Pune Institute of Computer Technology
Dhankawadi, Pune

A SEMINAR REPORT
ON

AUTOMATED HANDWRITING ANALYSIS TO PREDICT HUMAN BEHAVIOR USING MACHINE LEARNING

SUBMITTED BY

Name : Aryan Patil
Roll No. : 31353
Class: TE-3

Under the guidance of
Prof. A.A.Chandorkar

DEPARTMENT OF COMPUTER ENGINEERING
Academic Year 2021-22

DEPARTMENT OF COMPUTER ENGINEERING

Pune Institute of Computer Technology
Dhankawadi, Pune-43

CERTIFICATE

This is to certify that the Seminar report entitled

"Automated Handwriting Analysis to predict Human Behavior using Machine Learning"

Submitted by
Aryan Patil Roll No. : 31353

has satisfactorily completed a seminar report under the guidance of Prof. A.A.Chandorkar towards the partial fulfillment of third year Computer Engineering Semester I, Academic Year 2021-22 of Savitribai Phule Pune University.

Prof. A.A.Chandorkar
Internal Guide

Prof. M.S.Takalikar
Head
Department of Computer Engineering

Place:Pune
Date: 09/11/2021

ACKNOWLEDGEMENT

It is my pleasure to present report on "Automated Handwriting Analysis to predict Human Behavior using Machine Learning". First of all, I would like to thank our Seminar Coordinator Prof. D.D.Kadam, Head of Department Prof. M.S.Takalikar and Principal R.Shreemati for their encouragement and support.

I would also genuinely express my gratitude to my guide Prof. A.A.Chandorkar, Department of Computer Engineering for her constant guidance and help.She has constantly supported me and has played crucial role in completion of this report. Her motivation and encouragement from beginning till end to make this seminar a success.

Last but not the least I would thank all the faculty,my parents and friends who have helped me.

Contents

List of Tables

List of Figures

Abstract

Handwriting is unique and exclusive to an individual and provides important insights into human behavior and physical, mental and emotional states. Graphology is the science of analysis of handwriting to determine personality traits by evaluating features of handwriting. Features used for such analysis includes slant of alphabets, spacing, pen pressure, baseline, word size, slope, etc. Graphology has a wide scope including recruitment, psychology tests, medical diagnosis, forensic as well as criminology.

While human intervention in Handwriting Analysis is effective, it is costly and prone to fatigue. To automate the process of analysis, making it more reliable and less prone to fatigue of graphologists, a machine learning approach is considered for behavior recognition using Handwriting analysis. The proposed methodology automates this analysis by computerized image processing, feature extraction and mapping of features to determine personality traits.

Keywords

Personality Traits, Graphology, Image Processing, Feature Extraction, Machine Learning.

1 INTRODUCTION

Handwriting Analysis or Graphology is the scientific study of handwriting patterns, which identifies the psychological state of a person and evaluate characteristics of their personality. Our personalities affect the way our handwriting develops after being taught to write. This is because handwriting is the pattern of our psychology expressed symbolically on a page and these symbols are unique. These unique patterns provide insight in human behavior as well as physical, mental and emotional states of the person. Each handwriting pattern is associated with a distinctive neuromuscular movement associated with the brain patterns and occurs unconsciously while writing. Thus, the science of graphology maps these patterns with certain behavioral and personality traits. Behavioral Analysis through Graphology is used for personality assessment in areas of human activity where people interact as well as fields including recruitment, security checking, career guidance, personality profiling in law, criminology and forensic analysis. Graphology is also useful in psychiatric diagnosis of Dyslexia, Schizophrenia and Parkinson's disease.

Prime traits and patterns used in handwriting analysis by graphologists includes loop of letters L and e and y, height and crossing over of letter t, pen pressure, margin gaps, slant of letters, baseline, writing regularity, spacing between words and word size. Each characteristic of strokes in handwriting is useful in detecting a unique personality. These patterns identified in handwriting are sequentially analyzed to deduce the results i.e., personality traits.

