

Performance Evaluation of various Machine Learning Algorithms for Personality Detection using Handwritten Document

PUTTASWAMY B. S.¹ VAISHNAVI R²

¹Assistant Professor, Dept. of ISE, P. E. S. College of Engineering, Mandya

²Student, Dept. of ISE, P. E. S. College of Engineering, Mandya

Abstract - Handwriting is solitary of the individual characteristics to express all the things in ours mind, to speak with one another. Handwriting expose actual characters including behavior, psychological expenditure, self-respect, temper, fears, honesty, creativity, uprightness, anxiety, defense and plenty of other personality traits. Graphology is the methodology as understanding and evaluates along with grab the writer's traits through the configuration, and structure of the word within the handwriting. A Characteristic is often recognizing through several handwriting features like size of letters, pen pressure, baseline, top margin, slant of letters.

Key words: character, behavior, graphology, personality traits.

1. Introduction

Handwriting likewise named as cerebrum composing is a helpful measure in recognizing the trademark character attributes of a person. Penmanship investigators otherwise called graphologists can analyze a person's penmanship to anticipate the character qualities of the essayist. Computerized penmanship investigation can be utilized to inspect individual attributes of up-and-comers during interviews precisely as the exactness of an expert exceptionally relies upon his range of abilities. Likewise employing a graphologist to break down many examples for enrollment reasons will be tedious and not be practical financially. This work examines about a strategy for breaking down genuine transcribed content examples with guide of innovation. The investigation is accomplished for explicit highlights of the examples for deciding different trademark conduct characteristics of the individual. Different boundaries of the features taken here are top margin, baseline, pen pressure and slant will be contemplated to decide comparing characteristics.

Proficient penmanship analyst is the person who distinguishes the character through the penmanship test are called graphologist. We can decipher the inward brain research and conduct through the manner of seeking (discourse), look, signals, stances and habits of dressing. Penmanship is likewise one of the expressive ways that ways that tell about your tendency, brain science and conduct of the essayist. Penmanship is interesting to every person. Also, it will be same and remarkable for an author whether the person in question has composed with their foot, hand or mouth. Penmanship is composed by the mind, and not by the feet or hand. Thus, the penmanship is otherwise called cerebrum composing. Every character include has neurological mind design in the human cerebrum. Each neurological cerebrum design configuration conveys exceptional neuromuscular

development which is something very similar for every person who has that specific character include. Any motion or development in penmanship uncovers a specific character highlight. Graphology is the science which recognizes these motions in penmanship and depicting the relating character include. Author acknowledgment is utilized for various purposes, for instance, for security, financial action, criminological and uses as access control, examination of penmanship reports can be utilized to pass judgment on the offenders in the criminal equality associations.

2. Literature Survey

Hemalatha et al. [1] proposed an HAS tool which stands for “Handwriting Analysis System” which is used to identify the personality of a person by giving his or her handwriting scanned sample as an input and input features are zone, baseline, size and word spacing which are used to detect the personality trait of the person using HAS tool.

BMGarlapati et al. [2] separated handwritten textual content & revealed text and then every have been given in separate documents. SVM have turn out to be used to categorize them into a category. It separated best absolutely one of a kind instruction. It used 10-fold pass-validation technique to discover accuracy. Some of the crucial limitations of the above papers have been that their tool grouped best super patterns, their tool regarded out for simplest one personality trait at a time, the accuracy became low due to the fact the bounding packing containers overlapped, the inclination of baseline has become measured using the skewness of the baseline which commonly get delivered in some unspecified time in the future of image preprocessing affecting the general accuracy.

A. R. M et al. [3] proposed a HABIT system which means “Handwritten Analysis Based Individualistic Traits Prediction” which predicts the personality of a person based on various features like slant, size, pressure, upper zone loops, lower zone loops and many more. This system proves the person’s personality irrespective of language they had used.

Champa H N et al. [4] proposed a method to predict the personality of the writer by using his or her handwriting. The parameters used here are baseline, the height of the t-bar on the stem of the letter ‘t’, the shape of the loop of letter ‘y’. They have measured baseline using the polygonalization and to calculate the loop of ‘y’ they had used GHT which means Generalized Hough transform method.

Prachi Joshi et al. [5] proposed a methodology to predict the accurate personality of the user by using the attributes like baseline, margin and slant of the words. These features are extracted from the sample taken from user which was then compared with trained data set to give accurate results.

Devesh Sony et al. [6] suggested a method to predict the human behavior using his handwriting. They have used only one kind of pen for this process. And they had used CNN algorithm for classification.

