only implication of user authentication that is based on fingerprint functionality towards application security. Such approach is discreet and adaptable for use in personal learning environment, such as e-learning or workplace sharing platforms. Systems personalize interaction, identify users by natural interaction, show preferences,

improve interface usability towards better understanding and user engagement in digital realm.

To understand such complexity in this reform, people should know about their programs, cross-sector uses and key decisions about biometrics collection and use. This research provides insights into current debates about different ways to authenticate computer users and looks forward at potential directions for future e-security based on use of fingerprints.

1.1.2 USER IDENTIFICATION BASED ON KEYSTROKE DYNAMICS

The user authentication using keystroke dynamics is a technology that focuses on the input typing habits of users in order to recognize their identity. At first, during the enrollment process the system collects and analyzes various habitual typing features of a user, such as the timings between keystrokes and the time a user holds their finger(s) on the key(s) as well as the time between keying. This kind of information is used to determine user-specific typing profiles on the account of extracting this type of features. The machine learning algorithms use these features as a training set to learn a model of any individual user. During authentication, the system senses user's real-time typing behavior, comparing it with the previous only to compute the confidence value in the authentication process. To gain access the system will check the keyboard motions and if a best fit can be found within an acceptable range the access is granted. Finally, this method has got a number of advantages such as the absence of invasiveness, continuous authentication, and being highly secure (since the identity is established using biometric data) with low cost. Though, obstacles that threaten including, fluctuate of typing behavior, user privacy and potential rollback attacks is a key issue to resolve. As a whole, user authentication and measuring of keystroke dynamics would ensure a more secure and convenient environment to authenticate people, especially in the cases of continuous monitoring and undetectable verification.

1.2 PROBLEM STATEMENT

The most important security risk lies in using legacies of security as part of the quickly changing digital environment. There will be a need for new IS considering the increase of demand for more power and alternative user authentication. Another aspect is that, as users identify themselves through various behaviors while interacting online, mouse finger

printing capabilities can help identify them. However, it is worth mentioning these shortcomings which are left unspoken of in this new process.

One Time Authentication

With so much being done as straight forward today when it comes to digital landscape an important issue of having secure access to online services and systems is conceivable. Apart from the most ancient authentication methods, such as passwords and PINs, which can be susceptible to a range of security threats, including phishing, password leaks, and brute-force attempts, sometimes, the strongest authentication method is still a password and PIN. To address these concerns, one-time authentication problems which are considered an effective solution. One-time authentication is a type of authentication in which the user has to produce temporary authentication credentials that are valid for the duration of a single login session or for a specified transaction. Such an authentication mitigates statistical identity theft risks which only single usage credentials might entail. Though it is relatively difficult to implement the one-time authentication systems properly, one should distinguish between various ways to accomplish this task. Those are, for example, the demand for the adherence to the usability requirements, scalability to the increase of user base and mitigation of the replay attacks. Lastly, the authentication infrastructure of the existing users also should be taken into the account. Essentially, the balance of security and usability is also essential such that one-time authentication methods do not bear rigorous burden on users so as to successfully prevent unauthorized access to information. Resultantly, a huge emphasis must be assigned to the designing and implementation of one-time authentication procedures which should match the level of the user's demands. This would significantly contribute to the improvement of the general level of cybersecurity and the avoidance of data leakage possible due to the actions of criminals. This problem statement will seek to reveal the most significant and notable issues surrounding the adoption and introduction of one-time authentication mechanisms, as well as suggest unique solutions that are designed to cut to the crux.

User Annoyance And Time Consumption:

With the unremitting development of the digital environment, the need for solid authentication systems for establishing highly-secured online systems and services is ever growing, greater and greater. Recognition of CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) as a testimony to the widespread use of

CAPTCHA in online security has been active for years now. CAPTCHA serves as a screen against the attacks of robots by differentiating between human users and the malicious bots. Nevertheless, CAPTCHAS has been able to resist from automated abuse, yet with the implementation there lies a few challenges and limitations. One of the major concerns is the user interface, because CAPTCHA can be a very uncomfortable and irritating experience for legitimate users, resulting in them quitting not having reached their goals, thus, a decrease in the level of engagement. On top of that, the usefulness of classic CAPTCHA methodology diminishes over time because of the fact that AI and learning algorithms have achieved more and more progress in bypassing the traditional captcha challenges with higher and higher rate of success. Moreover, availability issues concern those people with the disability who may not be able to voice or listen which can make completing visual or auditory CAPTCHA tasks difficult. Ensuring the sufficient level of security is possible only if there is execution of CAPTCHA-based authentication systems with the least impact of the users' experience and fulfillment of their accessibility requirements. Consequently, new techniques of CAPTCHA design and deployment, as well as usability and accessibility, must be explored. These techniques must entail strategies that can guarantee user convenience and that maintain accessibility for all users. The aim of this problem statement is to take a closer look at the core obstacles that are associated with CAPTCHA-based verification and suggest new approaches that can aid the process to become more effective, easy to use and also inclusive enough to uphold the affairs of the digital world.

Lack Of Continuous Monitoring :

Today's digital sphere might be seen as a volatile place with cyber criminals who are always inventing new techniques, which makes frequent security audits and only responding to incidents out of focus to completely protect priceless data and system. Another critical problem organisational faces is that they are not equipped with the tools to perform continuous monitoring of the computing systems, the networks that connect them and the applications that run on them. Monitoring in real time is done by the acquisition, of data, its analysis, and the correlation of relevant security themes to indicate possible threats, weaknesses, and unusual behavior in a timely fashion. Yet some companies make use of while the others take manual or traditional monitoring approaches that create holes in protection and a time frame during which cybersecurity events might be undetected.

There is a risk of leaving such open spaces, which can be used for advanced persistent threats, insider attacks and data breaches, causing financial losses, erosion of reputation and failure

to comply with regimes. On the other side, wider scope, as well as the complex nature of newest kinds of digital architecture, only worsen the issue, making it really hard to remain aware and in command of the security conditions. Additionally, the shortage of cybersecurity professionals with the adequate skills for monitoring leads to the problem since the organizations struggle with the difficulty of keeping and providing proper monitoring measures. And that is why there is a sense of urgency to bridge the gap between security and the digital world by introducing automated, proactive, and scalable monitoring solutions which yield immediate assurance of online security and firm awareness of possible threats and vulnerabilities. This problem statement is aimed at emphasizing the fact that continuous monitoring is the key to cybersecurity due to the dynamic nature of the cyber landscape. And the statement argues for investment in technologies, processes, and staff development initiatives to support their institutions generate an expert cybersecurity force.

Role Of Behavioral Biometrics:

In the area of digital authentication, outmoded methods such as passcodes and token users face the emerging threats of advanced cyber-crimes such as phishing, credential stuffing, and account takeovers. In the context of these weaknesses biodata characterization as a promising solution is among the most obvious, which essentially relies on the distinctive behaviour exhibited by the user to verify the identity. Although behavioural biometrics gives an opportunity to digitize authentication systems, the field has number of difficulties to address.

Another difficulty is the demand for capable and strict authentication algorithms that will be able to separate the approved users from the bad actors professionally. The building of these models requires an abundance of data, the analysis of user behaviour, such as their keystroke dynamics, movements of their mouse and on-screen touches. Not only that, but behavioural biometrics should cope with the changeable and unstable nature of the human behaviour that makes it rather hard to create well-fitting reference levels for individual authentication.

Besides, behavioural biometrics technology market uses rise also spurt privacy and consent security issues. To render behavioural data, the process of collecting and analysing this data by design requires having, what might be regarded as a sensitive information about individuals' interactions with digital systems, which leads to ethical and legal discussions about data protection as well as user consent. Besides that, scalability and interoperability of behavioural biometrics across different interface vendors and devices is among the main confrontations.

The binary digit of superintelligence is introduced as an outcome of the merge of machine intelligence with human intelligence, aiming for intellectual excellence significantly surpassing current levels.

Furthermore, it should be acknowledged that behavioural biometrics may be vulnerable to attacks that are difficult to identify by adversaries as well as evasion methods that are more complex. Impostors may try to role or impersonate a real user behaviour to get in the front of behavioural biometric authentication systems that is why constant studies of this systems and development of counter measures is critical to prevent such attacks.

1.3 OBJECTIVE

Develop A Robust And User-Friendly Authentication System:

An effective and user-friendly authentication system involves all the security improvement and recently there is the addition of user-friendliness. Through the use of such advanced authentication techniques like biometrics and single sign-on, organizations are able to efficiently secure their company's data and systems against internal and external unauthorized access, while at the same time minimizing the impact on user experience and convenience. The logging in of users becomes simple and convenient as the online bank minimizes the complicated issues thereby, increasing customer satisfaction. Similarly, the frameworks help spare expenses in dealing with password resets and account lockouts, meet compliance on implementing security measures, improve insider threat prevention, develop users` and customers` trust, and boost reputation. In conclusion, ensuring a system with both solid security feature but user-friendly experience to the customers is considered as a priority when mentioning the modern day's technology development.

Implement Continuous Authentication System:

Continuous authentication systems provide a holistic approach toward security which is achieved by reviewing user specific behavioural patterns and changing identity verification based on the security requirements, which in turn, greatly minimizes the opportunities of unauthorized access and solidifies system resistance against cyber-attacks. Through the use of the process of the caliber which is not anivativous while operating in the background and ensuring that the user experience is not disrupted, continuous authentication cultivates more protection while preserving the user productivity. Banking on the part of business analytics

and machine learning algorithms that sends forward accurate outputs, these systems successfully differentiate between the routine user actions and danger signs; thus, reducing the chances of false alarms, and supporting adaptive control mechanisms that keep on adjusting their access authority best of their ability based on real-time risk assessments. Wholly, sustain authentication constitutes compliance readiness by keeping chronic logs of users' transactions and authentication events, providing the companies essential power for audit and the show that they follow the data protection regulations and standards on security matters.

Check User Authenticity Using Mouse Movement And Keystroke Dynamics:

Users authentication via keystroke behaviour analysis and mouse movement involves identification of unique behaviour patterns displayed by the persons as they type and use mouse. By gathering and storing these details while the system streamlines and extracts features to prepare personal user models during registration. Machine learning algorithms build such models and use them for users' real-time authentication by comparing their pattern of presently used symbols and mouse manipulation with the stored templates. Such new strategies touch on the superior security by applying the biometric-like behavioural attributes that are very "challenging" to imitate, and basic success to assure the frictionless and transparent experience to the end-users involved. An added complexity comes with users with varying behaviour patterns and privacy considerations, which necessitate their resolution to guarantee success and ethical usage.

Exploring Positive Behaviors Over Time:

In terms of the investigation of positive behaviour through time, one should perform the research and analyse how individuals display constructional actions, habits, as well as ideas throughout some time span. Typically this reporting implies continuous studies or data analysis methods to capture the evolvement of social developments, systemization of events and track the changes in behaviour which happens over time. Positive behaviour can be defining by various acts, for e.g. kindness, working actively, resilient and personal progress. Through studying of a positive behaviour over a period of time, researchers and practitioners can find out critical factors that help building up such behaviour, and further interventions and strategies can be applied to assist the development of such behaviour. Furthermore, it is possible to identify the important role of positive behaviour that influence the well-being of

the individuals, among families, in the societies, to the immediate and far-flung communities in a very positive way by re-echoing and providing a knowledge that can foster resilience and positive change in various situations.

1.4 SIGNIFICANCE AND MOTIVATION

Improved Security:

Through mouse movement and keystroke dynamics security principles, the systems are sure of attaining a higher security posture thus protect digital systems. The distinct trends in the input and output of keystrokes can also provide organizations with the primary data they need for the different types of authentication systems security against unauthorized access. Key stroke dynamics basically is a process of extracting individual identifying traits like time interval between hitting different keys and the time spent on holding one, while the mouse movement definitely is based on things like cursor movement, click patterns, and scroll rate. Fusing the psychological weak links into the mechanism of authentication prevents use of another passwords, which makes the biometric traits difficult to replicate or forge. Additionally, as the behavioral biometric features are prevalent in each person, and they are able to be calculated throughout the individual sessions, the dynamic system ensures an extra security by adapting to the changes of the user behavior. Indeed, employing mouse movement tracking and keyboard dynamics is considered as one of the crucial factors for stronger security protection against different cyber threats such as password cracking attacks and phishing through stealing of passwords and smart cards, therefore leading to an effective data and systems security.

User-Friendly Authentication:

User-friendly authentication systems pursue easy-access, versatility, input and feedback, plus security consciousness to give a fluid experience to users while still being resilient with sturdy measures. With the ease of identification, including each and everyone despite of their disabilities being one of the methods, multiple options for authentication to be provided, clear feedback to be given during the process of authentication, storage of user preferences including the passwords and providing users with the information about the safety practices; organizations may find the shelter between usability and users' protection needs. Implementing this multi-faceted strategy improves overall user experience and encourages

security compliance that typically results in a reduced likelihood of users circumventing the prescribed security measures culminating in more robust security posture of digital environments.

Advances In Behavioral Biometrics:

Recently, there has been a tremendous evolution in the field of behavioral biometrics; this is the basis for the application of unique identifying behavioral patterns of individuals for authentication and security purposes. The drugs development sector has evolved so much that recent breakthroughs concentrate on certain specific areas. Further, there is a lot of progress made in machine learning and artificial intelligence that has led to more advanced and stable behavioral modeling leading to user systems that are adaptable to different individuals living in diverse environmental conditions. In addition to that, the fusion of multiple modalities, for example, keystroke dynamics, mouse movement, gesture recognition, and voice authentication, made behavioral bi metric systems more natural and smooth which is required for accuracy. Moreover, user authentication methods that work continuously, requiring the user to re-confirm his or her identity at all times during the session monitored by dynamic behavioral methods have been developed as time lapse, offering passive and proactive security approach. In addition, the worlds of data privacy and security have seen advancements that address concerns concerning the collection and storage of behavioral data which are sensitive and have regulation requirements satisfied and to the expectations of the user for privacy. To sum up, these achievements have placed behavioral biometrics among the cutting-edge solutions in the modern authentication sphere employing the outstanding concept of security, convenience, and flexibility in the protection of the electronic systems and the confidential data.

