IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 22, Issue 3, Ser. II (May - June 2020), PP 13-19
www.iosrjournals.org

A Comparative Study on Classifier Algorithms Based on Mouse Gesture Data for User Detection

Masud Karim[1], Md. Amir Hasnat[2], Hasnain Heickal[3], Md. Hasanuzzaman[4]
[1]*Department of Computer Science & Engineering, University of Dhaka, Bangladesh*
[2]*Department of Computer Science & Engineering, University of Dhaka, Bangladesh*
[3]*Department of Computer Science & Engineering, University of Dhaka, Bangladesh*
[4]*Department of Computer Science & Engineering, University of Dhaka, Bangladesh*

Abstract–*In the age of information security, user detection methods are highly recommended for technological innovations. Continuous detecting a real user is a challenging issue. For this reason, researchers are giving importance on mouse gesture pattern to detect user. Different classifier algorithms are using in this technology. This paper triumphs a comparative study on classifier algorithms, Support Vector Machine, K-Nearest Neighbor and Naive Bayes based on mouse gesture for user detection process. Benchmark data of mouse gesture are collected and for our own testing more dataare captured by using Jitbit Macro Reader. Mouse gesture features are generated from both benchmark data and our collected data. Classifier algorithms are applied for these features to detect user mouse gesture pattern. It is found that K-Nearest Neighbor classifier shows the best performance to detect user.*

Keywords—*Classifier Algorithm, Support Vector Machines (SVM), K-Nearest Neighbor (kNN), Mouse Gesture, Naïve Bayes, Behavioral Authentication.*

Date of Submission: 07-05-2020 Date of Acceptance: 21-05-2020

I. Introduction

This paper presents a comparative data analysis of user detection from the experiences of mouse gesture data. In terms of highly secured smart system, information security management is dynamically changing. This paper will be a good reference for comparative analysis of different classifier algorithms on mouse gesture data for user detection.

This study gives a primary introduction to mouse gesture features and examples of feature values. We have shown brief discussions on three classifier algorithms are Support Vector Machine, K Nearest Neighbor, Naïve Bayes. Then visualizes the resultant figures of all kinds of stimulated data analysis and describes the comparative study of different classifier algorithms in favor to this paper as well. Finally, the conclusion summarizes the results and gives the future projections for further research works.

II. Related Work

Chinmayee.KS et al. [1], have proposed a framework to authenticate user from mouse movement data that covers four modules: gesture creation, data acquisition and preprocessing, feature extraction and classification. They also worked on static authentication. They have analyzed mouse movement data by using Hidden Markov Model. The mouse movement data captured by asking user to draw a mouse gesture. First gestures are kept as template and next time user replaced gesture several times and these compared with template. The proposed system is recommended for complex environment like e commerce or e learning.

Anam Khan et al. [3], have published a survey on performance analysis of mouse movement based user authentication researches. The resultant is compared based on FAR and FRR. Mouse dynamic features are shown during their study and also the basic features are identified. Algorithms and classifications are used in various related features those are briefly described. This paper shows how to using tools and techniques of using mouse movement based authentication.

S. Suganya et al. [4], have addressed a method that creates a database of containing mouse dynamic data like co-ordinate value time stamp value and mouse operation as well. From these features, features vectors are achieved. The referred dataset contains static mouse behavior data of 20 users which is collected from available open source.

By using diffusion map algorithm, they have proposed to reduce the dimension of future vectors. The author used neural network classes which prepared from the number of simple and highly interconnected processing elements. Performance analysis is shown by comparing existing proposed systems as well as FAR and

FRR where the resultant values are FAR=4.15% and FAR=5.05%. The performance of proposed system observed better.

G. Muthumari et al [6], discussed on classification where the extracted features are analyzed and performed authentication using Learning Vector Quantization (LVQ). The classifier algorithm LVQ is used to find the identity whether the given sample is authorized or unauthorized. The data captured from user mouse movement behavior where a drawing interface is used to capture data. The mouse dynamic data consists of x-y coordinates and elapsed time of users when drawing some predefined letters on illustration area. In this method LVQ is compared with SVM (Support Vector Machine) by calculating FAR and FRR. The performance of the proposed system provides values FAR=7.25% and FRR=6.25%. The test result demonstrates that, the proposed method LVQ gives low error rates along with good accuracy than the existing method SVM classifier.

