2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI) | 979-8-3503-4891-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICDSAAI59313.2023.10452641

A Deep Learning Method of Personality Prediction System Based on Signature

T. C. Sankar
Assistant Professor
Department of IT
Chennai Institute of Technology,
Chennai, India
sankartc@citchennai.net

A.R.Kavitha
Professor
Department of IT,
Chennai Institute of Technology,
Chennai, India
kavithaar@citchennai.net

C. Sankari
Assistant Professor
Department of EEE,
Chennai Institute of Technology,
Chennai, India
sankaric@citchennai.net

Abstract: Personality characteristics are **essential** ***to comprehending human behaviour and may yield insightful knowledge in a variety of fields, such as psychology, marketing, and human-computer interaction. The five factors of the big five personality's extraversion, agreeableness, conscientiousness, neuroticism, and openness are identified using a deep learning approach shown in this research that uses writing styles. Our method seeks to automatically and precisely forecast a person's personality traits by analyzing written content, such as essays, blog posts, or social media updates.***

I. INTRODUCTION

The Big Five model is widely accepted in psychology as a comprehensive framework for personality assessment. Traditional methods of personality assessment, such as questionnaires, interviews, or self-reports, can be time-consuming and prone to biases. Leveraging deep learning techniques can potentially provide a scalable and automated approach for personality trait identification.

Keywords: deep learning, personality traits, Big Five, writing styles, natural language processing, recurrent neural networks, forecast attention mechanism.

The Big Five (or Five-Factor Model) is a well-known and useful framework for describing a person's personality. It consists of five core traits that have been further broken down into several sub-factors. Through the analysis of traits like neuroticism, agreeableness, extraversion, conscientiousness, and openness to new experiences, this model offers a thorough explanation of human nature.

A. Openness to Experience: This describes people who can express their emotions without hesitation and who are drawn to novel ideas, artistic creations, and exciting adventures. This group of people is typically evaluated along a continuum that runs from constancy to curiosity.

Conscientiousness: This refers to people who are dependable, have a preference for carefully thought-out actions, and have an orientation toward achieving results. People are rated along a continuum from being well-organized to being careless within this spectrum.

B. **Extraversion:** People with extraversion are talkative, like being around other people, are assertive, and easily express happy feelings. People are rated according to how gregarious or reclusive they are within this framework. Sociability is a relevant and related term for extraversion.

C. **Agreement**: Nice people are those who are helpful and calm, and who show empathy rather than suspicion. The binary characteristics of compassion and detachment are used to evaluate people in this dimension. The word "amiability" has the same meaning as "agreeableness."

D. **Neuroticism:** This personality trait is typified by a lack of emotional stability and control, an inclination to become upset or nervous quickly, a vulnerability to depression, and other unpleasant emotions. Using this scale, individuals are ranked according to their level of confidence or lack thereof. One important concept connected to neuroticism is emotional resilience.

Data Collection and Pre-processing:

Dataset

Based on results from the aforementioned study, it is clear that widely available standard datasets, whether commercial or public, do not include contributors' identification labels along with crucial demographic annotations and emotional status data. This observation inspired the creation of a custom dataset for experiments designed to address particular issues identified in the research gaps. Individuals of various age ranges provide samples of their handwritten signatures. These samples include English, Hindi, Kannada, Marathi, and scripts in multiple languages.

• Ten signatures were recorded on separate white A4 paper sheets with ballpoint pens in the colors blue or black in order to reduce the geometric differences between each person's signature. Then, the scanned papers with the sample signatures were recorded and scanned at 300 DPI using the EPSON DS1630 color scanner. The final dataset consists of 4,790 handwritten signatures from 479 subjects (229 female volunteers and 250 male participants).

Handwriting experts used an empirical estimate approach to divide offline signature samples into 13 categories

