Advanced Production and Industrial Engineering
R.M. Singari and P.K. Kankar (Eds.)

doi:10.3233/ATDE220778

Analysis and Recognition of Handwriting Patterns for Personality Trait Prediction Using Unsupervised Machine Learning Approach

ASHISH SARASWAL[a] and URVASHI RAHUL SAXENA[b]
[a]*Department of Computer Science and Engineering, J.S.S. Academy of Technical Education, Noida, India*
[b]*Department of Computer Science and Engineering, J.S.S. Academy of Technical Education, Noida, India*

Abstract. The study of the link between handwriting and personality is known as handwriting analysis. It is a specific manner of using handwriting to convey one's cognitive process. That's why it's also known as "brain writing". Handwriting, like a person's fingerprint, is a unique trait that cannot be duplicated by two distinct people. It also aids us in determining a person's emotional state. Handwriting recognition is usually done in a traditional manner, in which a professional handwriting expert examines a person's handwriting samples for various personality features, which takes a lot of time and effort. As a result, the focus of this study is on developing an automated system that can predict personality characteristics utilizing the machine learning algorithm KNN to identify a certain individual based on their handwriting sample. In today's modern society, predicting personality traits using face, voice, handwriting, and other techniques is becoming increasingly essential. To develop this approach, we look at a document's many features such as letter size, word spacing, and slant of words, pen pressure, and baseline, and so on to anticipate the writer's personality.

Keywords. K Nearest Neighbour (KNN), Handwriting recognition, Personality Traits, Machine Learning, Pre-processing.

1. Introduction

Handwriting is a lifelong process that a person begins learning in childhood. Handwriting analysis is the study of various handwriting styles. It provides information on human behaviour in areas such as social skills, accomplishment, thinking style, and habit. The study also reveals how a person performs under various conditions. A specialist looks for qualities that may be used to predict their personality. The traditional way is to ask the person to write on a piece of paper and assess the handwriting for personality, however this method can take too long. The focus of this paper will be focusing on some of them for personality. Handwriting analysis also help us protect our health because our handwriting and physical health are linked. Similarly, an offline system can be built that is writer independent and can be used without fear of connectivity. However, because it is offline, it contains a lot of noise in the data, which can be overcome by using Gaussian de noising [1] and threshold models to remove unwanted noise from the given data. Handwriting analysis can reveal a person's strengths, talents and interests as well as

identify their weaknesses. It help us to better understand ourselves and others. Several psychological tests have been created by psychologists for personality. Some of the most well-known test used for psychological examinations are Big five test. [2], Disc assessment, MBTI test etc. Personality is a person's character and conduct in various settings, as well as their choice of books, music, and films, among other things. Furthermore, it influences interactions with others and the environment; it can also affect an individual's sense of self-worth

Major contributions of this work are:

(i) On the basis of handwriting prediction, determine an individual's personality traits in a simple and accessible approach.
(ii) It may also be used as an alternative to the questionnaire that asks for a handwriting evaluation of a specific individual.
(iii) Handwriting will be used to identify personality using six different elements.
(iv) In comparison to the usual procedure used by graphologists, time savings will be feasible.
(v) To show how to compare and contrast different techniques.

The rest of the paper is organized as follows: Section II depicts the literature review, Section III depicts the proposed work, Section IV depicts implementation details, and Section V depicts the results and conclusion.

2. Literature Review

Handwriting, signatures, and sleeping moments can all be used to forecast a person's personality. Our study focuses on handwriting analysis as one of these methods. Xi.W et al introduced the Gaussian algorithm, which eliminates noise from input photos and improves image contrast using only an adaptive threshold technique, unlike other algorithms that require two algorithms to remove noise and enhance contrast [1]. Anamika et al examine characteristics such as handwriting flow, size, and slant. It was created using the MATLAB instructions,. The system has pre-defined features. Only one attribute was compared at a time, not all of them. With 95% accuracy, the tilt and handwriting size were determined [2]. Saha et al proposed the technique was tested on over 550 text images from the IAM database and test handwriting images created by various authors. It is based on a different horizontal and vertical projection that can fragment content lines and words regardless of whether they are covered or multi-skewed. [3]. N.Majumdar et al created a document modeling-based approach that employs a CNN extractor to predict a person's features. In their architecture, they employed max pooling and a concatenation layer [4]. D.S Chauhan et al used Hindi character recognition with RBF neural network using handwritten. They used gradient feature extraction to normalise scanned images into 30*30 boxes. 1000 photographs were used in the experiment for training and testing of a particular sample, with 90% accuracy. [5]. Srinivasa Rao Chalamala et al used handwritten and printed text were separated and placed in separate documents. SVM was used to categories them into multiple categories. The 10 fold cross validation procedure was utilized to obtain accuracy [6]. T Gunwan et al In this experiment, we constructed an English handwriting acknowledgment system. In the testing dataset, digit recognition accuracy was 98 percent and letter recognition accuracy was 88 percent. [7]. Manoj et al thrived HABIT, an offline handwriting tool for