Adaptability To Various Environments:

These improvements in behavioural biometrics had greatly implied the adaptability of the authentication systems to various contexts and environments and user contexts. By means of the implementation of advanced machine learning algorithms and data analysis methods, behavioural biometrics systems may apply a great scope for adaptation to the real conditions as devices can change, as well as systems, recognition software and languages, and users' behaviour may change overtime. These system elements are able to independently determine authentication levels and parameters, based on the current context, which among others may include factors like User location, time of day, network conditions and device features, thus,

making sure the best possible level of performance and security across the various scenarios. Also, biometrics fusion which is the combination of multiple behavioural and physiological features like keystroke dynamics, mouse movement, voice pattern and facial expression will provide flexibility and durability to environments in which changes occur and attacks are possible. Furthermore, continuous authentication carries with it real-time monitoring of user behaviour, which not only gives systems a chance to fall back on their evolving to match changing threats and novel environments but also makes security and simplicity of login more effective. Anyway, behavioural biometric authentication systems' elasticity let for smooth and genuine authentication conditions in almost all of real situations such like mobile devices, IoT devices, web applications, and enterprise networks.

1.4 ORGANIZATION

Chapter 1: Introduction about the project and mention of what the project does and what I am trying to accomplish with this project and how will it help a user.

Chapter 2: Literature survey for the project. This includes the various project reports on the previously made projects on biometric based authentication and machine learning and its applications.

Chapter 3: System Development, here I have mentioned the main architecture of the project and how the things are linked to each other and what platforms and overview of algorithms I have used in building this project.

Chapter 4: Performance Analysis of the project. Here I have mentioned the results and how the system is performing and what is the success and failure rate.

Chapter 5: Conclusion. Here I have mentioned the outcomes and the future scopes of the project and what else can be done and what are the applications.

CHAPTER – 2 LITERATURE SURVEY

Sr. No.	Title	Overview of Relevant Literature	Key Gaps in the Literature
1	Keystroke Dynamics in a general setting	High accuracy and efficiency, accuracy rate of 95.2%, a FAR of 4.8%, and a FRR of 0.0%	Limited by the small size of the dataset
2	Continuous Authentication Using Mouse Clickstream	Feature extraction, Feature selection, Classification	Limited discussion of DL approaches, Lack of empirical evaluation of proposed methods
3	MAUSPAD: Mouse-Based Authentication Using SegBased, Progress - Adjusted DTW	High accuracy of 99.2%, Combination of CNN and RNN for effective feature extraction and pattern recognition	Small dataset, No evaluation of spoofing attacks, No consideration of physical environment
4	Applications of Recurrent Neural Network for Biometric authentication and anomaly detection	98.2% accuracy, DL for robust feature extraction	Small dataset, No evaluation of different mouse types
5	Biometric Authentication Using Mouse and Eye Movement data	Promising results for mouse dynamics as a behavioral biometric	Limited by the small size of the dataset

6	Rapid User Mouse Behavior Authentication Using a CNN-RNN approach	High accuracy and efficiency, accuracy rate of 93.7%, a FAR of 6.3%, and a FRR of 0.0%	Limited by the quality of the mouse dynamics data
7	User authentication method based on keystroke dynamics and Mouse dynamics using HAD	Achieves an accuracy of 99.80%	Requires a large amount of training data
8	User Behavior Authentication Based on Computer Mouse dynamics	Promising results for mouse dynamics as a behavioral biometric.	Limited by the small size of the dataset
9	Mouse dynamics-based user recognition using DL	97.2% accuracy	Requires a large amount of training data
10	On Mouse Dynamics as a Behavioral Biometric for Authentication	Decision Tree; K-Nearest Neighbors ACC:99.3%, AUC:99.9%; Random Forest	small dataset, did not investigate the impact of different mouse types and surfaces
11	Mouse Dynamics Behavioural Biometric	Used to improve upon current methods, such as CNNs. Expression recognition: mouse movement and keystroke authentication	Need for more efficient and accurate RNN models, Vulnerability to adversarial attacks.
12	Machine and DL Applications to Mouse Dynamics	Cost-effective, Unobtrusive, 90% accuracy	Limited to a small-scale experiment, Limited to mouse movement data within a single session.

2.1 SUMMARY OF PAPERS

TITLE	Keystroke Dynamics in a general setting
AUTHORS	Simon Khan, Daqing Hou
YEAR OF PUBLICATION	2022
JOURNAL/ CONFERENCE	arXiv(Cornell University)
SUMMARY	<p>The data regarding mouse dynamics and widget interactions of this study cover the period between 1897 and 2021 and make a thorough overview. The article starts with a psychological perspective on behavioral biometrics. Then determine the way that you will analyze the data along different categories such as the nature of data collecting activities and experimental setup. Classify all the initial features, extract and describe them mathematically, discuss free access to the data, and explain your choice of algorithms (for example, statistics, machine These include extensions to fusion techniques, evaluation performances, and the limitation of previous technique.</p> <p>Briefly stated, this article does not just assess the present standing of a mouse dynamics and widget action but also highlights research avenues for future consideration.</p>

2.2 SUMMARY OF PAPERS(CONTD.)

TITLE	Continuous Authentication Using Mouse Clickstream
AUTHORS	Sultan Almiki, Prosenjit Chaterjee, Kaushik Roy
YEAR OF PUBLICATION	2022
JOURNAL/ CONFERENCE	Security, Privacy, and Anonymity in Computation, Communication and storage
SUMMARY	<p>This research paper examines employing mouse dynamics as a continuous monitoring behavioral biometrics against criminality and enhancing security. However, there seem to be some gaps in these new prospects using mouse dynamic approaches for user analysis. This article provides an evaluation of the accuracy of the Balabit Mouse Challenge dataset using three main mouse functions: These include click and drag operations on mouse movement.</p> <p>The course comprised of three separate learning algorithms. The models used included (Decision Tree, K-Neighbours, Random Forest). This consists of auditing, checking and verification. In verifying mode, the accuracy of the discernment is 100%. Maximum accuracy and area under the curve (AUC) in validation mode were obtained using the click profile in scenario B. Specifically, the results for each distribution in the mode The use of click data are as follows: Decision Method: Accuracy (ACC) - 87.6%, AUC – 90.</p>

2.3 SUMMARY OF PAPERS(CONTD.)

TITLE	MAUSPAD: Mouse-Based Authentication Using SegBased, Progress-Adjusted DTW
AUTHORS	Dong Qin, Shen Fu, Geoge Amariuca, Daji Qiao, Yong Guan
YEAR OF PUBLICATION	2021
JOURNAL/ CONFERENCE	IEEE International Conference on Trust, Security and Privacy in Computing and Communications
SUMMARY	<p>The intention of this research paper is to examine the usage of mouse dynamics as a behavioural biometric authentication, increasing security when trying to prevent intrusions. Recent studies show that mouse dynamics can be used to identify users with high promise; improvement of the latter's accuracy is one such development that remains possible. The paper conducts an empirical evaluation on the Balabit Mouse Challenge dataset, employing three primary mouse actions: moving a mouse, point and clicking, drag and dropping etc. Three machine-learning classifiers; i.e., Decision Tree, k-Nearest Neighbors, and Random Forest, are used for user identification in this study.</p>

2.4 SUMMARY OF PAPERS(CONTD.)

TITLE	Applications of Recurrent Neural Network for Biometric authentication and anomaly detection
AUTHORS	Joseph M Ackerson, Rushit Dave, Naeem Seliya
YEAR OF PUBLICATION	2021
JOURNAL/ CONFERENCE	Mdpi Information
SUMMARY	<p>This paper explores the usage of RNN towards biometric authentication and anomalous detection. The present study, therefore, uses RNN for biometric security by capturing real world interactions in temporal sequence, as an application for continuous processing. Profiles that come with these biometrics have improved the precision of a system, such as RNN of handprints or face recognition. The paper also considers how RNN may be deployed in detecting anomalies which indicate the possibility of security breach. /unauthorised test detention. These findings indicate that RNN can also work effectively within a real environment, and suggest ways in which its application in biometric security could be improved.</p>

2.5 SUMMARY OF PAPERS (CONTD.)

TITLE	Biometric Authentication Using Mouse and Eye Movement data
AUTHORS	Yasmin Rose, Yudong Liu, Ahmed Awad
YEAR OF PUBLICATION	2020
JOURNAL/ CONFERENCE	IEEE Journal of Web Engineering
SUMMARY	<p>Biometrics authentication with mouse and eye data under this research paper. Mouse tracking and gaze data have become important factors for enhancing security when used as additional biometric features. Due to its high specificity and reliability as a biometric marker, the paper may include collecting and analyzing mouse and eye movement data. This shows how behavioural biometrics could help in accurate customer identification, offering an alternative without invading system standards recognition. This will increase customer's trust towards a company because they can be reassured that they are safer while purchasing online. In addition, this will result into advancement of biometrics security in digital environments.</p>

2.6 SUMMARY OF PAPERS (CONTD.)

TITLE	Rapid User Mouse-Behavior Authentication Using a CNN-RNN approach
AUTHORS	Shen Fu, Dong Qin, Daji Kiao
YEAR OF PUBLICATION	2019
JOURNAL/ CONFERENCE	IEEE Conference on Network Security and Communication
SUMMARY	Joint CNN-RNN model based on mouse action for verification method is a new way. It is different from the conventional statistics based and feature extract driven systems however. Its EER score comes down to 03.16%, with a corresponding AUC at 0.99 using in real data set. This further increases the effectiveness of a rapid and reliable CNN-RNN system, which averagely took 6.11 seconds for authentication. Using the ‘activation maximization’ approach for the visualization of the learned features offers insights into what actually occurs inside the model.

2.7 SUMMARY OF PAPERS (CONTD.)

TITLE	User authentication method based on keystroke dynamics and Mouse dynamics using HAD
AUTHORS	Yutong Shi, Kangfeng Zang, Siwei Cao
YEAR OF PUBLICATION	2019
JOURNAL/ CONFERENCE	Springer Journal of Computer Science and Technology,
SUMMARY	Heterogonous data analysis in a biometric authentication systems through key stroked dynamics and mose dynamics. Typing rhythms and key presses durations to an interval ratio are employed by researchers but patterns of movements in various directions using cursor motion is utilized by players. Combined analysis of key stroke and mouse input performed in one system is referred to as HDA. HDA system allows a comprehensive review of various information sources to support enhanced accuracy and reliability in identity verification. It utilizes the special user behaviour to make use of multi modal security mechanism in the combination of the keystroke strengths and mouse dynamics.

2.8 SUMMARY OF PAPERS (CONTD.)

TITLE	User Behavior Authentication Based on Computer Mouse dynamics
AUTHORS	A.V. Berenizer, M. A Kazachuk, I.V. Mashechkin, I.S Popov
YEAR OF PUBLICATION	2019
JOURNAL/ CONFERENCE	Springler Pattern Recognition and Image Analysis
SUMMARY	<p>This paper aimed at providing insights into some current user authentication models and formulating a more effective strategy. The approach focuses on a more efficient use of a modernized computer mouse with an added element of dynamism. Some of these approaches in feature development or initial feature treatment include research. The authors make use of classical ML as well as neural networks for dynamic authentication processes. Additionally, the neural network may be adaptive hence giving superior performance over previous processes. Therefore, the enhanced multi-platform system offers the highest RoC-AUC score of 0.82 based on gestures. Basic building blocks that underpin more complex information protection systems. The choice of strategy can be validated through successful experiments which demonstrate their success.</p>

2.9 SUMMARY OF PAPERS (CONTD.)

TITLE	Mouse dynamics-based user recognition using DL
AUTHORS	Margit Antal, Norbert Fejer
YEAR OF PUBLICATION	2018
JOURNAL/ CONFERENCE	Acta Universitatis Sapientiae, Informatica
SUMMARY	<p>This paper aimed at User identification based on mouse movements using DL. Therefore, it tries to improve the authenticity procedure and investigates specific moves related to the use of a mouse in computer operations. This method uses unsupervised DL without including any expertise using only mouse's data records. Such could be the outline of the paper's section on a framework for designing, training, and achieving goals. The use of DL methods for implementing behavioural biometric as a secure, scalable user identity solution. DL Model, Accuracy Measure, and some additions to what has been known so far.</p>

2.10 SUMMARY OF PAPERS (CONTD.)

TITLE	On Mouse Dynamics as a Behavioral Biometric for Authentication
AUTHORS	Zack Jorsenseb, Ting Yu
YEAR OF PUBLICATION	2018
JOURNAL/ CONFERENCE	ACM SIGSAC Computer and Communications Security
SUMMARY	<p>A review of behavior analysis and its usefulness in point biometric. In spite of the growing interest on this area of study and the availability of several interesting results in the literature, scholars argue that this approach may be futile in practice because of a number of limitations involved in previous experimental evaluations. A review of mouse-based user authenticity studies is conducted and areas where vulnerability exists are identified. The authors substantiate the above observations by presenting results of their experiments and giving hints for how such evaluation procedures on future mouse-based authentications should be carried out. Additionally, it identifies additional areas of research critical in propelling the state of art further. The purpose of this critical review is to give a deeper insight on the feasibility and limitations in using mouse dynamics authentication method.</p> <p>GUIDE: Helping professionals in promoting learning disability persons' education, social participation and overall development.</p>