The proposed system automates the process of handwriting analysis by using image processing, feature extraction and predicting individuals' personality based on the features. The average time to perform manual preliminary analysis by a graphologist is 4-6 minutes and the analysis is also prone to fatigue. Thus, the system reduces error margin and is also economical and time saving. The dataset used for training is the 'IAM Handwriting Dataset'. Total of seven features which are used for personality detection are t op margin, pen pressure, baseline angle, letter size, line spacing, word spacing and slant angle. These features are passed through a rule-based system to categorize into personality traits. Finally, Machine Learning approach like SVM, KNN and Decision Tree are implemented for improving efficiency of classification of personality traits.

2 MOTIVATION

In recent times, automation with the help of machine learning is being used in various professions to improve efficiency and reduce workload. Such systems have been observed to have better accuracy as they are not open to exhaustion along with being economical and time saving.

Today, process of handwriting analysis is offline, and graphologists manually analyze handwriting on paper. This process requires 4-6 minutes for preliminary analysis and around 30 minutes for detailed behvioral analysis. Analyzing a large number of handwriting samples by a graphologist will be time consuming and prone to fatigue.

The inference of these results should be considered carefully because of the sensitive nature of the subject. Thus, for an effective and efficient analysis, an automated system using machine learning for graphology would be useful for fast and accurate result.

3 LITERATURE SURVEY

The Following table shows the literature survey by comparing techniques propose in various references:

Table 1: Literature survey

No.	Paper Title	Features	Method used	Gap
1	Automated Human Behavior Prediction through Handwriting Analysis	Personality traits based on baseline, pen pressure, letter 't', letter 'y' loop, slant of writing	evaluation of baseline uses polygonalization method and pen pressure uses grey-level threshold value height of stem of 't' and slant of writing is calculated using template matching results are input to a rule-based system for prediction	Larger Dataset should be used. Other features like word-spacing,margins, word size could be used as input for personality detection
2	Handwriting Analysis for Detection of Personality Traits using Machine Learning	Personality traits based on baseline, letter 't', margin slant of writing	Methods for feature extraction include polygonalization, thresholding algorithm,template matching.These feature vectors are generated with an Artificial Neural Network. KNN classifier is used for final personality classification	Dataset used has only 100 samples Other features including concavity of letters could be used for better accuracy.Other Supervised Machine Learning models can be used to check for better results
3	Handwriting Personality Recognition with Machine Learning: A Comparative Study	comparative study of various approaches using Deep Learning techniques in handwriting personality prediction	Along with study of seven different papers,proposed a general outline of their own model for predicting the Big five personality traits	More elaborate description of proposed model should be given

No.	Paper Title	Features	Method used	Gap
4	Personality Prediction Based on Handwriting using Machine Learning	A Web-Application used for determining Personality Traits including optimistic, pessimistic, independent, pragmatic, balanced, shy,etc. based on features including word spacing, slant of letters, baseline.	Uses a CNN model for Feature Extraction from input images of handwriting samples.Returns different Personality Traits along with its percentage where sum of each trait is 100	The CNN model developed requires high computation power and time for training, hence tuning the parameters of the model is required. System can switch from offline samples to online handwriting inputs using touch-screen devices
5	Automated Career Guidance using Graphology, Aptitude test, and Personality test	Features include word spacing, Capital size, letter 'I', average size of middle zone letters, margin spacing, slant of letters.	Provides description for individual Algorithms for Image preprocessing, Feature Extraction, word spacing calculation etc. The results are then sent to a model which uses this handwriting analysis along with results of Aptitude Tests and Psychometric tests to give a final career options report	Product can be scaled to industry-domain specific models with help of domain expert.

4 A SURVEY ON PAPERS

4.1 Automated Human Behavior Prediction through Handwriting Analysis

This paper proposes a method to identify personality traits based on handwriting revealed by pen pressure, baseline, letter 't', lower loop of letter 'y' and slant of handwriting. These features are the input of a rule base which predicts the personality traits of the individual.

Evaluation of pen pressure is by grey-level threshold value and evaluation of baseline using polygonalization. Height of letter t and slant of letters is calculated by template matching. Generalized Hugh Transformation method is used for the shape of loop of letter y. For the data, 120 different handwriting samples were obtained from 120 people who were asked to write a simple paragraph which covered almost all alphabets of the English language.

4.2 Handwriting Analysis for Detection of Personality Traits using Machine Learning

This paper adopts the method of Artificial Neural Network with back propogation to create a trained dataset. The features used for handwriting analysis are Margin space, letter slant, baseline and stem of lower case letter 't'.