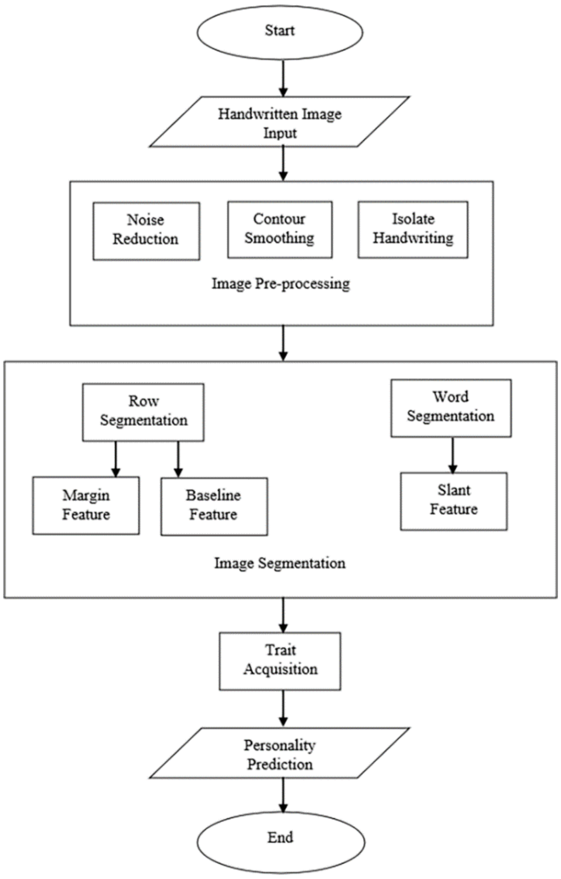
[23]. Examples of these classes include Presumptive Angle, Ascending Stroke, Complicated/Negative, Curved Slant, Descending Stroke, Garlands, Legible, Middle Stroke, Overlapped, Shell/Camouflage, Signature with additional text, Signature with Underline, and Signature with unnecessary Dots. For both male and female students in these seminars, personality descriptions are given. Tables 1(a) and 1(b) provide detailed examples and personality descriptions for the male and female participants, respectively.

Signature With More Text		• It represents a person's dependent streak and efficiency.
UNDERLINE		• Having unique idea and thinking, need support to make decisions and have reliability in the lead
Unnecessary Dots		• It represents reflective individuals with verifying prudence, • Desire of exceeding in a perfectionist.

To be trained, our deep learning model needs a sizable dataset of written text samples along with the labels of corresponding personality traits. Numerous websites, such as social media platforms, publicly accessible blogs, and online forums, can be used to collect data. Ensuring data anonymization and obtaining participant consent when necessary are crucial. Pre-processing the collected text data must include tokenizing, normalizing, and cleaning the text to remove noise and extraneous information.

Table 1(b): Signature type and description for female subjects

Signature Type	Signature Image Sample	Description
Predominant Angle Signature		• It depicts aggression, • Conclusive and reluctant to change and diversity. Inability to adjust.
Ascending Stroke signature		• It is sign of ambition, optimism creativity and vitality
Complicated/ Negative Signature		• It represents distrustful personalities, • Very conflicting.
Curved Slant signature		• More Social, extroverted, affectionate and sensual with a tendency to laziness. • It indicates flexibility cordiality kindness and ability to adjustment.
Descending Stroke signature		• The person is skeptic, Pessimistic or the person is in a little struggle and feeling down
Garlands		• It represents more predisposition to extroversion, transmits spontaneity kindness and compression. • Diplomatic skills that find the exit to any situation.
Legible		• It reflects equilibrated personality with clear goals and sense of duty and responsibility.
Middle Stroke		• It represents the possessiveness of a person.
Overlapping		• It represents the invasive feelings of a person
Shell/ Camouflage Signature		• It creates self-limiting beliefs. • A person often gets defensive in public and he/she often gets defensive, over protective.
Signature With More Text		• It represents a person's dependent streak and efficiency.
UNDERLINE		• Having unique idea and thinking, need support to make decisions and have reliability in the lead
Unnecessary Dots		• It represents reflective individuals with verifying prudence, • Desire of exceeding in a perfectionist.



II. INITIAL PROCESSING

The second part of the prediction system is called pre-processing. Spurious noise may be present in the scanned handwritten signature images. The technique mentioned above The process of improving an image's quality is called pre-processing. Greyscale image created from the input colour image. To resize the signature images with an aspect ratio of 150x150 while preserving the original signature's shape, normalization is performed, as illustrated in Figure

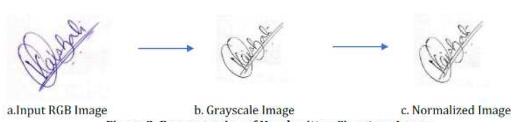


Figure 2. Preprocessing of Handwritten Signature Image

III. EXTRACTION OF FEATURES

Capturing the necessary shape information from the Handwritten Signature image and converting it into a feature vector is the process of feature extraction. A few feature extraction techniques include the following: geometric and shape features (such as area, aspect ratio, maximum and minimum axis lengths, eccentricity, horizontal and vertical projection, Euler number, and bounding box) and local directional patterns (LDP) and local phase quantization (LPQ).After being extracted, these features are categorised to forecast a person's personality and assign them to a specific class.

Classification :

Classification stage is the final step of the personality prediction system. After the feature extraction, the labelled feature vectors are used to train the model, which is then used for classifying the test feature vectors. In the experiments, different classification techniques were used, but the results were not satisfactory. To improve accuracy, we applied the ensemble classifier, namely, Random Forest Classifier that yielded relatively good accuracy.