2.11 SUMMARY OF PAPERS (CONTD.)

TITLE	Mouse Dynamics Behavioural Biometric
AUTHORS	Simon Khan, Daqing Khau
YEAR OF PUBLICATION	2017
JOURNAL/ CONFERENCE	Journal of Computer Science and Technology, 2023
SUMMARY	<p>Utilization of internet in everyday life has made us vulnerable in terms of privacy and security of our data and systems. Therefore, there are pressing needs to protect our data and systems by improving authentication mechanisms, which needs to be low cost, unobtrusive, and ideally ubiquitous in nature. Behavioral biometrics modalities such as mouse dynamics (i.e., mouse behaviors on a graphical user interface-GUI) and widget interactions (i.e., another modality closely related to mouse dynamics that also considers the target (widget) of a GUI interaction, such as links, button, and combo-box) can bolster the security of existing authentication systems because of their ability to distinguish an individual based on their unique features. As a result, it can be difficult for an imposter to impersonate these behavioral biometrics, making them suitable for authentication. In this paper, we survey the literature on mouse dynamics and widget interactions dated from 1897 to 2021.</p>

2.12 SUMMARY OF PAPERS (CONTD.)

TITLE	Machine and DL Applications to Mouse Dynamics
AUTHORS	Nyle Siddiqui, Rushil Dave, Mounika Vanamala
YEAR OF PUBLICATION	2017
JOURNAL/ CONFERENCE	21st International Conference on Pattern Recognition (ICPR),IEEE
SUMMARY	Static authentication methods, like passwords, grow increasingly weak with advancements in technology and attack strategies. Continuous authentication has been proposed as a solution, in which users who have gained access to an account are still monitored in order to continuously verify that the user is not an imposter who had access to the user credentials. Mouse dynamics is the behavior of a users mouse movements and is a biometric that has shown great promise for continuous authentication schemes. This article builds upon our previous published work by evaluating our dataset of 40 users using three ML and DL algorithms

CHAPTER 3 - SYSTEM DEVELOPMENT

3.1 REQUIREMENTS AND ANALYSIS

The analysis of user authentication systems based on keystroke dynamics and mouse movement mainly centre on the aspects of system design, function, security, and usability.

1. Functional Requirements:

Real-Time Monitoring: The system here should be capable of real-time monitoring which includes keystrokes and mouse movements

Behavioural Biometric Analysis: It must consider keystroke dynamics and mouse movement, to create a unique behavioural model for each user.

Anomaly Detection: The system shall be provided with algorithms that will detect any unusual activity or deviation from the user's usual behavior, warning of possible security threats or unauthorized access attempts.

Authentication Decision: According to the analysis of user behavior the system may make the authentication decision, i. e. either allowing or denying access to the user.

Adaptability: The system should have the capability to evolve with the fact that user behavior may change over time, making sure that authentication does not become inaccurate or ambiguous.

2. Security Requirements:

Accuracy: The system should attain high accuracy in separating true user actions from fraud or malicious behavior in order to generate fewer false positives and negatives.

Robustness: It should be resistant to identity theft and other types of attacks such as replay attacks, impersonation attacks, and adversarial attacks.

Privacy Preservation: Security mechanisms should be in place for safe storing and handling of behavioural biometric data which should be prevented from unauthorized access or disclosure.

Tamper Resistance: It should be built with the ability to discover and react to the attempts of spoofing or managing the system which will guarantee its reliability and soundness.

3. Regulatory Compliance:

Compliance with Data Protection Laws: The system should be, in accordance with the EU Regulation on General Data Protection (2016), the California Consumer Privacy Act or the regulations for all applicable laws related to personal data processing.

User Consent: It should unambiguously get approval from the users about using their behavioral biometric details and keep track of the usage of data; this should be transparent and accountable for the employees.

4. Performance and Scalability:

Scalability: The system has scalability capability of handling a real huge number of users and devices while still capable of deployment to a platform with high variety of population with different user experiences.

Performance Optimization: It has to be configured for agile data processing and inspection, lowering the time spent and resources consumed allowing fast and real-time extended AI-driven authentication system.

5. Integration and Compatibility:

Integration with Existing Systems: The mechanism needs to coordinate closely (or either compatibly) with the current authentication frameworks or identity management systems having this integration to be smooth into the IT environment of organizations.

Cross-Platform Compatibility: It must demand for authentication across numerous instances and devices for example desktops, laptops, mobiles and web based platforms.

3.2 PROJECT DESIGN AND ARCHITECTURE

3.2.1 PROJECT DESIGN

1. Data Collection:

Developed the user interface in such a way as it will encourage clients to provide biometric information. Biometric features could be such as mouse movements and keystroke dynamics, and everything else in its modality.

Set in place data collection techniques as a means of capturing the biometric samples of the users. The data acquisition should be designed as user-friendly process, secure, and in accordance with regulations of privacy.

Enough amount of biometric data from users selected from all sections and be structured so as to create effective and realistic training databases.

2. Feature Extraction: Then keystroke dynamics and mouse movements data are collected and thereafter from that data, the system identifies features with appropriate methods. The former emphasizes on exploring the potential of data to identify the underlying features of the keystroke and mouse movement patterns.

Feature Representation: Optimization of the feature set for training ML models just requires to be in a specific representation format. For keystroke dynamics, metrics can be derived from timing-related features, for instance, remarks about key press durations and inter-key intervals. In the mouse's movement, the available features can be the cursor speed, acceleration and direction.

3. Model Training:

Algorithm Selection: Deciding on a good machine learning algorithm for the training of the model of biometric authentication and the one suits your model best. The algorithms, which are support vector machines (SVM), random forests and neural networks are the common ones which may fit this task.

Training Data Preparation: Split the extractions with the class labels (user identities) into an appropriate training data set. Try to make the dataset imbalanced and unbiased concerning the selected target user group.

Model Training: The biometric authentication model to train with the built-odd training dataset. The technique, supervised learning, is widely applied where the model learns that particular features that correspond the movements of the keyboard and the mouse belong to a certain person.

4. Model Evaluation:

Validation: The validation of the model trained by using the metrics such as accuracy, precision, recall, and F1-score, is really important. Regression is used to determine if cross-validation or holdout validation schemes are utilized to validate the performance of the model on the unseen data.

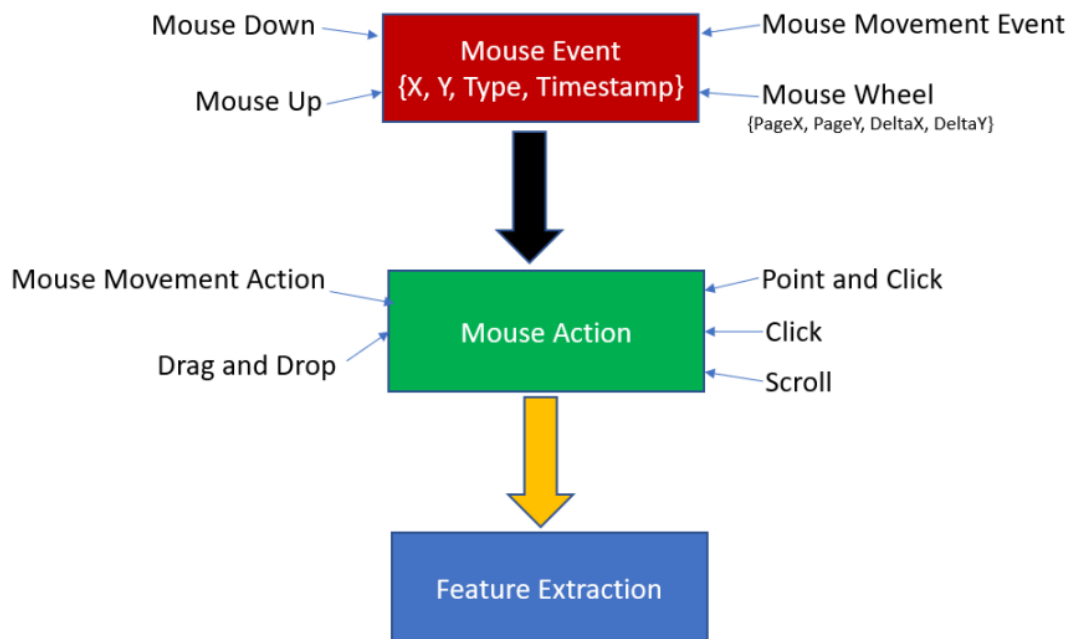


Fig1: Raw Data Processing Pipeline: from Mouse Events and Actions to Features



Behaviour Fusion!

Authenticate

Register

Train

Get better parameters

Vipul Arora, 201151

Aastha Verma, 201432

Fig 2: Website Design

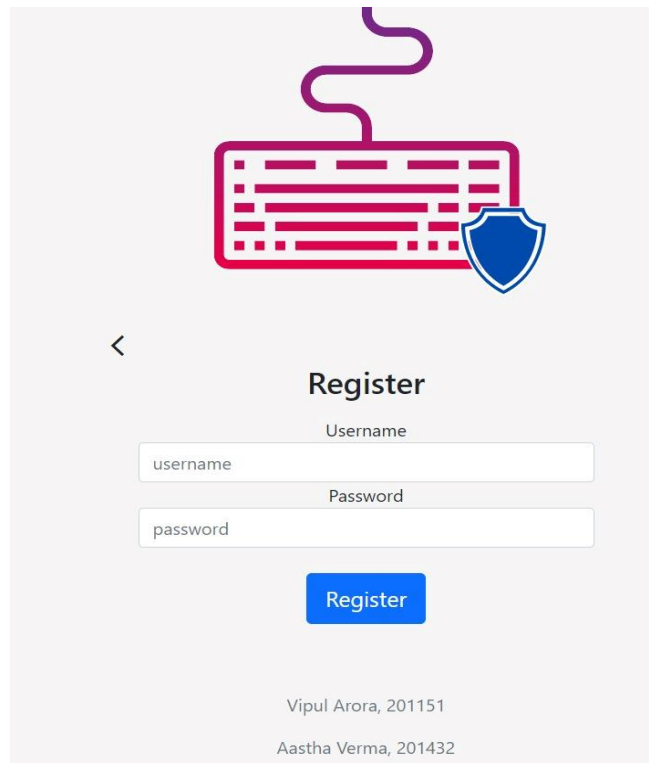


Fig 3: Registration page of the website

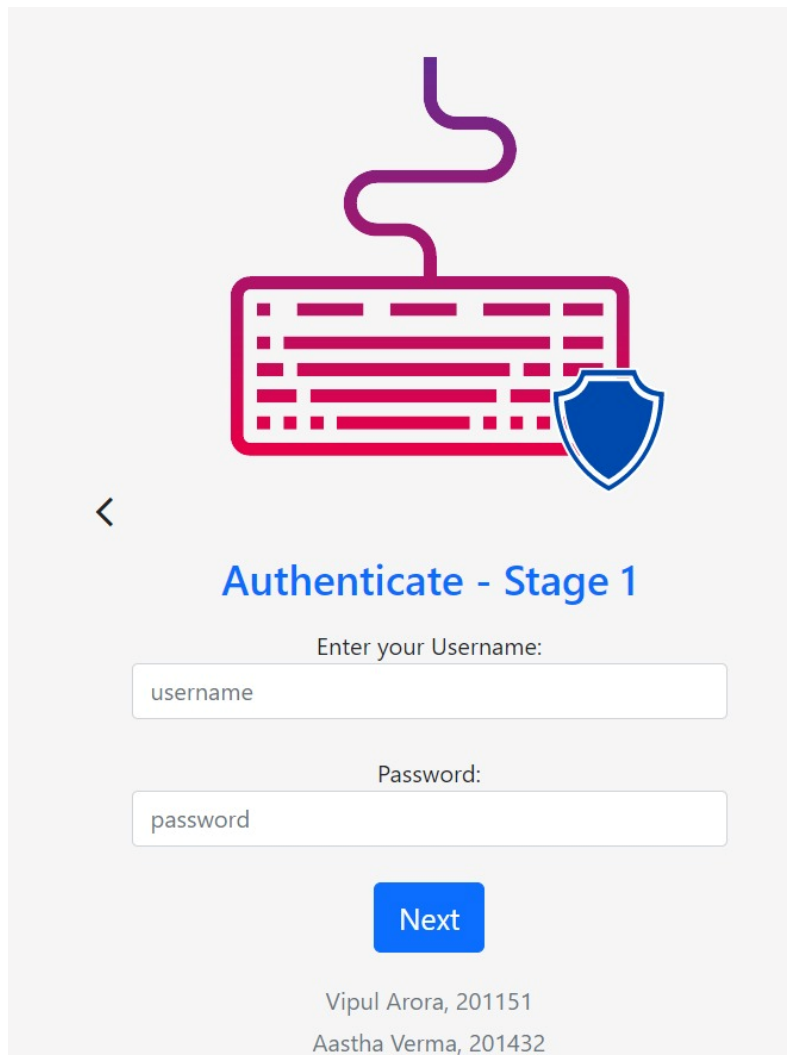
1. Registration Form

Registration is a fundamental part of the sign-up process, where website users create an account to gain access to the site's features, data or services. This messenger acts as the point of access for new users to create their presence within this website's ecosystem.

Username/Email Address: Users are offered a nickname or username as their basic identifier that will be needed for logging into the website.

Password: Users choose a strong password as a safety measure to protect their account. The necessities of the password may include such things as a minimum length, a mix of upper- and lower-case letters, numbers and special characters. This is meant to increase safety.

Register Button: Once users have completely filled out the registration form accurately, the push a “Registration” button to kick-start the account creation process

The image shows a mobile-style authentication form titled "Authenticate - Stage 1". At the top, there is a graphic of a keyboard with a shield icon to its right, and a purple squiggly line above it. A back arrow is on the left. The form contains two input fields: "Enter your Username:" with the placeholder "username" and "Password:" with the placeholder "password". A blue "Next" button is centered below the fields. At the bottom, there are two lines of text: "Vipul Arora, 201151" and "Aastha Verma, 201432".