Polygonalization, thresholding algorithm and template matching are the methods used to extract these features from handwriting sample. A supervised machine learning model like KNN-classifier is trained using the trained dataset for the prediction of personality traits. 100 handwriting samples examined by professional Graphologist, were used for the training data.

4.3 Handwriting Personality Recognition with Machine Learning:A Comparative Study

This paper presents a comparative study of different approaches and methods using Deep Learning algorithms for personality recognition using handwriting analysis, across seven different papers. A range of features from handwriting text were introduced including Page margin, word space, line space, connecting strokes, V to H ratio, word size, lower case letters 'f' and 't'.

This study also presented use of different Deep Learning algorithms for handwriting personality recognition including Multi-Layer Perceptron with Back Propagation Algorithm(MLP), Back Propagation Neural Network(BPNN), Feed Forward Neural Network with Back Propagation(FFNN) and Convolutional Neural Network (CNN). The paper also provided a generic workflow algorithm for handwriting analysis using deep learning approach.

4.4 Personality Prediction Based on Handwriting using Machine Learning

The proposed workflow in this paper includes the use of Convolutional Neural Networks for grouping of similar temperament traits based on features extracted from handwriting sample after image preprocessing. The proposed method is a Web-Application which takes input of the handwriting sample from the user and returns their personality traits predicted. Image preprocessing methods uses Black and White with Contrast filter, Noise Removal and Thresholding while CNN training includes filtering, pooling and ReLU activation. The paper also provides a better understanding of mapping of personality traits with extracted features. It returns the personality traits as percentages where sum of each trait is 100

4.5 Automated Career Guidance using Graphology,Aptitude test,and Personality test

This paper provides a novel approach to automate career guidance based on integration of analysis using Aptitude Test, Psychometric Test (Myers-Briggs Type Indicator – MBTI) and handwriting analysis. The Graphology module focuses on subconscious way of mapping personality. Features extracted are slant, size, speed, word spacing, line spacing, upper zone, pressure, capital size, lower loops, left and right margin. A report is generated which provides bar charts, pie charts for aptitude scores obtained by the applicant. The personality type and description is provided. Finally the suggested jobs/career is mentioned in the report.

5 PROBLEM DEFINITION AND SCOPE

5.1 Problem Definition

To propose a system that automates handwriting analysis to predict personality traits using Machine Learning.

5.2 Scope

Manual preliminary handwriting analysis for human behavior prediction, even by an expert, requires around 4 to 6 minutes for one sample of handwriting text. Thus, to perform handwriting analysis on a larger database will be time consuming and graphologists will be exposed to fatigue. Handwriting analysis is now used in a lot of fields and requires higher precision and timely result to be used in any domain. .

Currently, such analysis is performed offline without the help of automation. For the above purpose tools to automate graphology using machine learning models to save time and for accurate results, are under research to support the work of graphologists.

6 FEATURES USED FOR HANDWRITING ANALYSIS

6.1 Baseline

In handwriting patterns, baseline is the line, real or imaginary, established by writer, along which letters rest. Graphologists use Baseline to predict emotional stability and is categorized as even (levelled), erratic, sloped up and sloped down.