Anand Motwani et al. [7], presented two phases, one is enrollment phase and another one is verification phase. At enrollment phase different mouse and keyboard events data are captured to extract the features and creates profile database. In verification phase user's features are matched with that profile database. A probability and knowledge based generation step applied here. The probability phase recognizes who is the most likely user and in knowledge generation phase the users are classified on valid or invalid user which is more accurate. Three sample of each user and 27 features are taken in this research.

III. Background

A. Mouse GestureData

Many researchers have derived different features regarding mouse movement pattern. Calculations and analysis of these features are also varies from researcher to researchers. [5] [8] [12] According to background study this paper works with twelve mouse gesture features those are listed below and proposed in our another research paper [13]. These features are considered to calculate from user's mouse gesture dataset.

1) Number of Points in the Trajectory
2) Delay Time
3) Number of Delay
4) Number of Action
5) STDEV of the Trajectory Length
6) Total Length of Trajectory
7) STDEV of Slope (m)
8) STDEV of Difference Between Each of Slopes
9) Number of Curvature
10) Curvature of the Trajectory
11) Number of Changes in Horizontal Position
12) Number of Changes in Vertical Position

** STDEV: Standard Deviation*

This paper experiences the studies of how the other researchers have work with mouse gesture features. The mouse gesture data arises from some actions like mouse click, highlights, mouse idle time, mouse movement, drag and drop, system interrupt by mouse event, mouse gesture direction etc. [2] [7] [9] [10] [11].

In this research data are captured based on mouse event and captured data are processed as shown calculation in [13]. From the processed data, features values are generated in real number format. In Table 1 features value of two uses are shown.

Table no 1: Example of features' value of two users

Sl.	Features	Value of User A	Value of User B
1	Number of Points in Trajectory :	13	12
2	Delay Time :	502	72
3	Number of Delay :	7	6
4	Number of Action(no of rows) :	20	18
5	STDEV of Trajectory Length :	12.629556	6.491907
6	Total Length :	345.29009	713.8437
7	STDEV of Slope :	0.0840552	0.181709
8	STDEV of Difference between each of Slopes:	0.0855285	0.052777
9	Number of Curvatures :	12	12
10	Curvature of the Trajectory :	0.6971007	9.614984
11	Number of Changes in Horizontal Position :	12	13
12	Number of Changes in Vertical Position :	12	13

IV. Learning and classification algorithm

Three classifier algorithms are used in this research.

1) Support Vector Machine (SVM)
2) K Nearest Neighbor (kNN) and
3) Naïve Bayes.

These are all supervised learning algorithm. These algorithms are useful in cases where an attribute or property is available for a certain dataset and needs to be predicted for other instances. To train the system, a set of instances are labeled with appropriate classes which is called training set. The algorithm processes these training data and predicts class label of unknown instances based on these training dataset. For this reason, these algorithms are called supervised. In this study some part of data from both benchmark data and our own generated data are used for train up the algorithms. The remaining part of data are used to test the performance of algorithms. Tested output results are described in next sections.

V. Comparative Analysis of classifiers

This section illustrates output status of classification algorithms. Performance of three classifiers, SVM, kNN and Naive Bayes are shown here in terms of *visual classifiers error* curve and *margin curve*. These two curve play important role to observe the performance of classifier. We have shown the performance of output result in comparatively for both our collected data and benchmark data[14].

The v*isualize classifier error curve*(Figure 1 - Figure 3)brings up a visualization window that plots the results of classification. Correctly classified instances are represented by crosses, whereas incorrectly classified ones show up as squares.