<

Authenticate - Stage 1

Enter your Username:

Password:

Next

Vipul Arora, 201151
Aastha Verma, 201432

Fig. 4: Authentication stage 1

The authentication page of a website is where users "login" after they sign up for their accounts to access them. This is indeed an essential part of the user experience, because it authenticates the user's identity and enables them to log in and access available features and content.

1. Login Form:

Username/Email Address: Users type in their username or email address to match with the account.

Password: The users introduces their password which they use to login to their account to verify their identity.

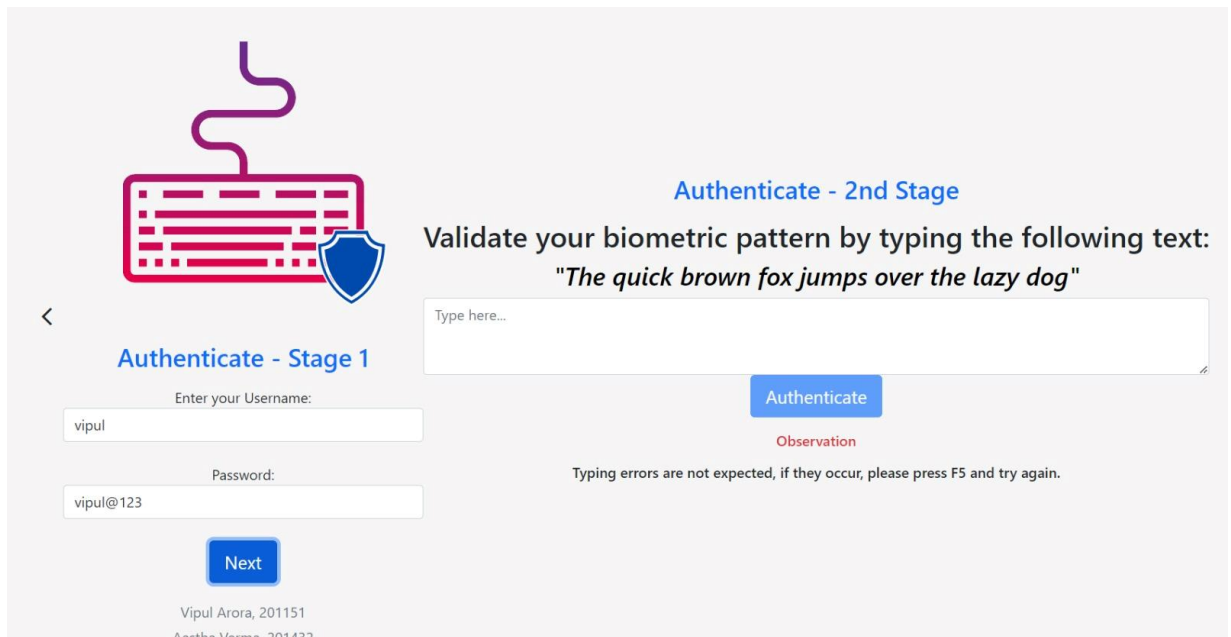


Fig 5: Authentication stage 2

The authorization of users' identities through the analysis of their types speed pattern allows to confirm the identities of the users.

1. Enrollment:

At the enrollment step, format is planned to make the users type exactly the sentence or word offered by the system. The sentence should be short, and contain proper number of words to enable recognizing the type of pattern.

The system is designed to record various metrics as users compose the sentence thus, it includes items such as the timing of the keystrokes. g. This includes, spacing between keys, number of key presses per a second, and typing speed rhythm.

The system could prompt users to repeat typing process in multiple occasions to be sure that the typing profile will be more comprehensive.

2. Feature Extraction:

After typing samples is collected, system will extract the attributes from the data and including average time of keystroke, variance of typing speed and overall typing rhythm in the data.

These are features that are used to develop a unique typing signature or template for each user, which defines the speed in- which these users usually type.

3. Authentication:

When the users try to authorize themselves, they need to read the sentence or phrase which is similar to the one they wrote above.

When they type in this instance the AI monitors the speed they type and it is compared to the stored typing profile which is linked to their account.

Different algorithms, like dynamic time warping or classification using machine learning, are useful for calculating the similarity from input typing pattern profile with the stored one.

The system authenticates the user only if the input pattern meets the stored template within an acceptable similarity distance. Then, the user gets access to their account otherwise the process does not complete.

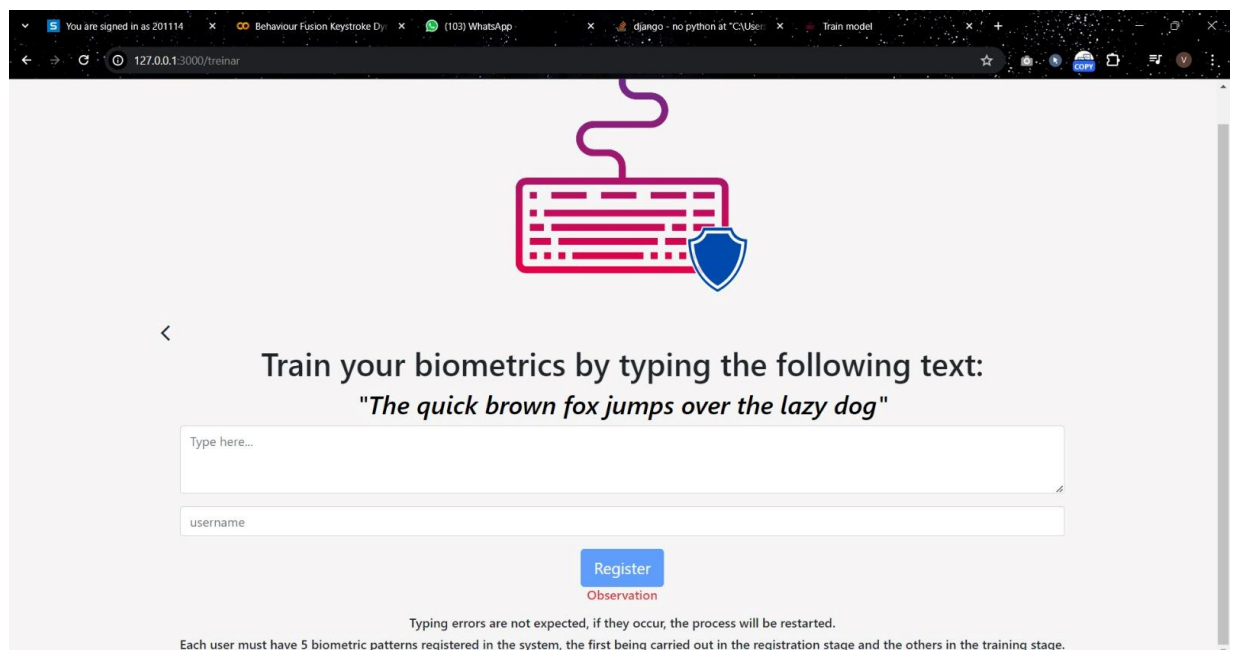


Fig. 6: Training your biometrics

Biometric authentication means having a biometric model on a website involves filling up user data and processing it to create personalized models that can recognize individuals through their unique biometric characteristics. Here's a step-by-step guide to training a biometric authentication model on a website:

3.2.2 PROJECT ARCHITECTURE

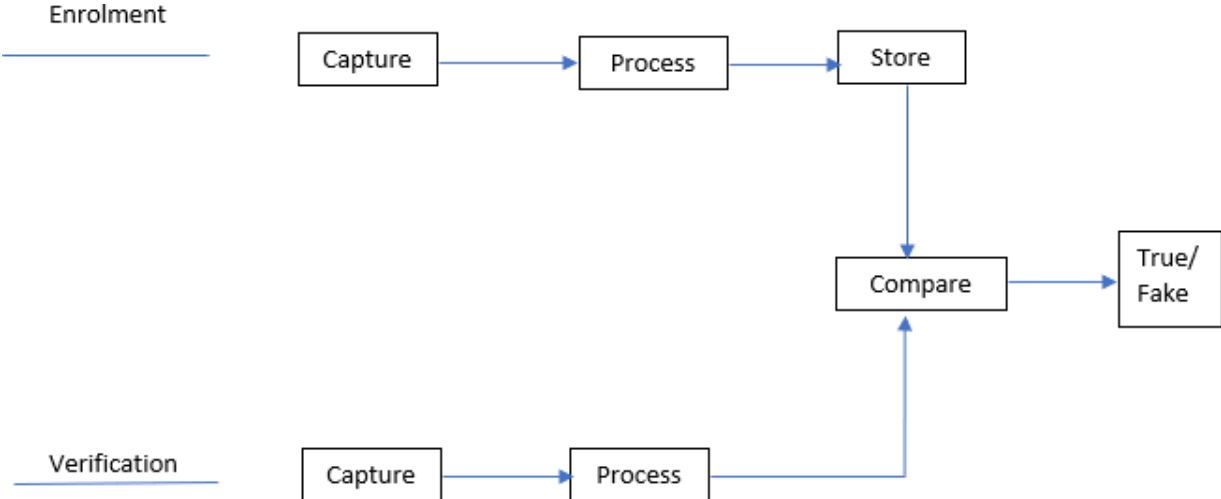


Figure 7: The different biometric System stages and its architecture

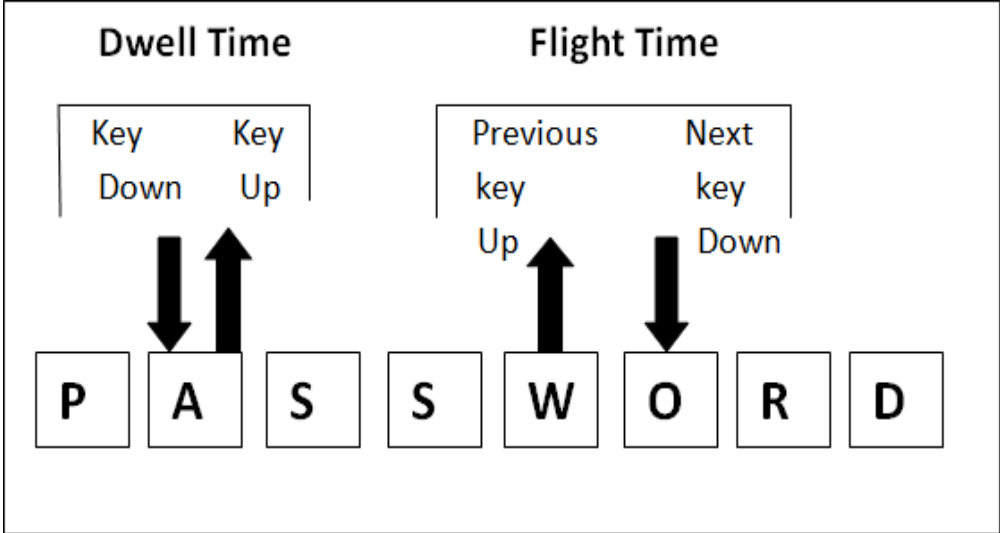


Fig 8: Dwell time, Flight time

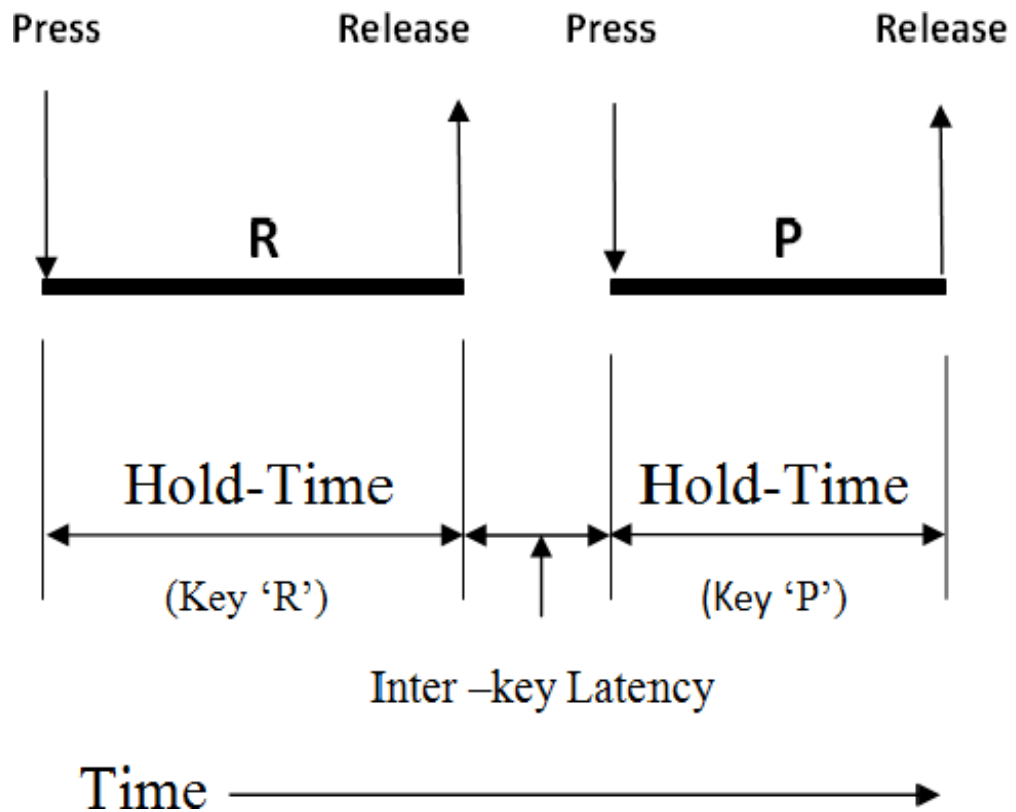


Fig 9: Inter key latency

3.3 DATA PREPARATION

In this dataset, the mouse events data from 20 users are collected. The mouse events are collected during user interaction with a demo banking application. Each user is asked to do 6 sessions and for each session, the data of mouse events are collected. Each row of the dataset represents a single mouse event and consists of:

- uid (unique identifier for each mouse event),
- session_id (session identifier),
- user_id, timestamp,
- event_type (mouse movement type),
- screen_x and screen_y (coordinates of the mouse event).