By combining several different classification techniques or one technique with multiple parameters, the Random Forest Classifier is essentially a composite of tree-structured classifiers. This ensemble learning technique, which is primarily used for pattern recognition classification and regression, combines classifiers h1(x), h2(x),..., hK(x), which are independent identically distributed random vectors. For the class that is most common at input x, each tree provides one unit of vote. The description that follows outlines the methodology.

Main Algorithm:

Input: Offline Handwritten signature image

Output: Personality Prediction, i.e., predicted class

Step 1: Provide an image of a handwritten signature for input.

Step 2: Implement pre-processing on the input image, including the conversion to greyscale and normalization of the greyscale image to a dimension of 150x150.

Step 3: calculate individual local features such as Local Directional Patterns, Local Phase Quantization, and Shape features. Subsequently, integrate these features through a fusion process for a comprehensive representation.

Step 4: Employ the Random Forest Classifier to categorize the amalgamated features using Random Forest Classifier, KNN, and SVM.

Step 5: Optimize hyper parameters for each classifier, If necessary, using the validation set.

Step 6: Output the final predicted personality class based on the model's decision.

IV. RESULT:

$$\text{Precision}=\frac{TP}{TP+FP}$$

$$\text{F\_Score}=\frac{2*Precision*Recall}{Precsion+Recall}$$

$$\text{Recall}=\frac{TP}{TP+FN}*100$$

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN}*100$$

True positives are represented by TP in this context, true negatives by TN, false positives by FP, and false negatives by FN.

The graphical representation of the feature fusion experiment results is shown in the figure.

V. CONCLUSION

The proposed deep learning method offers a promising approach for identifying Big Five personality traits using writing styles. By leveraging large-scale text data and powerful neural networks, this method can automate personality assessment and potentially provide valuable insights across various domains. However, ongoing research is required to further refine and improve the accuracy, interpretability, and ethical considerations of this approach.

REFERENCE

[1] Pathak, A. R., Raut, A., Pawar, S., Nangare, M., Abbott, H. S., & Chandak, P. (2020). Personality analysis through handwriting recognition. *Journal of Discrete Mathematical Sciences and Cryptography*, *23*(1), 19–33. https://doi.org/10.1080/09720529.2020.1721856.

[2] Remaida, A., Moumen, A., El Bouzekri El Idrissi, Y., Abdellaoui, B., & Harraki, Y. (2021). The use personality tests as a pre-employment tool: A comparative study. *SHS Web of Conferences*, *119*,05007. https://doi.org/10.1051/shsconf/202111905007.

[3] Chaudari,k., & Thakkar, A.2019.Survey on handwriting-based personality trait identification. Expert Systems with Applications,124,284-308.https://doi.org/10.1016/j.eswa.2019.01.028.

[4] Haridas, A., S, S., R, A., N R, M., & Muralidharan, R. (2021).Personality Prediction based on Handwriting using CNN & MLP [Review of Personality Prediction based on Handwriting using CNN & MLP]. Volume 09(Issue-07). https://doi.org/10.17577/IJERTCONV9IS07019

[5] Joshi, P., Agarwal, A., Dhavale, A., Suryavanshi, R., & Kodolikar, S. (2015).Handwriting Analysis for Detection of Personality Traits using Machine Learning Approach. *International Journal of Computer*

Applications,*130*(15),4045.
https://doi.org/10.5120/ijca2015907189.

[6] Katiyar, S., Walia, H., & Kumar, S. (2020). Personality Classification System using Data Mining. *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*. https://doi.org/10.1109/icrito48877.2020.9197803.
[7] Gavrilescu, M., & Vizireanu, N. (2018). Predicting the Big Five personality traits from handwriting. *EURASIP Journal on Image and Video Processing*, *2018*(1). https://doi.org/10.1186/s13640-018-0297-3
[8] Biswas, K., Shivakumara, P., Pal, U., Chakraborti, T., Lu, T., &Ayub, M. N. B. (2022). Fuzzy and genetic algorithm based approach for classification of personality traits oriented social media images. *Knowledge-Based Systems*,*241*,108024.
https://doi.org/10.1016/j.knosys.2021.108024.
[9] Esposito, A., Amorese, T., Buonanno, M., Cuciniello, M., Esposito, A. M., Faundez-Zanuy, M., Likforman-Sulem, L., Riviello, M. T., Troncone, A., & Cordasco, G. (2019). Handwriting and Drawing Features for Detecting Personality Traits. *2019 10th IEEE International Conference on Cognitive Info communications (CogInfoCom)*.
https://doi.org/10.1109/coginfocom47531.2019.9089985.