Sixth International Conference on Emerging trends in Engineering and Technology (ICETET'16)
www.ijera.com
ISSN: 2248-9622, pp.64-70

RESEARCH ARTICLE OPEN ACCESS

Anticipate Human Mistakes in Statistical Typing Using Data Mining Methods and Human Manners Modeling

Mrs. J.V.Anchitaalgammai[1], G.Gokula Sharmi[2]
[1]*Assistant Professor Department of computer science and Engineering Velammal college of Engineering and Technology*
[2]*PG scholar Department of computer science and Engineering Velammal college of Engineering and Technology Madurai*

ABSTRACT
Numerous types of electronic devices with alphabetical or numerical keyboards have become very important tools have been used now days. An erroneous keystroke can be easily caused by many reasons, such as the operators' inexperience, fatigue, and carelessness. Numerical typing errors can lead to serious consequences, but various causes of human efforts and the lack of the contextual clues in numerical typing make their prediction difficult. Find an error prediction to avoid financial loss in statistical typing, using a threshold values the error rate will be classified and find an error and nonerror task features. Modeling features produced by an enhanced human behavior model were combined to improve error classification performance by Linear Discriminant Analysis (LDA) classifier. Find an error prediction to avoid the financial loss in the statistical typing, preprocessing the dataset and To better predict human errors, a computational performance model that involves abstract, cognitive, and motor processes may be necessary. Using the bank dataset of human behavior modeling and psychophysiologicaldata's, find a prediction values and using an LDA classifier then the behavior modeling features can improve the error prediction by data classification better than apparent task features. To improve the accuracy level uses a root mean square classifier

***Index terms*:** linear discriminant analysis, statisticaltyping, datamining, humanmistakes, behavior modeling.

I. INTRODUCTION

Human mistakes in statistical typing tasks can induce financial loss, safety threats, or even fatalities in critical systems. A BCI is a system which allows a person to controlspecial computer applications. Then the behavioral, psychophysiological, and biomechanical evidence should converge to explain how humans commit errors.

Error detection is to be difficult because lack of contextual clues. In this paper focused on reaction errors under time pressure conditions, which yielded an error rate permitting a comparison of correct and error trials. The preparation, actual operation and mental imagination of limb movements activate similar EEG changes at sensor motor areas on the scalp. In successive responses during numerical typing, a slow reaction in one keystroke can delay another keystroke, and typing with single or multiple fingers may imply different difficulty. If a model can predict final behavior outcomes (e.g., human errors) based on environmental inputs, such as external task demands and time stress, then the intermediate outputs of the model that are used to determine the final outcomes should be classified. In these four apparent task features were selected to be test with EEG features alone can produce similar results to that of combining modeling features with EEG features, they are Number to be pressed, Quickness of previous keystroke, Fitts' difficulty index, Movement angle of the keystroke. The main objectives of the paper are, it presents our system which is able to integrate the latest technologies, in order to find a statistical typing error prediction framework based on integrated psychophysiological and behavioral modeling features. Find an error prediction to avoid the financial loss in the statistical typing, preprocessing the dataset and To better predict human errors, a computational performance model that involves abstract, cognitive, and motor processes may be necessary.

Using the bank dataset of human behavior modeling and psychophysiological, find an prediction values and the LDA classifier then the behavior modeling features can improve the error prediction by data classification better than apparent task features.

ISSN: 2248-9622, pp.64-70

II. METHODOLOGY

Preprocessing:

Preprocessing is commonly used to tokenizing the dataset into a set o keywords commonly the dataset is to be in text or number oriented if text means for preprocessing have to be remove an stopping and stemming words. In this preprocessing the unwanted spaces, commas etc are to be removed.To avoid processing redundant patterns, we partition the total set of contrast patterns into equivalence classes (EC), each consisting of all patterns sharing a common matching data set. Since patterns having the same matching data can be considered as having the same behavior, it suffices to consider just one pattern per EC. Technically, a pattern Pis a member of an EC of patterns defined as EC (P) = fQj mds (Q) = mds (P) g. It can be shown that each EC can be described by a closed pattern (the longest in the EC) and a set of minimal-generator (MG) patterns (minimal with respect to); so an EC contains all patterns Q satisfying "Q is a superset of some MG and Q is a subset of the closed pattern, of the EC".

Predict Rating:

In the CPXR approach, used to first split D into two classes, LE and SE, consisting of instances of D where f makes large/small prediction errors respectively. CPXR then searches for a small set of contrast patterns of LE to optimize the trr measure, and uses that set to build a PXR model. Intuitively, the search aims to find a set of patterns, such that each pattern is highly useful in correcting prediction errors off, and when combined the patterns in the set give high prediction accuracy.