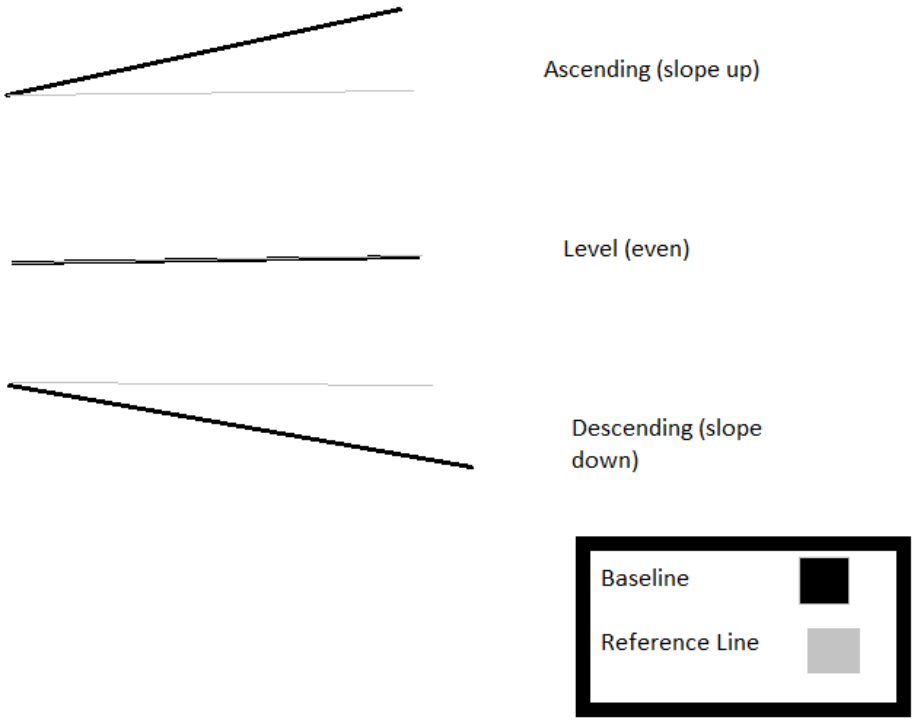

Figure 1: Baseline Features

Table 2: Baseline Angle

No.	Baseline	Personality
1	Ascending (slope up)	Optimistic
2	Level	Balanced
3	Descending (slope down)	Pessimistic

6.2 Top Margin

Top Margin is the gap between the start of the page and beginning of the first sentence on the page. Top margin is associated with the personality trait of modesty. Medium to high top margin is one of the factors which indicates the individual to be modest in nature.

6.3 Pen Pressure

The writing pressure is an important feature in handwriting analysis. The amount of pressure exerted on the page while writing reflects our depth of feeling or emotional intensity. It is a parameter to show our mental resistance ability and will power. Based on pressure applied, a writer is classified as light writer, medium writer, heavy writer.

Table 3: Writing Pressure

No.	Pen Pressure	Personality
1	Light Writer	This writer has a low depth of feelings. Emotional experience do not have a lasting effect on the writer and shows low will power.
2	Medium Writer	Writer has an average emotional stability and will power. Writer is moderately affected by emotionally traumatic experience
3	Heavy Writer	Forgives but doesn't Forget. Has high depth of feelings

6.4 Letter Size

Size of letters in handwriting is also associated with metal strength and also specific to how a person behaves in a society and how much the writer wants himself to be noticed. Size of letters are divided into small, medium and large handwriting.

Table 4: Size of Letters

No.	Letter Size	Personality
1	Small Size	Weaker mental strength. Often shy and withdrawn, however are meticulous.
2	Medium Size	Has good mental strength and will power. Generally well adjusted in the society.
3	Large Size	Has good mental strength and will power. Outspoken and People-oriented behavior.

6.5 Line Spacing and Word Spacing

Line Spacing and word spacing reflects individuals' openness towards criticism, flexibility and personal harmony. Evenness represents a stable mind and a medium level of spacing indicates the person to be self-harmonious and flexible towards suggestions.

6.6 Slant of Letters

Slant of letters is in relation to emotional direction and stability of a person. More erratic the slant, stability is less. The slant of letters is taken in consideration of baseline.

$$\theta = \tan^{-1} \frac{y_2 - y_1}{x_2 - x_1}$$

Table 5: Slant of Letters

No.	Slant	Personality
1	Extremely Left	Fear of future,defensive, early rejection.
2	Extremely Right	Lack of self control, impulsive, unrestrained low tolerance.
3	Left	Difficult in expressing emotions, independent, not sypethatic.
4	Right	Expressive, extrovert, Future Oriented.
3	Vertical	Trusts the head, Works independently