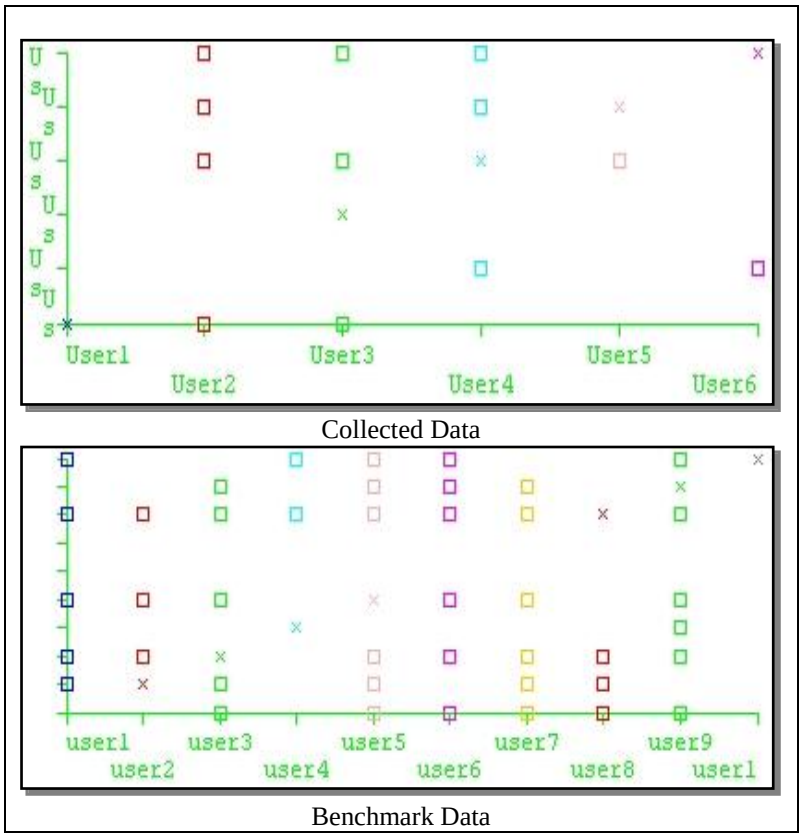

Fig 1: Visual classifier error of SVM

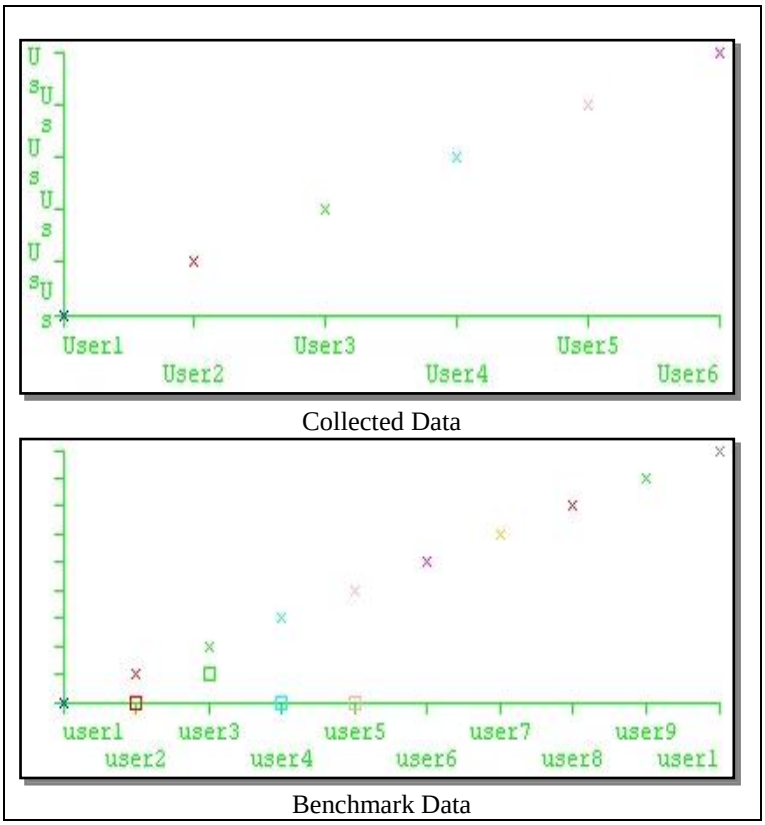

Fig 2: Visual classifier error of kNN

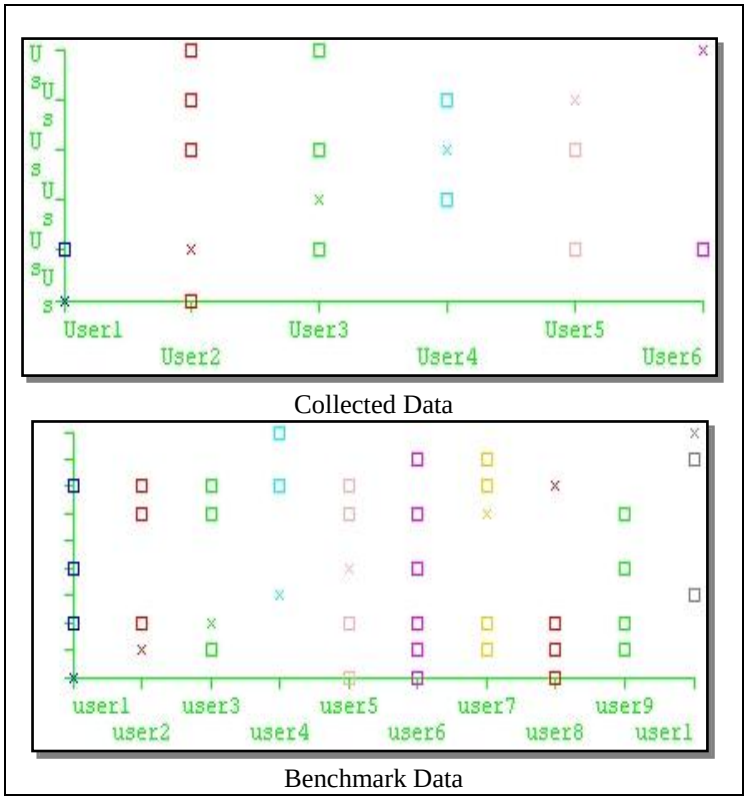

Fig 3: Visual classifier error of Naïve Bayes

The *visual margin curve*(Figure 4 – Figure 6)show the cumulative frequency. The prediction margin is found from the plot. The plot presents the difference between the probability predicted for the actual class and the

highest probability predicted for the other classes. Values below 1 presets error classification and such class is not the correct one. On the other side the values towards 1 presents better performance of classification.

In case of kNN classifier the margin goes towards 1. So this classifier shows the best performance.