Importantly, we associate with each pattern P a local regression model fP built using P's matching data, mds (P), as training data. There is no limit on what variables fP can use. While any regression method can be used to build f P, one may prefer linear regression or piecewise linear regression models, since the simplicity of such models helps ensure that PXR prediction models are easy to interpret.

Human Error:

To better predict human errors, a computational performance model that involves abstract, cognitive, and motor processes may be necessary. In continual responses during arithmetic typing, a slow retort in one keystroke can delay another keystroke, and typing with single or many fingers may imply different difficulty. Time weight also impacts the spatial changeability of activities adversely.

Classification:

Classification is a data mining task of predicting the value of a categorical variable (target or class) by building a model based on one or more numerical or categorical variables. In the classification algorithm there are four main groups one of that is covariance matrix. In these two types of algorithm are to be used. In these paper used a linear discriminant analysis classifier.

LDA:

Linear Discriminant Analysis is a classification method. It is simple mathematically robust and often produces models whose accuracy is as good.Compute the *d*-dimensional mean vectors for the different classes from the dataset. Compute the scatter matrices (between-class and within-class scatter matrix).Compute the eigenvectors (**e_1, e_2, ..., ed**) and corresponding eigenvalues (λ_1, λ_2, ..., λ_d) for the scatter matrices. Sort the eigenvectors by decreasing eigenvalues and choose **k** eigenvectors with the largest eigenvalues to form a $d\times k$-dimensional matrix **W** (where every column represents an eigenvector). Use this $d\times k$ eigenvector matrix to transform the samples onto the new subspace. This can be summarized by the equation

$$Y = X \times W$$

(Where$X$ is an $n\times d$-dimensional matrix; the *i*th row represents the *i*th sample, and Y is the transformed $n\times k$-dimensional matrix with the *n* samples projected into the new subspace).

ISSN: 2248-9622, pp.64-70

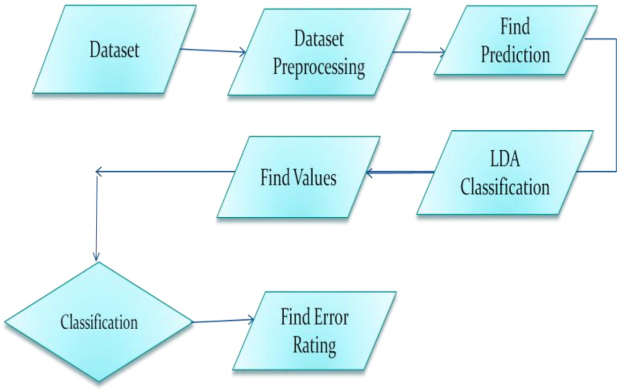

Fig 1.System Architecture

III. FUNCTION FLOW

This explains the entire process of the project. Using the human manners and psychophysiological dataset based on the banking process. Here collectinga raw data to predict human mistakes. From the raw dataset the data's are to be preprocessed.Then the advanced human mannersmodeling with real-time psychophysiological (EEG) features could produce better error anticipate results by an LDA classifier for human mistakes in statistical typing. Then predict the rating values for the attributes, and an error values and based on the range values the error and non error can be classified by applying LDA algorithm, thenclassify the error and nonerror process. Based on the range values, the range values is above the threshold then there is an error process, if the range values is below the threshold then there is a correct process.

IV. RESULT

The figure describes the initialization of the dataset choosing and start the execution of code.

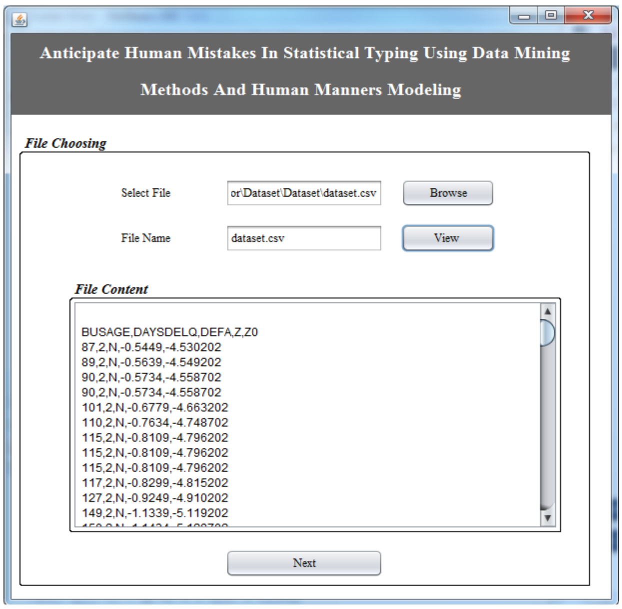

Fig 2.Initialization of dataset choosing

The figure describes the starting status of the dataset selection execution and view the dataset. The dataset is to be viewed and starting the execution to be proceeded. The dataset has to be chosen based on the human activity.