P. K. Grewal et al. [7] presented a method to detect the character of the user by taking his handwritten sample and giving it as an input. And extracting features like letters i.e. i and f, baseline and others. Taking this input as an input ANN which predicts the behavior of the writer.

3. Features and behavior

There are many features in the handwriting which we can use for the prediction of the human behavior. The features we are used here are size, slant of the letter, top margin, pen pressure and baseline angle.

a. Size

The size of the handwritten sample is measured by the height of the letters. At the end, it takes the average height of every letter and predicts the behavior. We have categorized the size into two which are small and large. If the letter size is large then it tells the person's behavior is that, they are sociable, extrovert, want to get noticed by the people in the crowd. On the other hand, if the letter size is small, then they are balanced, peace of mind.

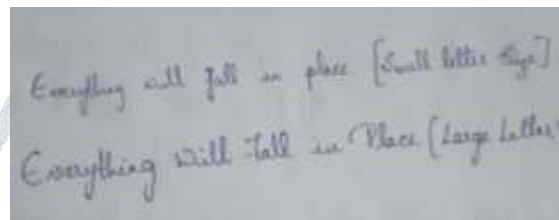


Fig. 1 Size of letters

b. Slant

The slant of the sample taken is determined is, in which way the letters are curved and it is calculated by the angle formed between the baseline of the sample and the letters curved direction. It is divided into three types which are right, left and no slant. The right slant indicates that they are friendly, impulsive and sentimental. The left slant tells that they are reserved, not an open minded and introspective. And no slant tells, they are very strong in their nature and pragmatic person.

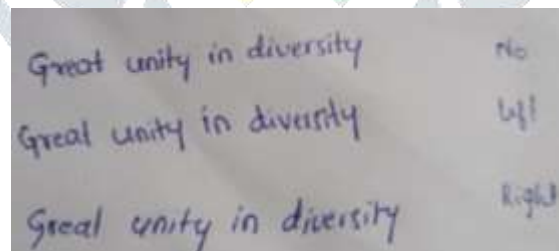


Fig. 2 Slant

c. Top Margin

The top margin means it is the space left by the writer at the top of the page they have written. The user assumes that there is margin in the page and then he or she starts writing. This top margin is of two types i.e., narrow top margin and wide top margin which tells the user behavior based on these criteria. The narrow top margin says that person won't give respect which means they are lack of respect and formality. They are lack of politeness in their behavior. And on the other hand, wide top margin tells that the person in simplicity in their nature, they give respect to all and they behave in formal way.

d. Pen Pressure

Pen pressure is the pressure applied by the writer at the time of writing. It is calculated by setting grey threshold value. The pressure can be heavy, medium and light. If the pressure is heavy, it means they are easily attracted to hot blooded actions which in turn they easily lose their cool and they are short of patience. Medium pressure tells that, they are lively in nature, they can easily gel with others, spirit full. If the pressure is light means, their will power is not so strong, they determination in their life.

e. Baseline

Baseline defines the emotional stability and nature of writer. Baseline can be of three types which are ascending, straight and descending baseline. Each of these tells unique characteristics of the person. Ascending baseline means they are positive in nature, confident, optimism and hopefulness. Descending baseline tells they are pessimistic, disappointment, and they easily to go depression. If the baseline is straight then the person is emotionally stable and he/she handles the situation in the perfect manner.