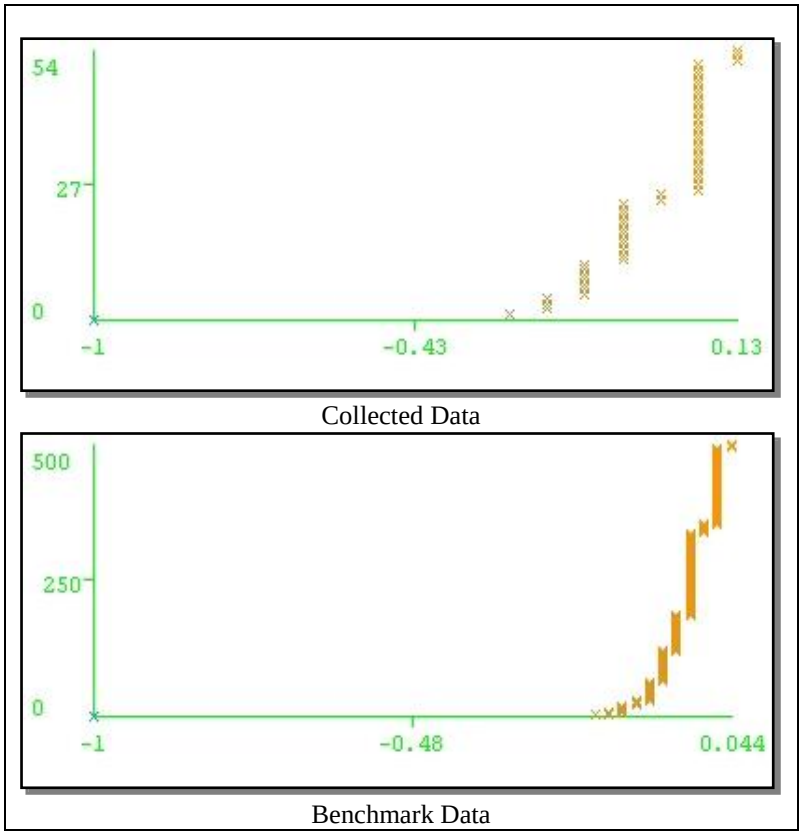

Fig 4: Margin curve of SVM

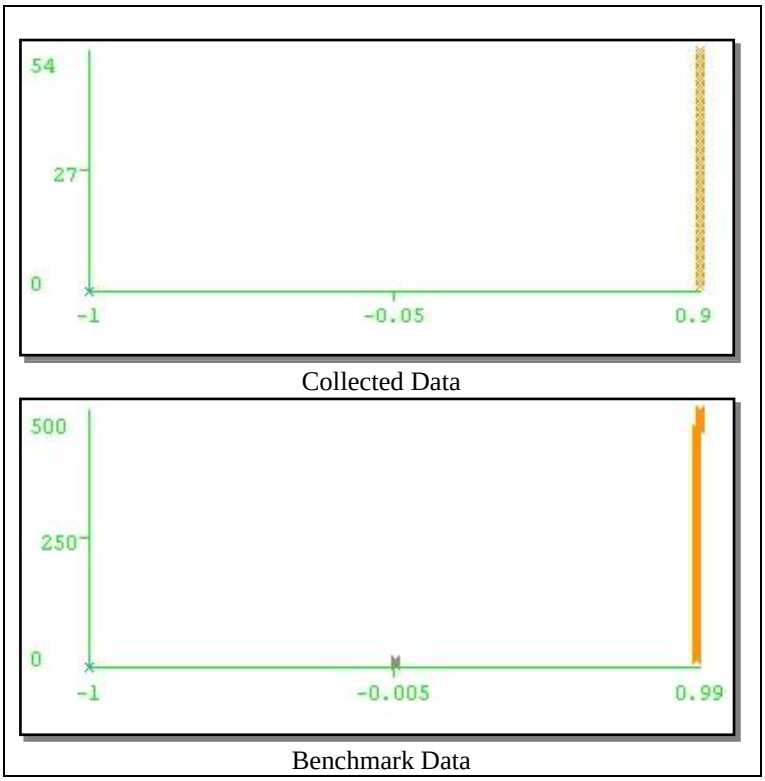

Fig 5: Margin curve of kNN

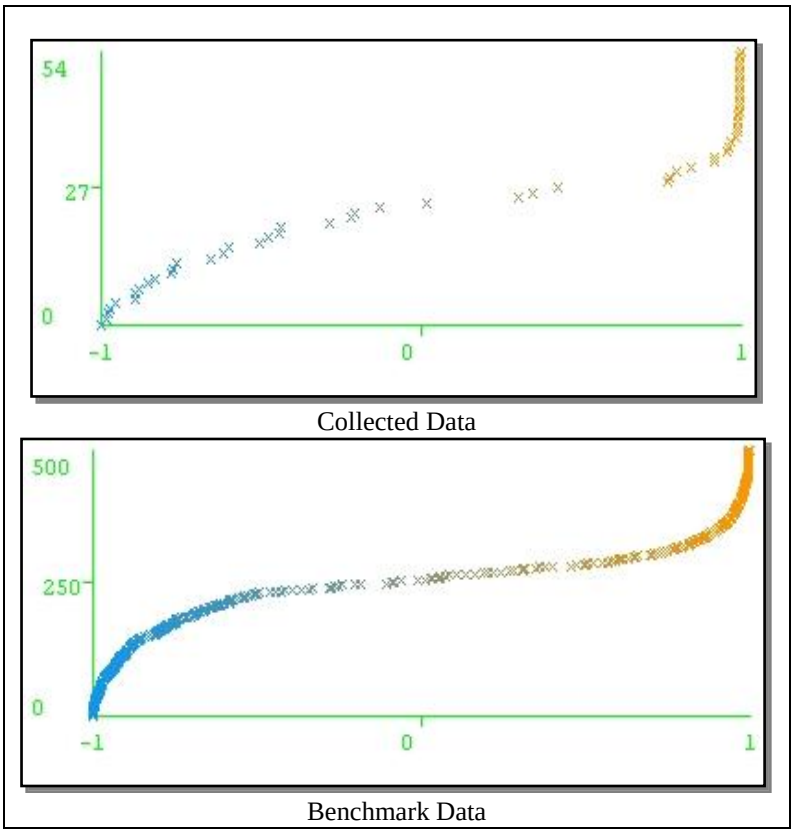

Fig 6: Margin curve of Naïve Bayes

From the performance of the above scenario, it is assumed that classifiers work better over the bench mark data. Results of benchmark data show lower error rate than result of our collected data and kNN classifier has the best performance.The accuracy of the test depends on how well the test separates the group being tested into those with and without the disease in question.

We have shown output of learning and classification from Figure 1 to 6. A total output data can be taken from these figures. We have observed the performance of classifiers in terms ofvisual classifier error (Figure 1 to 3) and visual margin curve (Figure 4 to 6) of SVM, kNN and Naïve Bayes classifiers. These outputs are shown for both experiments, with using data of our own and benchmark data. In both cases we have found that kNN classifier shows more accuracy than other classifiers.

VI. Conclusion

Successively security is becoming important issue. Valid user detection is a challenge task. We have analyzed user mouse movement data by using different classifier algorithms to detect user. Among of these algorithms best three are compared in this paper.Though our research objective is to find and comparative result among different classifier algorithms then we can consider the K-Nearest Neighbor based classifier would be a good reference in our further research work related to mouse movement based user identification.