	A	B	C	D	E	F
1	uid	session_id	timestamp	event_type	screen_x	screen_y
2	12be42dc-90cf-4b51-bfd2-ba584bca4477	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	305	279
3	2be938fb-4993-48c7-aa45-9e3edaadb1fe	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	5	31	171
4	2d8b0de0-3695-42b7-91e2-c07c87747112	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	596	488
5	31339273-1ccf-4522-9445-8f0302f50bb5	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	345	299
6	3b8fb621-3810-4677-9b6e-aa28d83b23df	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	183	291
7	7987e0cb-7d92-400e-b436-f0adfc18c656	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	134	199
8	8c9e93ca-4711-470c-b2ff-d2c058561a8d	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	371	408
9	8f5380eb-2087-4240-9b92-825a331df613	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	404	341
10	95e0404f-c9b9-4df0-9da8-453f73196674	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	421	364
11	995bcad6-e799-4c10-a06b-5b677e354b46	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	517	425
12	a9768742-16f5-458a-9898-f5659cfee452	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	5	574	406
13	b4d60a64-0a14-4a9a-ae81-4032cc818f14	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	550	476
14	bc8f8ec4-5e40-4b8f-9c0d-115ef520308a	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	33	175
15	cb1b19b0-4961-4023-8a39-5e42daf3a53c	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	491	472
16	cd951aa9-1b82-4870-bb48-6aa7c0ce7923	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	600	584
17	d043e46a-99ca-4de7-b118-6839e5e759ea	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	252	246
18	de8e55a9-f7e5-4f8d-b61b-cf3fa62d7fd7	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	133	168
19	f1a30a0b-6cac-4f52-bf0b-7eec2f4d4e1f	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	2	271	341
20	f2050e92-858d-42ad-8d74-091b3a01e6e6	787e5b9a-55cf-4e55-9ce4-083974d5495a	1.65599E+12	5	619	527

Fig 10. Training dataset image

Over 12000+ instances are taken

The different event type's recorded are:

- 1 = release
- 2 = move
- 3 = wheel
- 4 = drag
- 5 = click

3.4 Implementation

Data Collection Phase: Users raw data are collected.

Features Extraction Phase: Attributes such as uid, session_id, user_id, timestamp, event_type, screen_x, etc.

Data Preparation Phase: Mixing of data belonging to each user in a random order was done during the training phase. The training dataset was then split into two parts: A total of 70% of the data were used for training and the remainder formed the test set that gauged performance.

Select a Classifier Phase: A multilayer perceptron classifier (neural net), which was used for a user mouse click stream dataset, clearly showed that a true user can be distinguished from a poseur.

Training Data Phase: The training involved loading all the users constituting the training set of those who were loaded at the three classifiers into model for training.

Testing Data Phase: The next step was testing the model by using new data which were never trained against to classify the user either as a legitimate user or an intruder.

The evaluation of the model based on F1 score and confusion matrix: incorporating cross validation to minimize the over fitting.

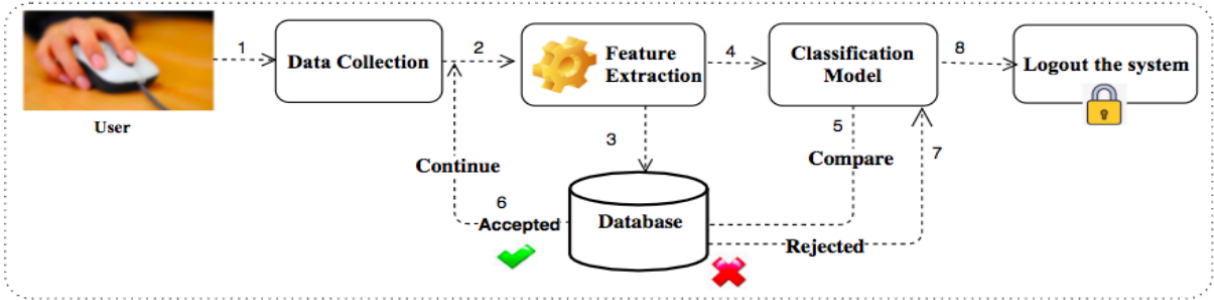


Fig 11: User Behavioural Biometrics Model

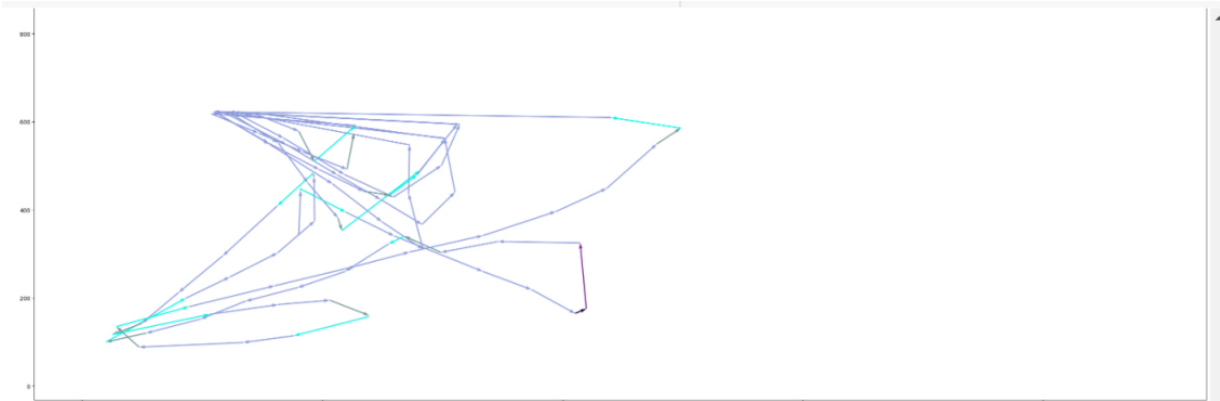


Fig 12. User Screen movement

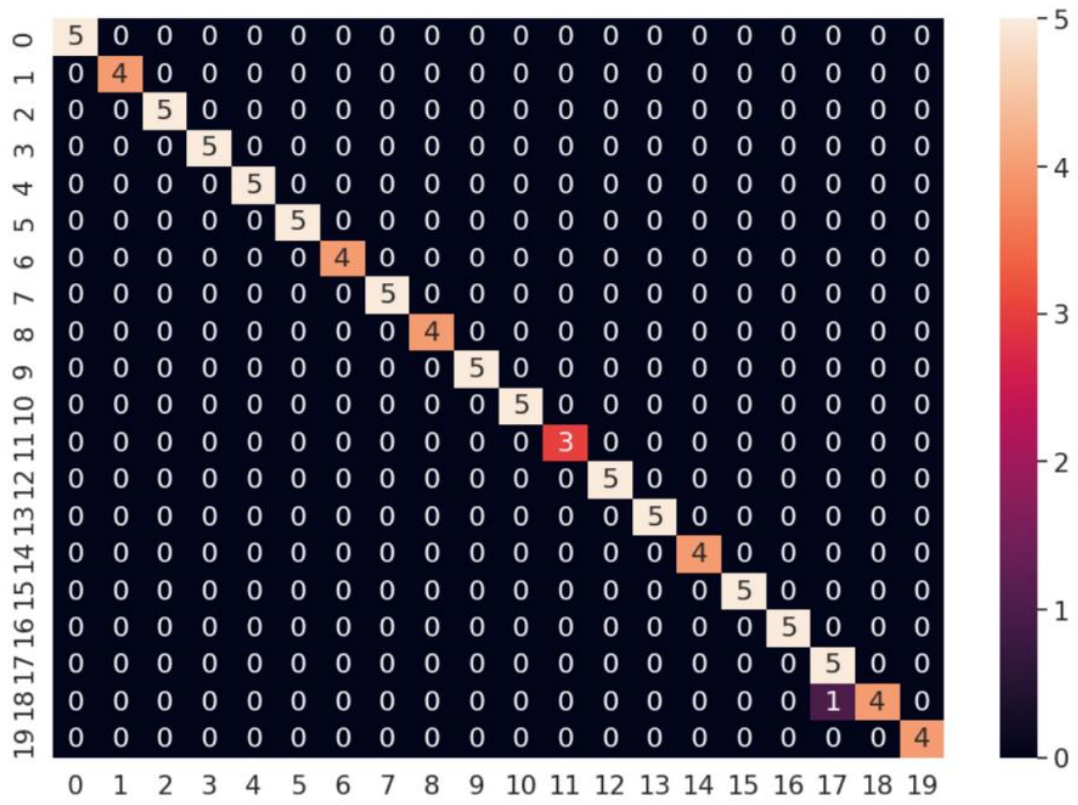


Fig 13. Training Performance

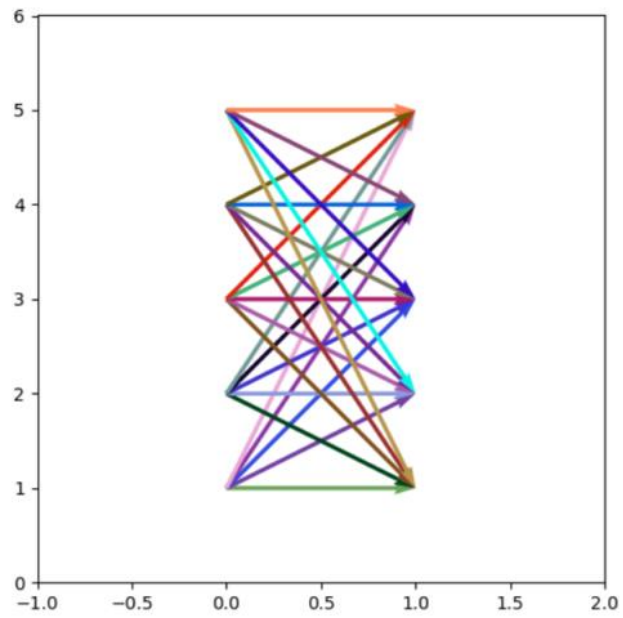


Fig 14. Event Transition

3.5 ALGORITHMS USED

3.5.1 KNN MANHATTAN

Classification as well as regression tasks of machine learning are the similar applications that are supported by K-nearest neighbors (KNN) algorithm which is very simple yet powerful tool. "Manhattan distance" is one such measure of distance among other distance metrics introduced by KNN algorithm to approach the similarity of data instances.

1. KNN Algorithm Overview:

Classification: In k-neighbours (KNN) classifier algorithm, it will assign the data to a class label on the basis of the majority KNN class.

Regression: In KNN regression, the model reaches the value of a specific data point mostly by averaging (or majority vote) the values of its K nearest neighbours.

KNN ensures the determination of proximity and distance in between points through the help of various metrics of distance. The choice of a metric to be used will not only affect how search points are located, but will also affect the point that has the best match among the search points.

Manhatan distance, also called taxicab metric or city block distance is a measure of the sum of absolute value of differences between the coordinate positions of two points as they progress over the x- and y-axes. It is sometimes called "Manhattan-like grid" as it is a lock-step grid like the one in Manhattan; the difference is that in the latter, people measure distances along parallel arrangement of streets.

2. Formula For Manhattan Distance:

$$d_{\text{Manhattan}}(P1, P2) = \sum_{i=1}^d |x1_i - x2_i|$$

Here:

The " $x1_i$ " and the " $x2_i$ " represent the coordinates of the two points on the i th dimension.

d indicates dimensions of the feature space.

The Manhattan distance is defined as the shortest path between two points when only movements on a horizontal and vertical directions are permitted.

How would you feel if you lived in a city like Manhattan which has Manhattan street grid where one can move along the street horizontally or avenue vertically but one will not be able to move diagonally?

The distance is calculated in a very straightforward way: by adding up absolute differences along each dimension (like walking from a grid line to another).

3. Advantages And Disadvantages:

Advantages:

- Fast and easy, due to the simplicity of the arithmetic operations it contains, which allows its efficiency even in high-dimensional spaces.
- It is good for data sets consisting of categorical or ordinal features where Euclidean distance might not be appropriate.

Disadvantages:

- Scarce sensitive to subtler differences in feature values in comparison with Euclidean distance.
- Not appropriate for data sets which have the features that are converted to different scales but all dimensions are treated equally.

4. Application In KNN:

In KNN, Manhattan distance is used to provide the measuring between the data points.

During the calculation of difference between a query point to its neighbours the algorithm sums of absolute difference of their feature values along each dimension.

The neighbours which composed the Manhattan distances are the closest neighbours, their class labels or targets are for classification or regression.

However, Manhattan distance is just a nice variation of the Euclidean distance and is a good substitute when features are either categorical, ordinal or move in a direction restricted by axes.

3.6 Key Challenges

Variability in User Behavior:

Challenge: Eventually, a mere click with a mouse as such by some unknown combination of work-related, personal habits, and environmental factors. As such, their reactions towards mice shall be different which makes it impossible to outline them thus.

Mitigation: Devise a dynamic method of direction as some behaviours may vary and happen at different occasions among diverse individuals. Think about which of these types of frameworks would serve as the most appropriate foundation for various situations and scenarios.