ISSN: 2248-9622, pp.64-70

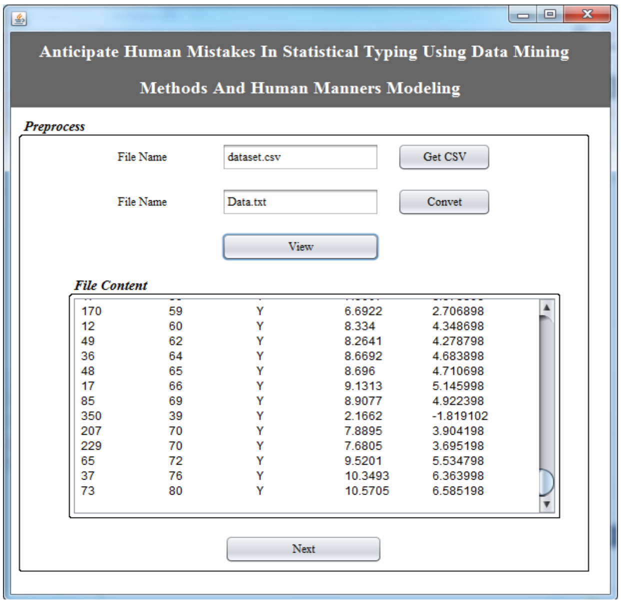

Fig 3. Preprocess the dataset

The figure is that shows the fully preprocessed dataset by removing the unwanted spaces, commas. Then the dataset has to be cleaned.