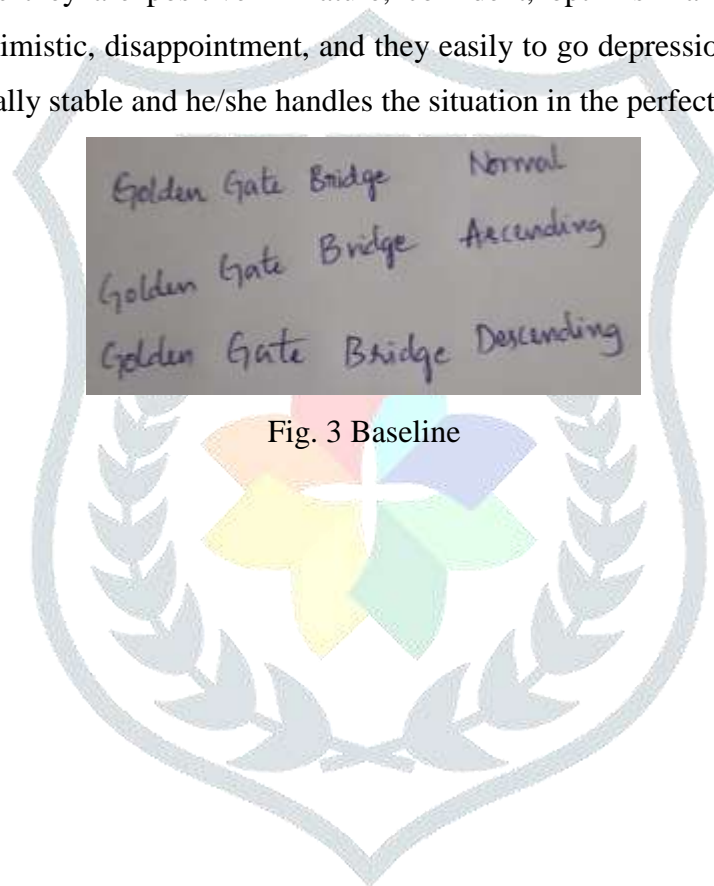
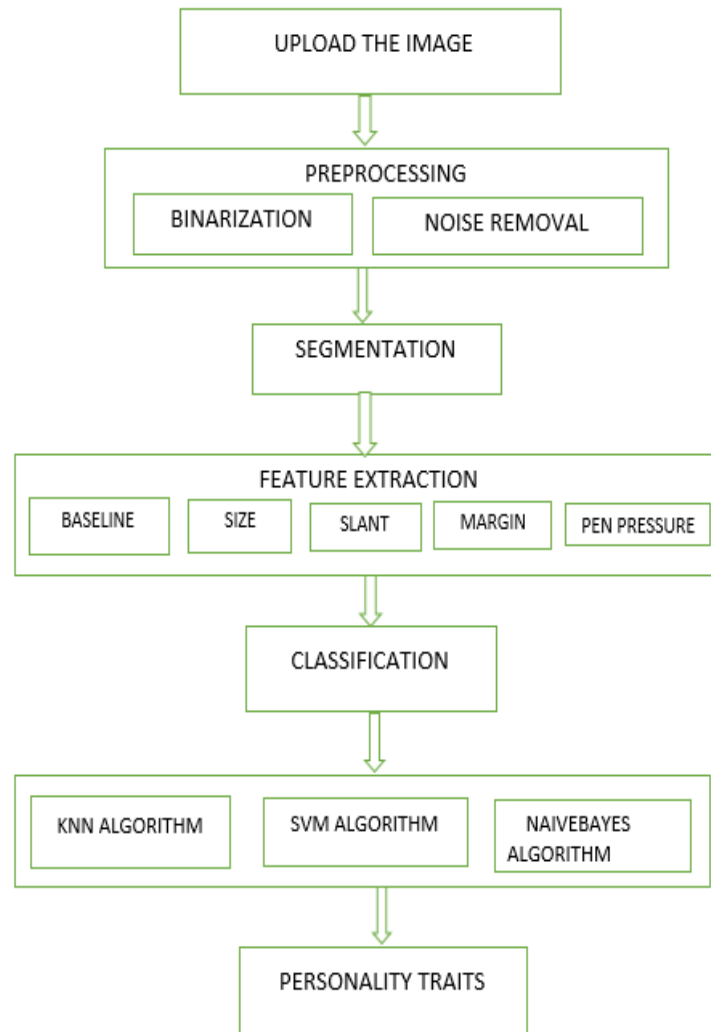


Fig. 3 Baseline

4. System Design

The below diagram shows, how the handwriting analysis takes place.



a. Upload the image

Here we need to upload the image from the dataset we have collected, it can be either Kaggle dataset or real-world dataset. The image file can be in either PNG or JPEG format.

b. Preprocessing

Pre-processing is the procedure on image to improve its quality for more processing. Here Binarization and Noise removal happens. Binarization is termed as the process which converts grey scale image into binary image. Noise removal is the procedure which is applied to remove the data which is not required and to improve the quality of the image. This process is done to remove the variations that are not required for the process which shapes and styles of the letter the writer used while writing.

c. Segmentation

It is the process of dividing the input image into numerous segments i.e. set of pixels. Segmentation is further divided into three types, they are line segmentation, word segmentation and character segmentation. Line segmentation means it segments the input image content into text lines, then word segmentation happens

where it segments images, word by word. After this, character segmentation comes into the picture, here the sample is segmented character by character.