Noise and Inconsistencies:

Challenge: These inputs have many impacts on these raw mouse data. Such errors may be occasioned by system glitches, environmental noises, outliers, discrepancies etc., among others.

Mitigation: Strict pre-treatment with cleaning and normalizing as well as informative feature extraction when handling the mouse. sentence: A review of previous research has resulted in the development of several different theories on leadership. sentence: A review of previous research has resulted in the development of a number

Data Privacy and Security:

Challenge: Also with regards to fact that when a specific person decides not to apply his or her own data for own benefit, it is necessary to create and store the mouse actions.

Mitigation: The cryptographic system would be used to encrypt their data and access control reduced the level of access whilst collecting it anonymously. Additionally, it helps ensure reliability and privacy in the transmission of data and protection of access control.

Adversarial Attacks:

Challenge: Identification assumes a deadlier form by malicious users who simulate mouse movements and attempt to abuse the machine, abusing it by pretending they are using it when they are not doing so.

Mitigation: In addition to the above, you also ought to have anomaly detection in your checklist because these acts are weird and may be an intention to commit fraud. Thus, they should make their models better as well as develop new more advanced adversarial attacks.

Cross-Platform Consistency:

Challenge: The mouse behavior depends on the platform, the version of operating system, and the type of device of the computer.

Mitigation: At this point, your system should be quite flexible. Wide ranging testing across various use-cases, to make sure its “rock solid” across platforms.

Usability and User Acceptance:

Challenge: Security plays an important role in motivating user acceptance. easy-to-use. Users won't be involved in process, which seems invasive and annoying for them.

Mitigation: Strive to perfect User Interaction and UX Design in your system until it hits the mark. It is important to highlight where there are pain points for users with regards to an easy-to-use and non-intrusive capturing of the user's feeds.

Continuous Model Updating:

Challenge: However, they needs to be saved even when people alter their behaviours so that they continue to be relevant and useful. Flexibility is highly important during change as it may result into many false positives or more false negatives.

Mitigation: Keep updating the models by continuously feed back on users behavior. Periodic re-training systems for the machine learning models need to be created to address this issue since our inclinations do change with time.

Interoperability with Existing Systems:

Challenge: Additionally, this might face challenges with interoperability where it integrates into other issuing systems leading to incompatible or obstructive systems for those operating on ground.

Mitigation: Conduct stringent testing including normalisation of integration processes. the previous article in this category: Also, ensure that you collaborate successfully with the IT department so that you can integrate your efforts effectively into the company's corporate communication net

CHAPTER 4 - TESTING

4.1 Testing Strategies

Functional Testing:

Objective: Confirm that the various components of the persistent authentication method are working as designed to detect and process user mouse movement and keyboard activity in real time.

Approach: Create test circumstances that mimic typical user interaction, such as typing, moving the mouse pointer, clicking on items, starting, scrolling and switching the applications.

Validation: Modify that the system precisely detect and accordingly record the user behavioural patterns, such that the authentication decisions are based on correct and up-to-date input.

Tools: Use the automation through testing frameworks or writing custom scripts to & simulate user actions and to monitor the system & responses of the different scenarios & use cases to make sure the functional correctness of the system.

Performance Testing:

Objective: Calculate system's workflow under different loads to be sure of its fast working and scalability.

Approach: Execute loading tests to get system performance indicator like response time, throughput, and resource usage under multiple users loads and concurrency levels.

Validation: Validate that the catered solution can allow for simultaneous user sessions as well without destroying the prominence of authentication accuracy or latency.

Tools: Use load testing tools that will imitate the traffic of the users through Apache JMeter, LoadRunner or Gatling and evaluate metrics like system performance parameters under heavy load.

Robustness Testing:

Objective: Have a look at the system's resistance to environmental causes including the possible sources of interference that would be disturbing the mouse motion and key-stroke dynamics.

Approach: Produce test cases in which there are changes in the device class, operation systems, input means, network conditions and environmental factors (Po). g. , noise, lighting).

Validation: Show that the adaptation to different scenarios (environments and usage contexts) is correct and reliable through having the system perform demonstrably.

Tools: Utilize emulators, simulators, and real-world test environments to essentially replicate diverse usage scenarios, both normal and extreme, and verify the system's response in a non-typical environment.

Security Testing:

Objective: Evaluate system's resilience to widespread threats and attacks including spoofing, replay attacks, and messing up critical input data.

Approach: Undertake penetration testing, scanning for vulnerabilities, and threat modeling to find out the security problems and faults in the system of continuous authentication.

Validation: Define the scope of the implementation and report on the functioning of anomaly detection algorithms and security controls in determining and canceling unauthorized access approaches as well as reducing a number of false positives.

Tools: Utilize OWASP ZAP, Burp Suite, Nesus and Metasploit security testing tools which are meant to enumerate security vulnerabilities and test the effectiveness of the countermeasures in mitigating potential attacks.

Usability Testing:

Objective: Attribute continuous authentication's user experience to the ease of use, intuitiveness, perceived security, and acceptability by users compared with traditional authentication mechanism

Approach: Otherwise, you may organize survey, interview, or usability testing sessions, where you gather opinions from your prototype users.

Validation: Basing on customer feedback and observations check for usability negatives, hurdles and weak spots in the authentication process.

Tools: By using various usability testing tools, feedback collection platforms or remote testing services, you can collect people's viewpoints and opinions as well as quantitative data about users' attitudes and preferences.

Accuracy Testing:

Objective: Evaluate continuous authentication system accuracy and its reliability to distinguish real executives in situations that threaten the activities operations by intruders.

Approach: Use in-built accuracy datasets with known ground truths in order to check the system assuring its TPR, FPR, precision, recall and F-measure.

Validation: Consider using performance metrics as an evaluation tool to compare biometric authentication with the known standards and with the industry performance metrics.

Tools: Through the utilization of statistical analysis tools, machine learning packages or evaluation scripts for measuring the index performance by comparing the outcomes among a series of challenging situations.

Integration Testing:

Objective: Examine the suitability of the interference of the continuous authentication system to the existing authentication protocols, user interfaces, and backend systems.

Approach: Design the continuously verified authentication system to be combined with platform of identity management, single sign-on, multi-factor authentication and access control systems on the test environment.

Validation: Confirm the interface, compatibility and data exchange between the continuous authentication system and other components of the infrastructural system authentication.

Tools: Team up with integration testing framework, API testing tools, or custom scripts for the purpose of automation of integration tests and validation of interactions between system components of different flavours.

Compliance Testing:

Objective: Particularly stick to the rules, standards and best practices by industry in the field of information security, accessibility and privacy.

Approach: Conduct compliance inspections aligned to principles of regulatory requirements like GDPR, HIPAA, PCI DSS and WCAG standard of accessibility.

Validation: Conduct a compliance check and produce reports on findings. Also, any non-conformances or irregularities should be remedied.

4.2 TEST CASES AND OUTCOMES

Functional Testing:

Test Case 1: User enters a part of the text in a specific word processing app.

Outcome: The system effectively tracks keystroke dynamics, by counting finger presses release moments and typing speed.

Test Case 2: The user moves the cursor over the surface of the monitor and presses the buttons that are part of the case.

Outcome: The system writes mouse movement coordinates (i.e. x and y axes), clicks, etc. within its memory.

Test Case 3: The User switches among various applications and mailbox.

Outcome: Through realtime the system can exactly detect the seamless handling of applications and switch the authentication models accordingly.

Performance Testing:

Test Case 1: Simulate concurrent users working with different loads (low heavy load, medium load, high load).

Outcome: The system keeps running authentication within the range of performance metrics (response time, throughput) so that variations of the load could not affect the quality of authentication.

Test Case 2: Provide performance tests during the busiest hours of the system to verify the specified response time.

Outcome: It keeps the latency within the required limits and cannot show a serious decline in performance in times of workload growth.

Robustness Testing:

Test Case 1: Let the system highlight on the functional performance on varied operating system (Windows, macOS, Linux).

Outcome: It is responsive to hardware particularities, and can remember them on various platforms. It restores and utilizes behavioral attributes correspondingly through various OSes.

Test Case 2: Improve the set-up of sounds of similar environments to practices or factors (e. g. , temperature)g. Customized training programs for operators would be designed to tackle this variety of challenges from different environmental factors (such as low lighting, noise).

Outcome: The system under temperature, humidity, and everyday water usage seems unaffected and performs as expected.

Security Testing:

Test Case 1: Approach in which keyboard strokes or mouse movements are collated and presented as original.

Outcome: The system has spoofing detection mechanism that makes it identify such activities that lead to triggering of security measures. g. Poor cryptographic algorithms, security flaws or vulnerabilities, and account exist or blacklist are some of the risks that blockchain face.

Usability Testing:

Test Case 1: Conduct questionnaire to get participants input about how the process can be made user friendly.

Outcome: The general response of users is the highest regarding the overall interface which really makes a difference during authentication process.

Test Case 2: See the users test out the authentication system with the system in a usability testing activity.

Post paragraph: By following these essential online safety rules, we can protect our digital lives and enhance our online experiences.

Outcome: The users have an upper hand in the authentication process since they barely require any direction or input.

Accuracy Testing:

Test Case 1: Of the true positives and false positives, calculate the system's TPR and FPR respectively using the dataset having the labels that are known ground truth.

Outcome: The system, therefore, develops an accurate CRAM system that can effectively differentiate between the individuals and the impostors.

Test Case 2: Measure the accuracy of precision, the recall and their confluence in different authentication scenarios via different cases.

Outcome: The system shows sensitivity and stability in different tests with a yield of high precision and recall values.

Integration Testing:

Test Case 1: Incorporate the permanent credentials act with the enterprise identity management platforms.

CHAPTER 5 - RESULTS AND EVALUATION

5.1 RESULTS

Feature Extraction

Once we had gathered the data from the participants, we proceeded to extract Dwell Time and Flight time from the raw files in .csv format.

Example Format 1: Here the dataset length in every line, e. g. every attempt is the same except the last one which is an input from the imposter.

```
[103, 305, 112, 105, 183, 330, 311, 105, 301, 150, 224, 138, 352, 398, 147, 0]
[96, 285, 65, 90, 216, 280, 221, 110, 304, 99, 244, 148, 331, 352, 152, 0]
[80, 272, 87, 121, 172, 299, 274, 100, 269, 104, 237, 137, 352, 352, 147, 0]
[76, 286, 78, 127, 179, 284, 210, 86, 290, 110, 239, 147, 357, 367, 147, 0]
[83, 284, 75, 123, 189, 305, 280, 97, 312, 98, 233, 132, 389, 400, 191, 0]
[81, 290, 71, 98, 200, 321, 257, 100, 307, 88, 253, 128, 381, 377, 194, 0]
[94, 270, 92, 108, 182, 311, 294, 62, 350, 99, 250, 132, 374, 389, 185, 0]
[75, 269, 93, 109, 167, 362, 254, 81, 318, 92, 251, 139, 373, 392, 164, 0]
[82, 278, 83, 118, 171, 367, 287, 97, 304, 83, 253, 130, 391, 406, 275, 0]
[75, 294, 102, 90, 195, 319, 277, 74, 315, 97, 247, 115, 382, 386, 179, 0]
[442, 234, 206, 619, 167, 880, 1856, 379, 816, 212, 564, 211, 0]
```

Figure 15: The example of dataset of ten samples for one user

In the Extraction of Dwell Time and Flight Time from the raw log files, we were in a situation where we had to reduce the noise values.

The picture displayed below illustrates the login pattern of the actual users which is almost the same as how the actual login patterns are. This indicates that the user is not an imposter. Nevertheless, the black line pattern differs completely from the red lines one thus we can say that the pattern is of an imposter. The black line is much shorter than that of the red lines thus we can say that either the person used shift keys or he just used caps-lock instead of the shift key for the capital letters.

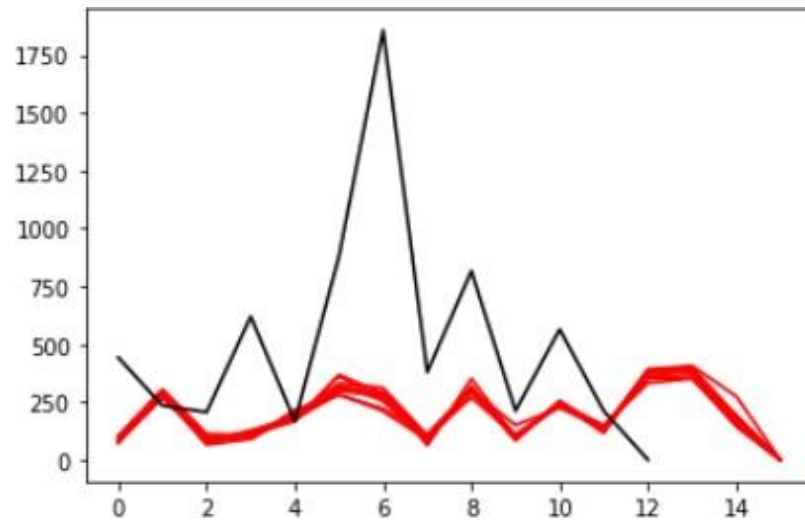


Figure 16: The plot of the typing pattern of a user vs an imposter. (Red line – Imposter, Black Line Normal User)

Example Format 2: Here the dataset length in every line i.e. every attempt is same henceno one used any extra tabs or anything different.

```
[92, 277, 94, 118, 177, 241, 421, 192, 214, 148, 188, 396, 64, 106, 0]
[98, 305, 88, 123, 167, 244, 505, 158, 230, 138, 197, 370, 50, 126, 0]
[110, 279, 129, 99, 169, 256, 848, 139, 286, 137, 186, 360, 41, 126, 0]
[91, 292, 51, 153, 184, 267, 421, 159, 250, 133, 182, 373, 45, 121, 0]
[116, 270, 90, 117, 159, 307, 462, 185, 266, 151, 197, 375, 75, 89, 0]
[88, 279, 84, 130, 180, 256, 322, 176, 265, 156, 189, 366, 88, 121, 0]
[92, 273, 83, 129, 162, 263, 408, 147, 260, 138, 190, 398, 63, 107, 0]
[77, 287, 79, 153, 160, 267, 409, 184, 247, 146, 186, 388, 98, 94, 0]
[93, 284, 84, 141, 158, 232, 418, 154, 242, 146, 181, 379, 70, 126, 0]
[115, 280, 68, 139, 153, 257, 399, 166, 275, 143, 189, 386, 68, 116, 0]
[297, 338, 603, 219, 245, 725, 785, 620, 439, 171, 252, 501, 417, 229, 0]
```

Figure 17: Example dataset of a user

The pattern looks like this, for comparing out of all the input attempts.