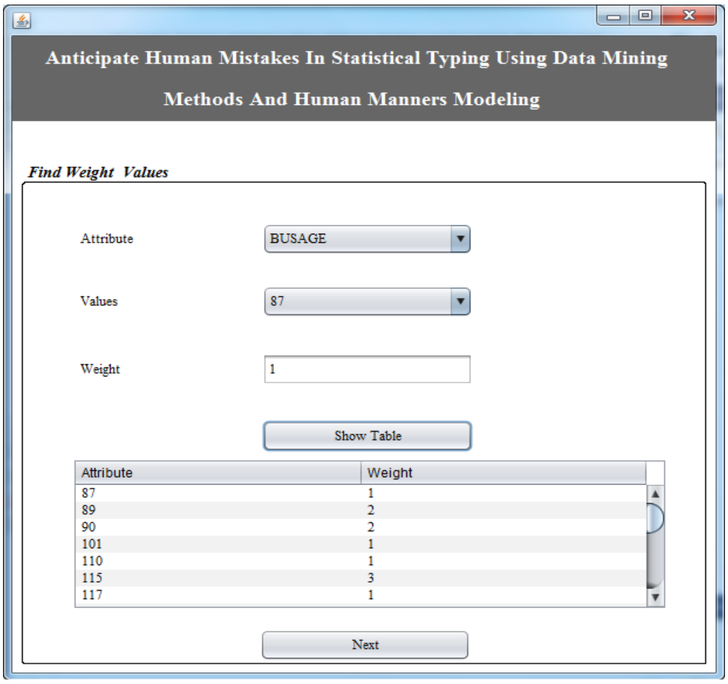

Fig 4. Finding an weight values

ISSN: 2248-9622, pp.64-70

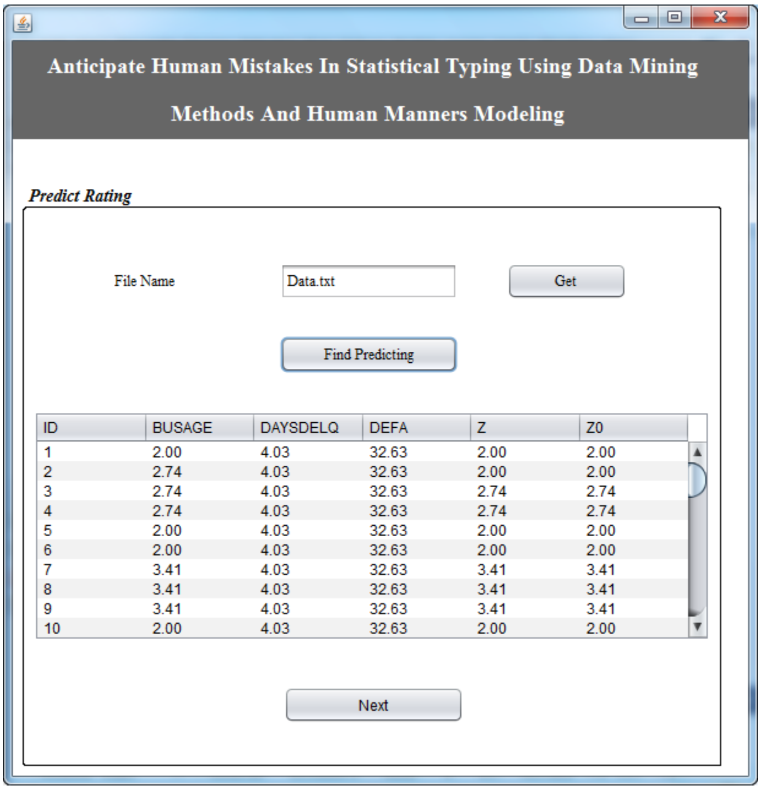

Fig 5. Predict the rating for attribute

The figure is that shows the rating for each attribute in the data and finding an error values for the predicted rating attributes.

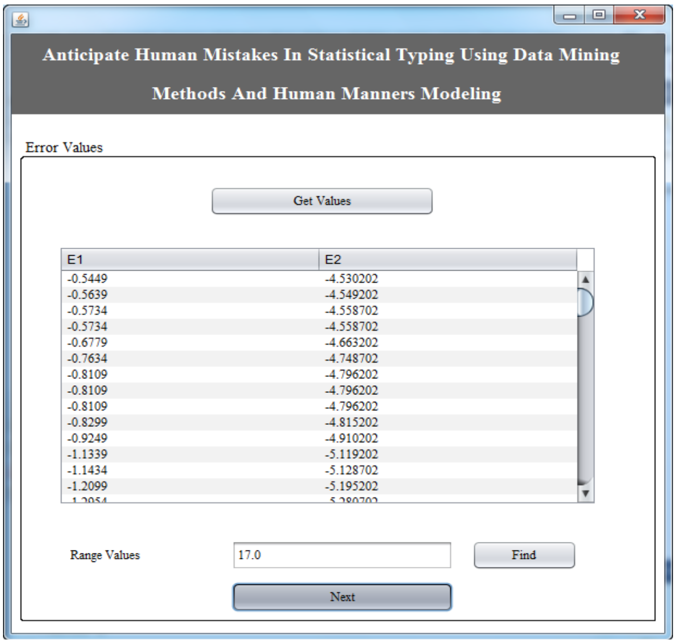

Fig 6. Finding error values

This figure describing the error values and the range values based on the range values the error and non error can be classified by applying LDA algorithm.