d. Future Extraction

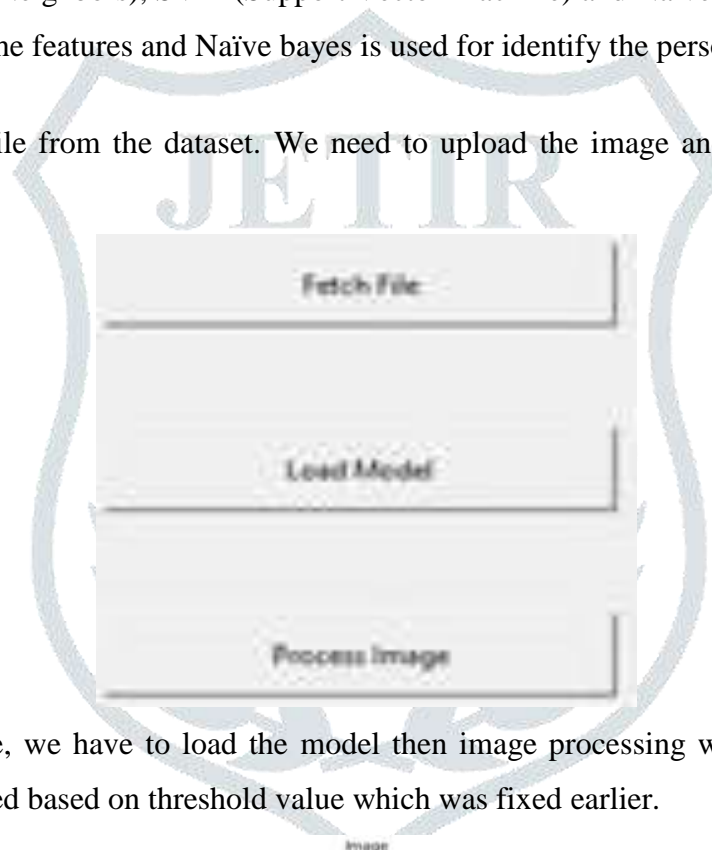
Future extraction is the process of extracting the crucial data from the high dimensional input image. The features extracted here are used to identify the behavior of the person. The features are size, pen pressure, top margin, slant and baseline.

e. Classification

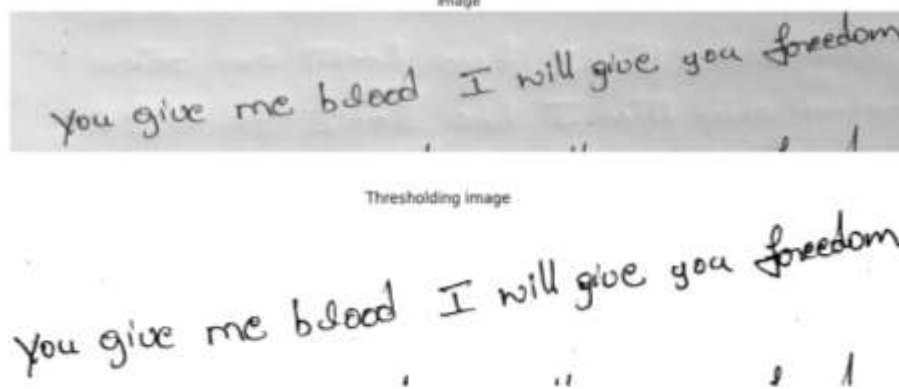
Classification is used to identify the personality of the person by using the features obtained from the feature extraction process. This output is given as the input for the algorithm we have used here. Algorithms we used here are KNN (K-Nearest Neighbors), SVM (Support Vector Machine) and Naïve Bayes algorithm. KNN and SVM are used to classify the features and Naïve bayes is used for identify the personality of the person.

5. Results and Analysis

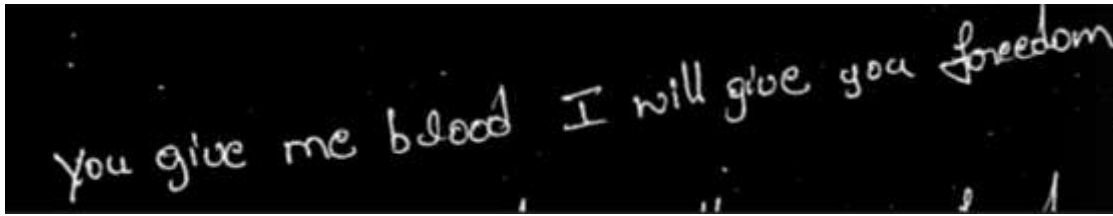
First, we need fetch the file from the dataset. We need to upload the image and then we have to load the model.