personality prediction that is writer independent. The machine was fed scanned pictures of writer handwriting as input [8]. Vaishali.R .Lokhande et al Signature was used to assess a person's personality. It evaluated personality qualities such as the baseline under the signature, the dot above the letter, and the final style of signature. For the baseline, 87 percent accuracy was achieved utilizing horizontal and vertical segmentation. [9]. Rinki Kumari et al projected an offline personality identification is the focus of this article. Extraction, skew edge, zones of handwriting, statistics, feature, slant angle, random and Hough transformations, and soft computing are some of the characteristics used for recognized evidence such segmentation and pre-processing. The handwritten text is classified by a back propagation neural network into a list of known authors [10]. Nithya Lakshmi et al projected an adaptive neuro fuzzy inference system was presented to estimate human writer's personality based on handwriting attributes. The proposed effort would use a Sugeno-based fuzzy inference model to predict personality.Spacing, size, slant, form, loop, dot, pressure, signature, zones, and page margin are all input variables. The fuzzy system was created with the help of the MATLAB 7.1 [11]. P.K Grewal suggested an artificial neural network-based method for behavior analysis that may predict personality attribute. Baseline, letter slant, letter I and letter f were all utilized. Standard back propagation was utilized as the learning algorithm. Based on forecasting behavior, this research exhibits an accuracy of 81 percent [12].Afnan et al conducted a survey on multiple handwriting analysis and discussed in detail what graphology is, as well as the benefits and drawbacks of using a computerized approach to accomplish it and they analyzed numerous studies that were about handwriting analysis but used different approaches like image processing or artificial intelligence to achieve it [13]. Aditya et al projected system based on handwriting recognition, created a histogram oriented gradient approach. The support vector machine method was utilized, and it was able to reach 80% accuracy utilizing five personality qualities features: energetic, extrovert, introvert, sloppy, and optimistic [14]. Mim Liu et al created a handwriting analysis to help with neuromuscular condition diagnosis. The data provided in the research is real-world data that is then utilized to validate the technique produced. It also aids the clinician in the diagnosis of neuromuscular disorders. It also help in early detection of problem [15].

3. Proposed Work

Graphologists are handwriting analyzers that can detect a person's characteristics attributes only by looking at their handwritten samples This process mainly include pre-processing of the handwritten picture, feature extraction using several approaches such as letter size, slant of words, pen pressure, baseline, and space between word and letter, and finally training and classifying them using the KNN algorithm for trait prediction. This work is divided into several module first module shown in Fig 1 where user write a sample which is uploaded in web application and then passed to handwriting analysis system.

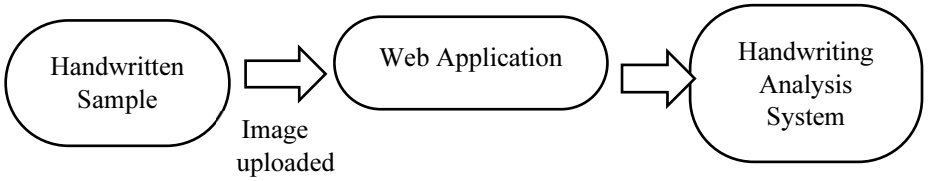

Figure 1. Model of input

In the handwriting analysis system which is second module, various steps are performed which are explained as following:

3.1 Preprocessing

In this process picture is pre-processed using several techniques to eliminate noise and smooth it out for better results. Because black and white images have a lower intrinsic complexity, the first step in image pre-processing is to transform the input picture to black and white, further thresholding process is used to eliminate noise from a picture and sharp it.

3.2 Feature Extraction

Feature extraction is dimensional reduction of data from high-dimensional to lower-dimensional input data which is the fundamental purpose of feature extraction to extract the most important information from the unique data and store it. Six features used are: letter size, slant of words, pen pressure, baseline, and gap between word and letter. Description given below.

3.2.1 Pen Pressure

Amount of force used when writing. It denotes a person's mental state. It determine a person's wrath, feelings, and energy.

3.2.2 *Slant of Words*

Basically three types: left slant says writer is self-reliant, right slant says writer is enthusiastic, and vertical slant says writer is independent.

3.2.3 Letter Size

Size of letter indicates goal in person think in different situation. Mainly two type: large size means person is honest and small size means self-consciousness.

3.2.4 Baseline

The emotional outpouring and temperament of a person are mostly depicted here. Mainly three primary types upward, downward and straight.

3.2.5 Space between Letters

Narrow letters are often associated with caution or self-control. Wide letters may indicate a desire for success and a lack of restraint.

3.2.6 Space between Words

The space between ending starting second word and ending of first word which mainly is associated to emotional state of person

3.3 Classification

Classification is dividing set of data into categories. We will be using a dataset of various images of different handwriting styles collected from EMNIST. Further Euclidean distance will be used to find the distance between the data points available and the value of k that will help classify the data.

3.4 Prediction Result

The output that is obtain after classification and training of data will help to predict personality of the user based on the six different features used to classify the handwriting..