II. Acknowledgement

We acknowledged department of Computer Science and Engineering, University of Dhaka for supporting in this project. We also thank to ICT Division, Ministry of Posts, Telecommunications and IT, Government of the People's Republic of Bangladesh for funding.

References

[1]. Chinmayee.KS, Vanishree C; "Continuous Authentication for Mouse Gesture Recognition using Hidden Markov Model"; InternatIonal Journal of electronIcs & communIcatIon technology; Vol- 8, Issue 4 OCT - DEC 2017.

[2]. Ahmed Awad E. Ahmed and IssaTraore, Member, IEEE; "A New Biometric Technology Basedon Mouse Dynamics"; IEEE Transactions on dependable and secure computing, Vol.4, No.3, July-September 2007.

[3]. Anam Khan, Suhail Javed Quraishi, Sarabjeet Singh Bedi; "Mouse Dynamics as Continuous User Authentication Tool"; International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8 Issue-4, November 2019.

[4]. S. Suganya, G. Muthumari, and C. Balasubramanian; "Improving the Performance of Mouse Dynamics Based Authentication Using Machine Learning Algorithm"; International Journal of Innovation and Scientific Research; ISSN 2351-8014 Vol. 24 No. 1 Jun. 2016, pp. 202-209.

[5]. Mr. Vishal S. Patil; Prof. P.R.Devale; "Biometric Validation by Storing different Patterns using Mouse Gesture Signatures"; International Journal on Recent and Innovation Trends in Computing and Communication; Volume: 3 Issue: 4; April 2015.
[6]. G. Muthumari, R. Shenbagaraj, M. BlessaBinolin Pepsi; "Mouse Gesture Based Authentication Using Machine Learning Algorithm," 2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT).
[7]. AnandMotwani, Raina Jain, JyotiSondhi; "A Multimodal Behavioral Biometric Technique for User Identification using Mouse and Keystroke Dynamics"; International Journal of Computer Applications (0975 – 8887); Volume 111 – No 8, February 2015.
[8]. Raina K. Jain, SanchetiShital D, PansareJyoti T, KawaleShubhangi K; "Mouse and Keystroke Based Behavioural Biometrics: Techniques, Problems and Solutions"; International Journal of Modern Trends in Engineering and Research ; e-ISSN No.:2349-9745, Date: 2-4 July, 2015.
[9]. B J Gorad, D.V Kodavade, "User Identity Verification Using Mouse Signature"; IOSR Journal of Computer Engineering (IOSR-JCE); e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 12, Issue 4 (Jul. - Aug. 2013), PP 33-36.
[10]. A. Kaklauskas, E. K. Zavadskas, M. Seniut, M. Krutinis, G. Dzemyda, S. Ivanikovas, V. Stankevič, Č. Šimkevičius, A. Jaruševičius; "Web-Based Biometric Mouse Decision Support System For User's Emotional And Labour Productivity Analysis"; The 25th International Symposium on Automation and Robotics in Construction; June 26-29, 2008; PP-69-75.
[11]. Youssef Nakkabi, IssaTraoré and Ahmed AwadE.Ahmed; "Improving Mouse Dynamics Biometric Performance Using Variance Reduction via Extractors with Separate Features"; IEEE Transactions On Systems, Man, And Cybernetics — Part A: Systems And Humans, Vol.40, No.6, November 2010.
[12]. Haitao Tang and Wang Mantao; "User Identity Authentication Based on the Combination of Mouse and Keyboard Behavior"; International Journal of Security and Its Applications; Vol. 10, No. 6 (2016) pp.29-36.
[13]. Masud Karim, Hasnain Heickal, Md. Hasanuzzaman; User Authentication from Mouse Movement Data Using Multiple Classifiers; 2017 International Conference on Machine Learning and Computing, DOI: http://dx.doi.org/10.1145/3055635.3056620.
[14]. Chao Shen, "Performance Evaluation of Anomaly-Detection Algorithms for Mouse Dynamics" (Computers&Security-2014). http://nskeylab.xjtu.edu.cn/people/cshen/ data-sets/behavior-data-set. [15 December 2015].

Masud Karim, et. al. "A Comparative Study on Classifier Algorithms Based on Mouse Gesture Data for User Detection." *IOSR Journal of Computer Engineering (IOSR-JCE)*, 22(3), 2020, pp. 13-19.