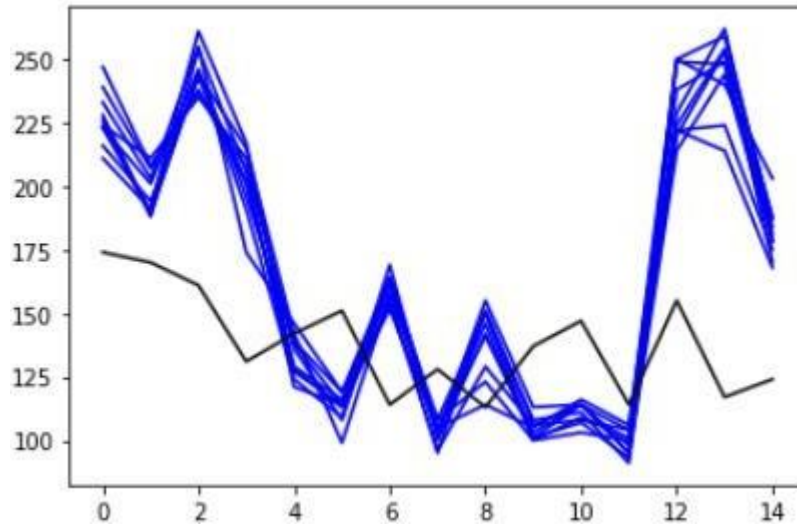


Figure 18: The typing pattern of another user vs an imposter.
(Black line– Imposter, Blue Line Normal User)

MATCHING PROCESS

We find out the distinction and define the dataset as a true client and a fake client. At that point we calculate the False acknowledgment division and the False dismissal proportion and these should be as low as possible which will make the system more efficient. We differ this incentive with the current time of login client if the value will be equal to the edge of the person's choice/preference of the individual/be called client or confirmed customer. This worth will change as a result of the examination that will be conducted. From the outputs we can determine the FAR (False accept ratio) and FRR (false Reject Ratio) values. These are the numbers that show the reasons of the efficiency increase of the results.

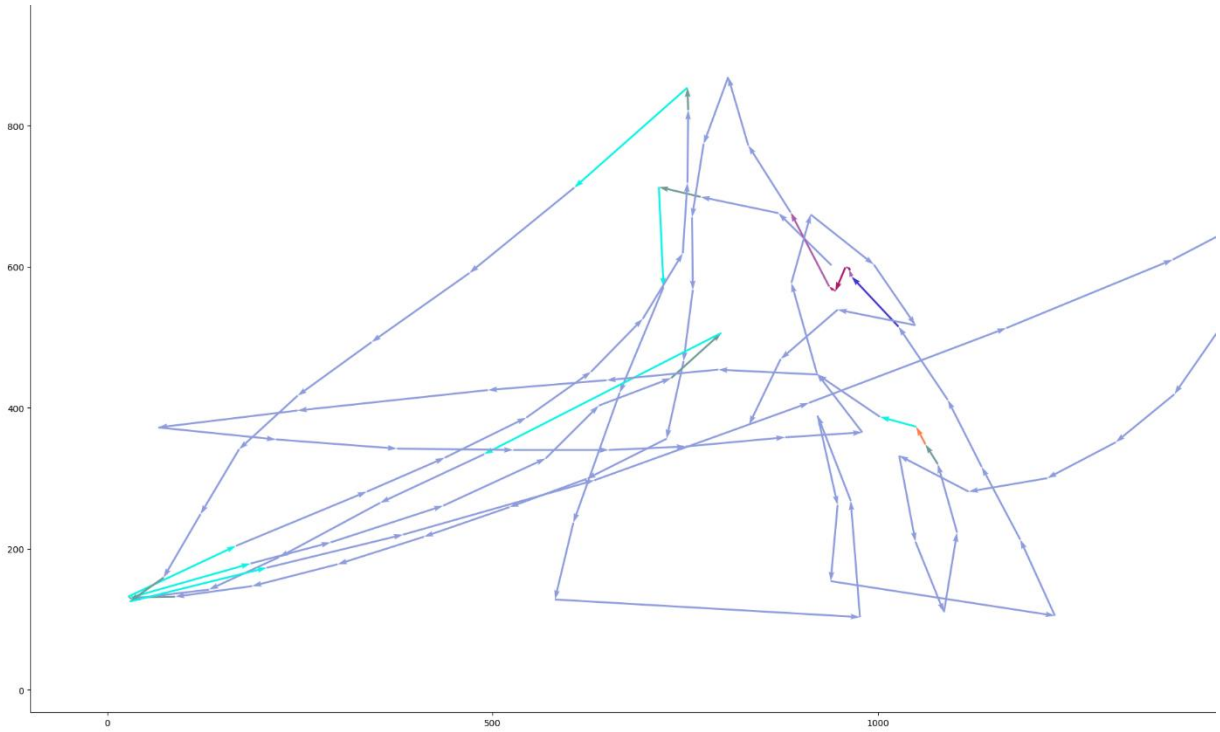


Fig19. User's screen movement using mouse dynamics

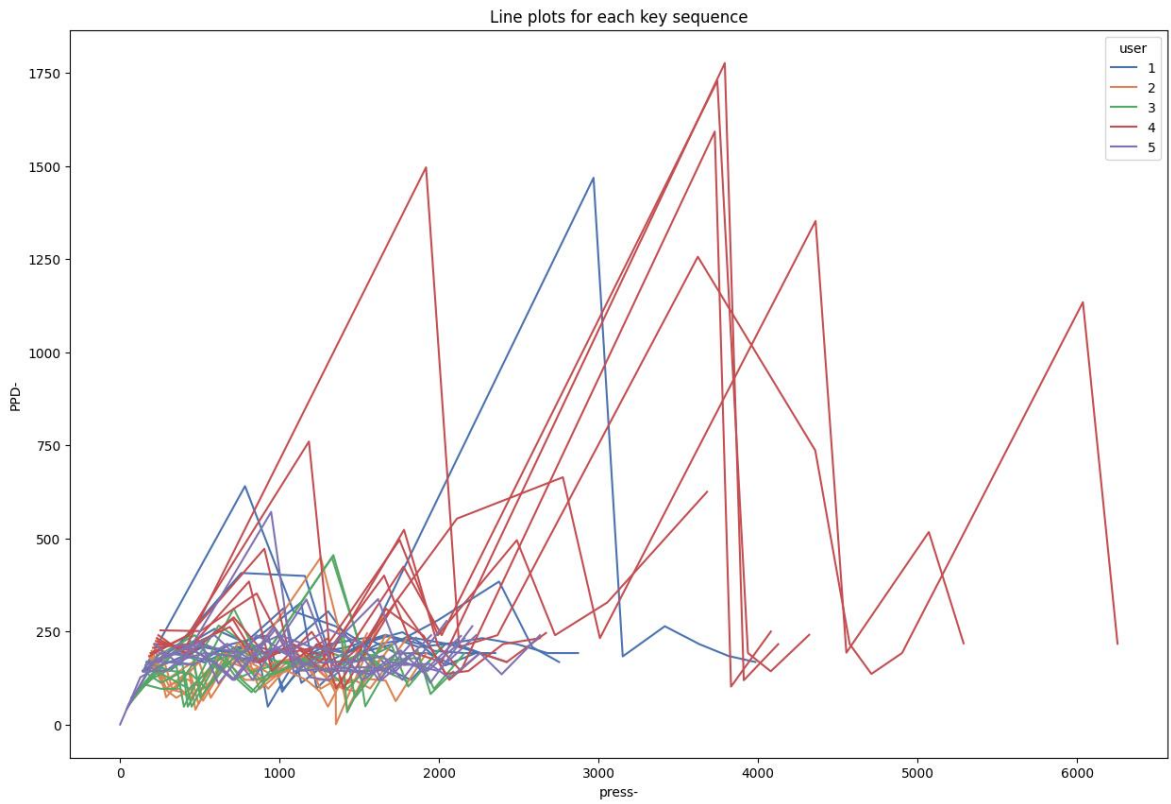


Fig 20. Different users line plot for mouse dynamics

5.2 PERFORMANCE EVALUATION

Comparison between KNN and XGBoost accuracies	
K-nearest neighbor (KNN) Accuracies	XGBoost Classifier Accuracies
0.6363636363636365	0.8959090909090909
0.5795454545454546	0.8729545454545454
0.5829545454545455	0.8556818181818182
0.5852272727272727	0.8859090909090909
0.5965909090909091	0.8443181818181818
Average Accuracy: 0.5951948051948052	Average Accuracy: 0.8629545454545454

Average Accuracy= 0.86 for the best estimator XGB Classifier

5.3 EVALUATION

Numerous studies have evaluated the performance of mouse dynamics-based authentication systems, demonstrating promising results.

Accuracy rates ranging from 85% to 95% have been reported, with FARs and FRRs typically below 5%.

These results indicate that mouse dynamics can be an effective modality for user authentication

CHAPTER 6 - CONCLUSION AND FUTURE SCOPE

6.1 CONCLUSION

Modelling and predicting of the user activity through the keystroke dynamic and movement of the mouse can help in perfecting the security, personalization, and the user experience in different applications, but that is their promise. This approach will enable the system to analyse the users' original patterns in the way of typing and using a mouse, as a result they will get an opportunity to give a distinctive user identification, intentions, and preferences. Still, the method could be something very powerful in humanizing the given sentence.

In an optimistic outlook, the keystroke dynamics and the mouse movements bring in an extra security layer in comparison to other conventional methods like passwords and biometrics. It pioneers in a way of exploitation of the extensive real time behavioural information, which makes it less vulnerable to both personality and fraud. In an additional way, it is also can adapt to shifting user behaviour as time goes by, consequently, it boosts the process of getting stronger.

Besides that, this technology also plays a major role in personalization and optimization of the reality user experience. With gathering of what the users are inclined to when using the interfaces, applications can thus easily personalize the experiences according to what each individual prefer, which consequently results in better usability and satisfaction. From adapting user interfaces to come up with recommendations based on the data analysed from mouse movements and keystrokes, there is a lot of improvement that can take place for the different sectors.

These barriers include the cost of installation and management of the systems and finding the funds to ensure that these systems scale globally. Data privacy issues relating to the retrieval of behavioural details and their safekeeping should be explained transparently and guarantee security of participants as well as compliance with regulations and laws. Moreover, the models'

accuracy and reliability depend a lot and even how good as the quality and volume of training data so it can result in data emergencies like the lack or noisy data.

Various aspects of modelling and predicting user behaviour by analysing their keystrokes and mouse movement analysis.

1. Security Enhancement:

Continuous Authentication: The use of keystroke dynamics and mouse movement evaluation is a substitute for the ongoing authentication quotation, which can act complementary or in lieu of the common static techniques, such as passwords or biometrics. While static biometric credentials are based on the way an individual looks or scans, behavioural biometrics provides real-time identification insights of user behaviour based on the way they interact with devices.

Resistance to Impersonation: Digital biometrics capitalizes on the idiosyncrasies of users as shown through the variation of typing and mouse movement patterns, which increases the chances that these fraudsters will be spotted and brought to justice. At a minimum, even when a hacker has stolen plausible password they would still probably find it challenging to mimic the user's action patterns.

Adaptive Security: These systems can come to fit the user's changing behavior over the whole time period, for example a user to speed up his or her typing or mouse movement pattern. Such a practice ensures that they become more resilient against unpredictable hazards such as evolving threats, a situation where an insider could attack or credentials compromised.

2. Personalization and User Experience Optimization:

Tailored Experiences: Through the usage of keystroke dynamics and mouse activity monitoring, a particular application is capable of capturing the users' preferences and behavior pattern. This information may then be used to customize the user experience by providing the user with a wide range of options, including tailored recommendations, content, or interfaces.

Improved Usability: Realizing how people go about interacting with interfaces will have positive impact on creation of intuitive and delightful applications. Take for instance IUIs that are able to adaptively shift functionalities, or even layout designs, to reflect the observed behaviour of the user thus improving usability.

Efficient Input Methods: It can even help designers of user interfaces develop faster input options. An illustration of this case is the predictive algorithms in text messaging that can recognize users' previous touch / patterns to enable them to type faster using their mobile devices or keyboards.

3.Challenges and Considerations:

Privacy Concerns: Collecting and analysing behavioural biometric data raises a question of privacy right, since everybody is tracked and everything is registered concerning the individuals' activity with devices. Having strong privacy measures such as anonymization, encryption, and user consent comes in Real handy because it helps in shielding sensitive data.

Data Quality and Quantity: The precision and stability of behavioural biometric models are ensured by sequencing and diverse training samples. The first step of building working models is to collect a large amount of data that represents the real experience of users in terms of behavioral patterns. These data might be disrupted by noise or inconsistencies which can determine model effectiveness.