4. Implementation & Results

System is developed in python using Jupiter Notebook Editor and computer with 8GB RAM. User can upload his handwriting after scanning using camera. Handwriting analysis system will show result according to personality matched from user handwriting through the uploaded text from a user of his handwriting. After the image is uploaded it further starts pre-processing which helps to remove unwanted noise and sharps the image to get the accuracy easily.The image is classified to get accurate result of the personality shows the result of accuracy of 89% of user image uploaded in system which gives his personality trait as confident and balance

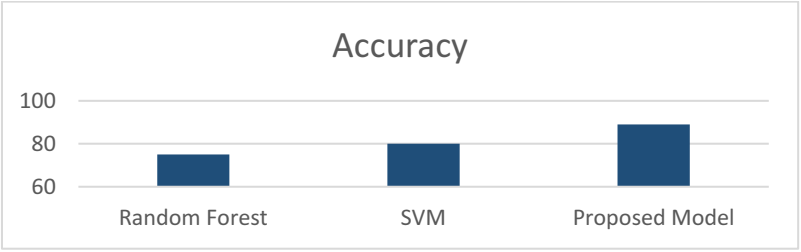

Figure 2. Accuracy comparison result

Fig.2 shows accuracy between different models. Our proposed work can help the graphologist to obtain personality of user quick and accurate. Proposed model build using KNN has 89% accuracy and SVM has 80% accuracy, while random forest has 75%.accuracy.

5. Conclusion and Future Work

In this paper different personalities traits like balanced, optimistic, pessimist, confident etc can be seen in different handwriting sample and our system was able to distinguish between the personality using KNN algorithms. Mainly used features to predict personality are letter size, slant of words, pen pressure, baseline, space between word and letter for the feature extortion process. The system was able to predict personality in real time. It further can be used by company recruitment process for hiring of people using their handwriting sample in future work and also it can also be used mass hiring of

people. In future it can also be used to detect early-stage mental problem as handwriting is brain writing.

References

[1] Xie, W. Huang, H. H. Wang and Z. Liu, "Image de-noising algorithm based on Gaussian mixture model and adaptive threshold modeling," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, 2017, pp. 226-22.
[2] A. Sen and H. Shah, "Automated handwriting analysis system using principles of graphology and image processing," International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, 2017, pp. 1-6.
[3] Bal andSaha, ― An Improved Method for Handwritten Document Analysis Using Segmentation, Baseline Recognition and Writing Pressure Detection‖,Proceeded Computer Science, 93, pp.403-415. 2016.
[4] Majumbder N., Poria S. Deep Learning Based Document Modelling for Personality Prediction from Text IEEE Intelligent Sysmen 2017
[5] Dayashankar Singh, Dr. J.P.Saini "Hindi Character Recognition Using RBF Neural Network and Directional Group Feature Extraction Technique "2015 IEEE An Analysis. In: Singh S., Wen F., Jain M. (eds) Advances in System Optimization and Control. Lecture Notes in Electrical Engineering,
[6] B. M. Garlapati and S. R. Chalamala, "A System for Handwritten and Printed Text Classification," UKSim-AMSS 19th International Conference on Computer Modelling & Simulation (UKSim), Cambridge, 2017, pp. 50-54
[7] T. Gunawan, A. Noor, and M. Kartiwi, "Development of english handwritten recognition using deep neural network," Indonesian Journal of Electrical Engineering and Computer Science, vol. 10, pp. 562–568, 05 2018.
[8] A. R. M and D. Varghese, "HABIT: Handwritten Analysis Based Individualistic Traits Prediction," no. 7, pp. 209–218, 2013.
[9] V. R. Lokhande and B. W. Gawali, "Analysis of signature for the prediction of personality traits," 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM), Aurangabad, 2017, pp. 44
[10] Rinki Kumari1, Gajanand Sharma2, Akhilesh Pandey3, ― An Intellectual System For Human Personality Identification Using Soft –Computing‖, ijates vol.no.3 Issue 10, October 2015.
[11] Nithya Lakshmi, A.Keerthana Pushpa R., ―Handwriting Analysis Based Human Personality Prediction Using Surgeon Fuzzy Model‖,International Journal of Scientific & Engineering Research Volume 8, Issue 5, 105 ISSN 2229- 5518 IJSER © 2017
[12] P. K. Grewal and D. Prashar, "Behavior Prediction Through Handwriting Analysis Behavior Prediction Through Handwriting Analysis," no. June 2012, pp. 13–17, 2016
[13] A. H. Garoot, M. Safar and C. Y. Suen, "A Comprehensive Survey onHandwriting and Computerized Graphology," 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, 2017, pp. 621-626.
[14] Aditya Chitlangiaa, G.Malathia Handwriting Analysis based on Histogram of Oriented Gradient for Predicting Personality traits using SVM INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ADVANCED COMPUTING 2019, ICRTAC 2019
[15] Min Liu, Guoli W Handwriting Analysis for Assistant Diagnosis of Neuromuscular Disorders 2013 6th International Conference on Biomedical Engineering and Informatics (BMEI 2013)