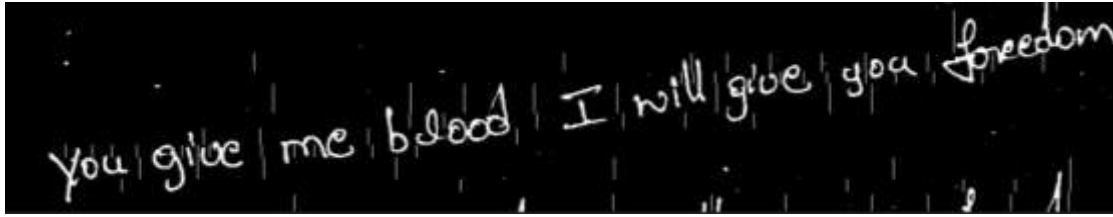
After uploading the image, we have to load the model then image processing will happen. In this process, binary image will be created based on threshold value which was fixed earlier.



Character segmentation into individual character to speed up result and to ensure satisfying recognition rates.



After uploading the image. we have to load the model then image p



Finally predict the personality of a person based on features.

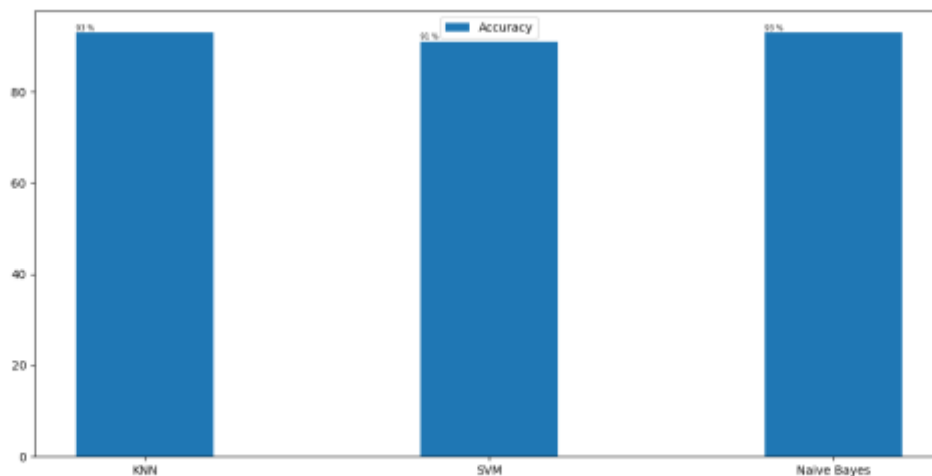
```

C:\Windows\system32\cmd.exe
Baseline Angle: 1
Top Margin: 1
Letter Size: 1
Pen Pressure: 83
Slant: 90
Balanced mind
Very solid and pragmatic
Will power is not strong
Optimism and Hopefulness
Modesty and Formality respect
Press enter to close.

```



Figure 1



6. Conclusion

This project has suggested a procedure to detect the personality of a writer from his/her handwritten sample. By using their handwritten sample, we extract some features and with the help of machine learning algorithms like KNN, SVM we classify the features and identify the character of the person using Naïve Bayes. Here we have extracted features like size of letters, slant of letters, baseline angle, top margin and pen pressure. This extracted feature is then compared with the previously trained dataset. Finally, it is mapped with comparable personality trait.

References

1.Hemlata, Manoj Sachan and Shailendra Kumar Singh, “personality detection using handwriting Analysis Review,” The Seventh international conference on Advance in computing, 2018,pp,85-89.

2. B.M. Garlapati and S.R. Chamala, "A System for handwriting nad Printed Text Classification ." UKSim-AMSS 19th International Conference on Computer Modelling and Simulation (UKSim), Cambridge,201,pp.50-54.
3. A.R. M ANS D. Varghese, "HABIT:Handwritten Analysis Based Individualistic Traits Prediction,"no. 7, pp.209-218, 2013.
4. H.N. Champa and K. R. Anandakumar," Automated human behavior prediction through handwriting analysis ", Proc-1stInt.Conf. Integr.IntellComput.ICIIC 2010,pp,106-109, 2001.
5. Prachi Joshi,Agarwal and Ajinkya Dhavale, "Handwriting Analysis for detection of Personality Traits using MachineLearning Approach", International Journal of computer Application , 2015,pp,40-45.
- 6.O.K. Grewal and Deepak Orashar," Behavior Prediction Through Handwriting Analysis" UCST2012, pp.520-523
7. S.H. Cha and C.C. Tappert, "Automatic detection of handwriting forgery ", Proc, -Int.Work, Front, Handwrit, Recognition, IWFHR, pp ,264-267, 2002.
8. Devansh Sony and Rakshitha Sawanth," Identifying Human Behavior Characteristics using Handwriting Analysis",IRJET,2019,pp, 4436-4439.
9. Aditya Chitlangai and G.Malathi," Handwriting Analysis based on Histogram of Oriented Gradient Personality traits using SVM," ICRTAC,2019.