Usability and Accessibility: The recognition framework of behavioural biometrics must enable the usage of the different types of hand movements (hereinafter distinct typing patterns), physical disability, and other forms of impairment. Creating an open system which serves to users from different backgrounds with no effect on security level is a key to making sure that the system is usable and accessible.

Balancing Security and Convenience: Keeping the golden middle by securing the favour of users and fulfilling requirements for security is extremely important. Behavioural biometrics

would offer heightened security though sometimes it can result in user experience becoming too difficult or cumbersome and therefore cause abandonment or frustration.

Finally, the key-board dynamics and mouse movement modelling and recognition play a vital role in security enhancement, personalization, and user's experience optimization. Nonetheless, dealing with concerns regarding privacy, data quality, usability, and strikes a balance between the security and convenience of the technologies, lies in the success of but not the deployments and adoption of such systems. The thoughtful thinking and practice of the quickest measures have behavioural biometrics that will play a key role in the transformation of automated authentication and user interaction in the future.

6.2 FUTURE SCOPE

The scope for future modelling projects looking into user behaviour recognition and prediction via keystroke dynamics and mouse movement measurements is inexhaustive, with chances of discovering new opportunities in several domains for creativity and advancement. Here are some potential avenues for future exploration

Here are some potential avenues for future exploration:

1. Enhanced Security Solutions:

Multifactor Authentication: Combination of behavioral biometrics with current authentication processes makes a security system more robust, as a result of providing a next layer of authentication. In further research, the introduction of the variables of keystroke dynamics as well as mouse cursor movement along with the list of other biometric modalities and/or authentication elements could be used for authentication systems of higher accuracy.

Adversarial Defense: The arising challenge of producing methods that can recognize and handle the delusive adversarial attacks against the motion/pattern/rhythmic, i. e. the behavioral biometric systems will play the key role in supporting their credibility and universal applicability against the sophisticated threat.

2. Privacy-Preserving Technologies:

Secure Protocols: Research studies will make secure protocols as well as alcohol cryptographic technologies appropriate for the use of privacy-preserving behavioral biometric systems. Thus, safeguarding strategies like, e. g. , homomorphic encryption or a secure multiparty computation help running the risk analysis without breaking the confidentiality of users' behavioural data.

Differential Privacy: Behavioral biometric data are considered vulnerable to privacy risks. Therefore, the application of differential privacy techniques encompasses an addition of noise to the data without affecting the analysis capability of them.

3. Behavioral Insights and Applications:

Predictive Behaviour Modelling: The recent progress in machine learning, apparent in the deep learning and reinforcement learning techniques, allows building models which are able to imitate and identify users by their keystroke patterns and moving the mouse. The models may find use not only in these areas but also in personalized recommendation system, fraud detection services as well as proactive user assistance.

Healthcare Monitoring and Diagnosis: In the context of remote health assessment, behavioral signal processing would prove to be an effective technology to pinpoint the level of health of an individual. The systems that observe the shifts in typing and mouse behavior by comparing them against time, could be used to identify early indicators of neurological disorders, cognitive deterioration, or mental health conditions which would enable timely intervention as well as treatment option.

4. User Experience Optimization:

Adaptive User Interfaces: A good thing to ponder what the next applications are in the future. One may like to deal with adaptive user interfaces which, by the means of monitoring the real-time behavioral biometric data, dynamical response can be achieved. The user interface could be customized to the content, layout, and system interaction which could be adapted by

individual users' preferences and needs to match so to maximize usability and eventually, overall user experience.

Context-Aware Assistance: Utilizing behavioral biometrics combined with artificial intelligence to make virtual assistants/personalities or chatbots context-aware and individualized interactions possible. They would learn how to accept changes made by the user, including the features he preferred and the circumstances in which he was acting, to make his assistance more relevant and useful.

5. Ethical and Regulatory Considerations:

Ethical Guidelines and Governance Frameworks: The concern for the future research is to craft ethical standards and regulatory systems that would be conducive for storing, as well as, using biometric behavioral data in a responsible manner. Such policies should try to solve relevant issues, including user consent, transparency, fairness, and accountability, with them enforcing protection of users' privacy and rights.

Regulatory Compliance and Legal Compliance: Besides these constant obligations to the developing standards for the protection of data, privacy, and security, respecting the relevant rules and regulations becomes primary if it is to be ensured that the process runs smoothly, and legal exposure is minimized. Next ventures must primarily address these laws so that they can protect the privacy and safety of users.

In a nutshell, the scope of future projects for this domain is vast and intertwined, including but not limited to security, privacy, and user experience as well as ethical questions. The targeting will involve tackling the challenges and identifying innovations as well to achieve the potential that is with keystroke dynamics and mouse movement analysis to predict user behaviour.

REFERENCES

- [1] I.H. Sarker, Y. B. Abushark, F. Alsolami, and A. I. Khan, “IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model,” *Symmetry*, vol. 12, no. 5, p. 754, May 2020, doi: 10.3390/sym12050754. [Online]. Available: <http://dx.doi.org/10.3390/sym12050754>
- [2] A. Al-Khazzar, N. Savage, Graphical authentication based on user behaviour, in: 2010 International Conference on Security and Cryptography (SECRYPT), IEEE, 2010, pp. 1–4.
- [3] M. Ahsan, R. Gomes, Md. M. Chowdhury, and K. E. Nygard, “Enhancing Machine Learning Prediction in Cybersecurity Using Dynamic Feature Selector,” *Journal of Cybersecurity and Privacy*, vol. 1, no. 1, pp. 199–218, Mar. 2021, doi: 10.3390/jcp1010011. [Online]. Available: <http://dx.doi.org/10.3390/jcp10100L11>
- [4] Y. Zhauniarovich, I. Khalil, T. Yu, M. Dacier, A survey on malicious domains detection through DNS data analysis, *ACM Comput. Surv. (CSUR)* 51 (4) (2018)
- [5] Z. Yang, L. Wang, X. Song, Secure model based on multi-cloud for big data storage and query, in: 2016 International Conference on Advanced Cloud and Big Data (CBD), IEEE, 2016, pp. 207–214.
- [6] L. Liu, Security and privacy requirements engineering revisited in the big data era, in: 2016 IEEE 24th International Requirements Engineering Conference Workshops (REW), IEEE, 2016, p. 55.
- [7] Amazon Web Services, Online, AWS [Online]. Available: <https://aws.amazon.com>, 2006
- [8] J.-H. Lee, Y.S. Kim, J.H. Kim, I.K. Kim, K.-J. Han, Building a big data platform for large-scale security data analysis, in: 2017 International Conference on Information and Communication Technology Convergence (ICTC), IEEE, 2017, pp. 976–980.

- [9] R. More, A. Unakal, V. Kulkarni, R. Goudar, Real time threat detection system in cloud using big data analytics, in: 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE, 2017, pp. 1262–1264.
- [10] S. Shiva, S. Roy, D. Dasgupta, Game theory for cyber security, in: Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research, ACM, 2010, p. 34.
- [11] S. Khan and D. Hou, "Mouse Dynamics Behavioral Biometrics: A Survey," Clarkson University, USA.
- [12] S. Almalki, K. Roy, and P. Chatterjee, "Continuous authentication using mouse clickstream data analysis," in 2011 4th International Conference on Information Technology and Communication (ICITC), IEEE, 2011, pp. 1-4.
- [13] D. Qin, S. Fu, G. Amariuca, D. Qiao, and Y. Guan, "MAUSPAD: Mouse-based authentication using segmentation-based, progress-adjusted DTW," in 2019 IEEE International Conference on Biometrics (ICB), IEEE, 2019, pp. 1-6.
- [14] J. M. Ackerson, R. Dave, and N. Seliya, "Applications of recurrent neural network for biometric authentication & anomaly detection," in 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2017, pp. 2715-2720
- [15] J. Rose, Y. Liu, and A. Awad, "Biometric authentication using mouse and eye movement data," in 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2014, pp. 2802-2807.
- [16] N. Siddiqui, R. Dave, M. Vanamala, and N. Seliya, "Machine and deep learning applications to mouse dynamics for continuous user authentication," in 2019 21st International Conference on Pattern Recognition (ICPR), IEEE, 2019, pp. 1-6.

- [17] Z. Jorgensen and T. Yu, "On mouse dynamics as a behavioral biometric for authentication," in 2016 IEEE International Conference on Biometrics (ICB), IEEE, 2016, pp. 1-6.
- [18] Y. Shi, X. Wang, K. Zheng, and S. Cao, "User authentication method based on keystroke dynamics and mouse dynamics using HDA," in 2017 IEEE International Conference on Cybernetics and Intelligent Systems (ICIS), IEEE, 2017, pp. 1-6.
- [19] A. V. Berezniker, M. A. Kazachuk, I. V. Mashechkin, M. I. Petrovskiy, and I. S. Popov, "User behavior authentication based on computer mouse dynamics," in 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2018, pp. 1383-1388.
- [20] V. Channarayappa, V. Monaco, and C. C. Tappert, "Mouse movement biometric system," 2013. [Preprint].
- [21] "Continuous Authentication Using Mouse Dynamics and Recurrent Neural Networks" by M.A. Sasse et al. (2022) , IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS)
- [22] "Mouse Dynamics-Based User Authentication Using a Multi-Scale Convolutional Neural Network" by Y. Xu et al. (2021), IEEE Transactions on Information Forensics and Security
- [23] N. Siddiqui, R. Dave, M. Vanamala, and N. Seliya, "Machine and deep learning applications to mouse dynamics for continuous user authentication," in 2019 21st International Conference on Pattern Recognition (ICPR), IEEE, 2019, pp. 1-6.
- [24] "User Behavior Authentication Based on Computer Mouse Dynamics" by A.V. Berezniker et al. (2021) Conference: IEEE International Conference on Computational Intelligence and Virtual Environment (CIVE)
- [25] "User Behavior Authentication Based on Mouse Dynamics and Keystroke Dynamics" by D. Hou et al. (2022) Journal: IEEE Transactions on Dependable and Secure Computing.

APPENDIX

Vipul report

ORIGINALITY REPORT

9%

SIMILARITY INDEX

8%

INTERNET SOURCES

4%

PUBLICATIONS

%

STUDENT PAPERS

PRIMARY SOURCES

1

www.ir.juit.ac.in:8080

Internet Source

3%

2

www.researchgate.net

Internet Source

2%

3

ir.juit.ac.in:8080

Internet Source

1%

4

researchr.org

Internet Source

<1%

5

export.arxiv.org

Internet Source

<1%

6

link.springer.com

Internet Source

<1%

7

par.nsf.gov

Internet Source

<1%

8

fastercapital.com

Internet Source

<1%

9

www.coursehero.com

Internet Source

<1%

10	web.archive.org Internet Source	<1 %
11	www.semanticscholar.org Internet Source	<1 %
12	"Security, Privacy, and Anonymity in Computation, Communication, and Storage", Springer Science and Business Media LLC, 2019 Publication	<1 %
13	dokumen.pub Internet Source	<1 %
14	www.mdpi.com Internet Source	<1 %
15	123dok.net Internet Source	<1 %
16	Amit Benbassat. "Evolving Lose-Checkers players using genetic programming", Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games, 08/2010 Publication	<1 %
17	developer.yodlee.com Internet Source	<1 %
18	"Table of contents", 2017 IEEE Security and Privacy Workshops (SPW), 2017 Publication	<1 %

19	docplayer.net Internet Source	<1 %
20	Jigyasa Handa, Saurabh Singh, Shipra Saraswat. "Approaches of Behavioural Biometric Traits", 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019 Publication	<1 %
21	Silvio Barra, Aniello Castiglione, Maria De Marsico, Michele Nappi, Kim-Kwang Raymond Choo. "Cloud-Based Biometrics (Biometrics as a Service) for Smart Cities, Nations, and Beyond", IEEE Cloud Computing, 2018 Publication	<1 %
22	healthdocbox.com Internet Source	<1 %
23	tudr.thapar.edu:8080 Internet Source	<1 %
24	financedocbox.com Internet Source	<1 %
25	www.education.eku.edu Internet Source	<1 %
26	Chenghong Huang, Theresa Liang, Shinji Harada, Eunsung Lee, Tobias Ritter. "Silver-Mediated Trifluoromethoxylation of Aryl	<1 %

Stannanes and Arylboronic Acids", Journal of the American Chemical Society, 2011

Publication

27	athenaeum.uiw.edu Internet Source	<1 %
28	athene-forschung.unibw.de Internet Source	<1 %
29	cps-vo.org Internet Source	<1 %
30	crypto-coemlogina.gitbook.io Internet Source	<1 %
31	ijns.jalaxy.com.tw Internet Source	<1 %
32	vdoc.pub Internet Source	<1 %
33	zenodo.org Internet Source	<1 %
34	Yutong Shi, Xiujuan Wang, Kangfeng Zheng, Siwei Cao. "User authentication method based on keystroke dynamics and mouse dynamics using HDA", Multimedia Systems, 2022 Publication	<1 %

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
PLAGIARISM VERIFICATION REPORT

Date:

Type of Document (Tick): PhD Thesis M.Tech Dissertation/ Report B.Tech Project Report Paper

Name: _____ Department: _____ Enrolment No _____

Contact No. _____ E-mail. _____

Name of the Supervisor: _____

Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): _____

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

- Total No. of Pages =
- Total No. of Preliminary pages =
- Total No. of pages accommodate bibliography/references =

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at(%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

(Signature of Guide/Supervisor)

Signature of HOD

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

Copy Received on	Excluded	Similarity Index (%)	Generated Plagiarism Report Details (Title, Abstract & Chapters)	
	<ul style="list-style-type: none"> • All Preliminary Pages • Bibliography/Images/Quotes • 14 Words String 		Word Counts	
Report Generated on			Character Counts	
		Submission ID	Total Pages Scanned	
			File Size	

Checked by
Name & Signature

Librarian

.....

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com