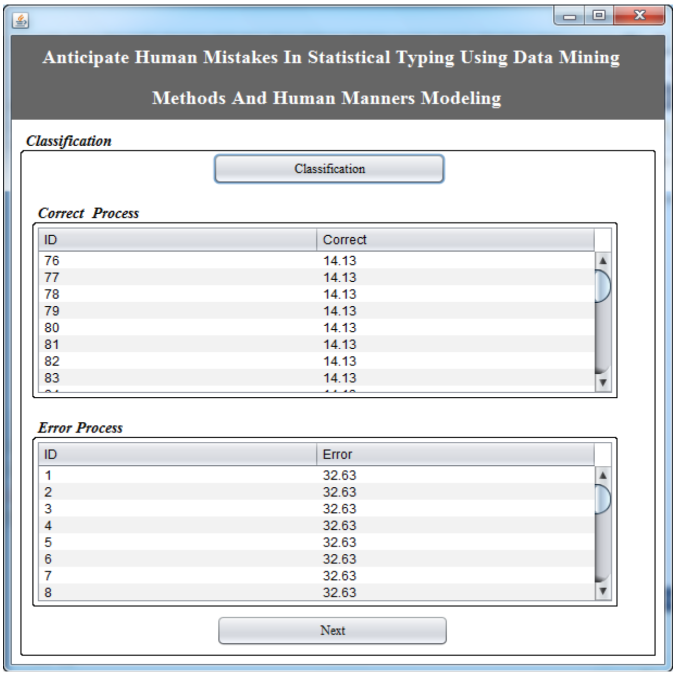

Fig 7.classify the error and nonerror process

V. CONCLUSION

In this project work can have an Integrating modeling features produced by an advanced human behavior model with real-time psychophysiological (EEG) features could produce better error prediction results by an LDA classifier. The data having an attributes based on that the data preprocessed and predict the rating for the attributes and finding an error values and the range values based on the error and the correct process are to be classified by the LDA classifier. Then the performance might indicate carelessness which could have caused less distinguishable psychophysiological feedback and resulted in outlier results in classification.Were to utilize both observable human behaviors and nonobservable behaviors by prediction and compare their effectiveness with apparent features available without having a model. Applications of data mining techniques to analyze error-related EEG patterns generally provide improvements in performance.

VI. FUTURE WORK

In this paper existing work used an LDA classifier to classify the data. In this paper the future work, using root mean square that can be used to classify the data. Root mean square is the one of the pruning technique. In Future using double pruning to improve the classification accuracy and remove the outliers from dataset. Also using pruning technique to improve the classification accuracy. Pruning is a technique in machine learning that reduces the size of data by removing sections of the data that provide little power to classify instances. So the prediction level of error can be increased. There are number of challenges in the field of Brain computer Interfacing will be addressed. One among them is to be predicting the human mistakes in statistical typing.

REFERENCES

[1]. H. Thimbleby and P. Cairns, “Reducing number entry errors: Solving a widespread, serious problem,” J. Roy. Soc. Interface,sOct. 2010.
[2]. J. M. O’Hara, J. C. Higgins, and W. S. Brown, “Human factors considerations with respect to emerging technology in nuclear power plants,” Brookhaven National. Laboratory., Upton, N.Y., USA, Rep. NUREG/CR- 6947, 2008.
[3]. N. Bohm, B. Gladman, I. Brown, C. Schaufler, and J. F. v. Schiller, “Banking and bookkeeping,” in Security Engineering: A Guide to Building Dependable Distributed Systems, R. J. Anderson, Ed. 2nd ed. New York, NY, USA: Wiley, 2008
[4]. M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein, “ERP components on reaction errors and their functional significance, 2000.
[5]. J. R. Fedota and R. Parasuraman, “Neuroergonomics and human error,” Theor. Issues Ergonom, 2010.

[6]. T. A. Salthouse, "Perceptual, cognitive, and motoric aspects of transcription typing," Psychol. Bull., 1986.
[7]. C.-J. Lin and C. Wu, "Factors affecting numerical typing performance of young people in a hear-and-type task," Ergonomics, 2011.
[8]. A. Subasi, "Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients," Expert Syst. Appl.,2005.
[9]. K.-Q. Shen, X.-P. Li, C.-J. Ong, S.-Y. Shao, and E. P. V. Wilder-Smith, "EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate," Clin. Neurophysiol.,2008.
[10]. R. Agarwal and J. Gotman, "Adaptive segmentation of electroencephalographic data using a nonlinear energy operator," in Proc. IEEE Int. Symp. Circuits Syst., 1999.
[11]. S. Wang, C.-J. Lin, C. Wu, and W. A. Chaovalitwongse, "Early detectionof numerical typing errors using data mining techniques," IEEE Trans.Syst., Man, Cybern. A, Syst., Humans, vol. 41, no. 6, pp. 1199–1212,
[12]. sNov. 2011.
[13]. B. E. John, "TYPIST: A theory of performance in skilled typing," Human Comput. Interaction, vol. 11, pp. 321–355, 1996.
[14]. D. E. Rumelhart and D. A. Norman, "Simulating a skilled typist: A studyof skilled cognitive-motor performance," Cognitive Sci.: Multidisciplinary J., vol. 6, pp. 1–36, 1982.
[15]. B. Blankertz, C. Schafer, G. Dornhege, and G. Curio, "Single trial detectionof EEG error potentials: A tool for increasing BCI transmissionrates," in Artificial Neural Networks, J. Dorronsoro, Ed. Berlin, Germany:Springer, 2002, pp. 138–138.
[16]. C. W. Anderson and Z. Sijerˇci´c, "Classification of EEG signals from four subjects during five mental tasks," in Proc. Conf. Eng. Appl. Neural Netw., 1996, pp. 407–414.
[17]. P. Ferrez and J. d. R. Mill´an, "You are wrong!—Automatic detection of interaction errors from brain waves," in Proc. Int. Joint Conf. Artif. Intell.,2005, pp. 1413–1418.