7 METHODOLOGY

7.1 Workflow

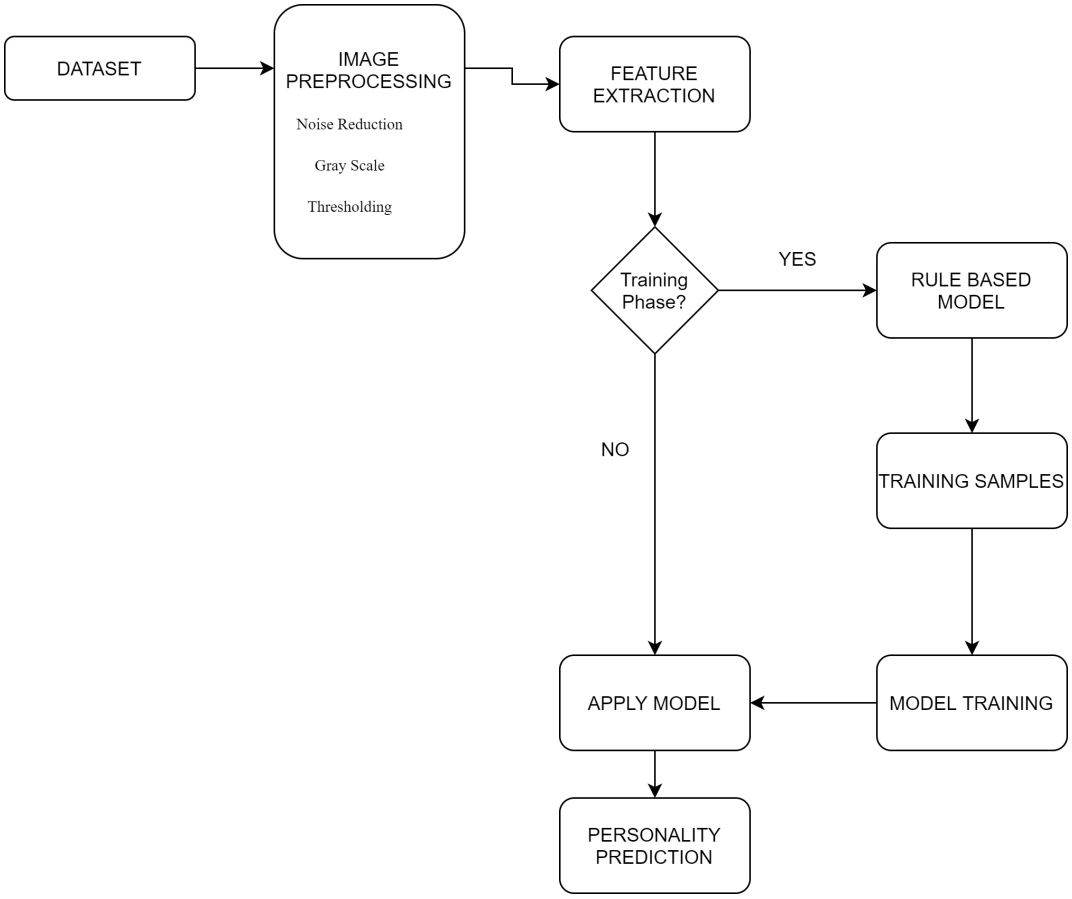

Figure 2: Overall Automated Handwriting Analysis: Workflow

7.2 Implementation and Result

After input handwriting sample image is preprocessed, seven features are extracted from it for further behavior prediction. For the Baseline angle, Polygonalization technique is used. In this technique a closed polygon is drawn around the handwriting text. Using the coordinates of this polygon, the slope is determined which is associated with the baseline angle. Thresholding technique is an image segmentation method in which a gray scale image is converted to a binary image. This method is used to find out pen pressure feature for graphology. Horizontal and vertical project along with thresholding are used to calculate other features including slant height, letter size, word spacing, line spacing and top margin.

For generating the Training Dataset, this feature matrix is further passed through a rule-based system which links different personality traits with corresponding features associated with it. Further, the obtained result is used to train supervised machine learning models to get the final behavior traits of the individual.

Table 6: Associating Extracted Features with Personality Traits

No.	Personality Trait	Associated Features
1	Emotional Stability	Baseline Angle, Slant Angle.
2	Mental Strength, Will Power and Emotional Depth	Letter Size, Pen Pressure
3	Modesty	Top Margin, Letter Size
4	Personal Harmony, Flexibility and Openness to Criticism	Line Spacing, Word Spacing
5	Lack of Discipline	Slant Angle, Top Margin
6	Poor Concentration	Letter size, Line Spacing
7	Weaker Interactive skills	Letter Size, Word Spacing
8	Social Isolation	Line Spacing, Word Spacing

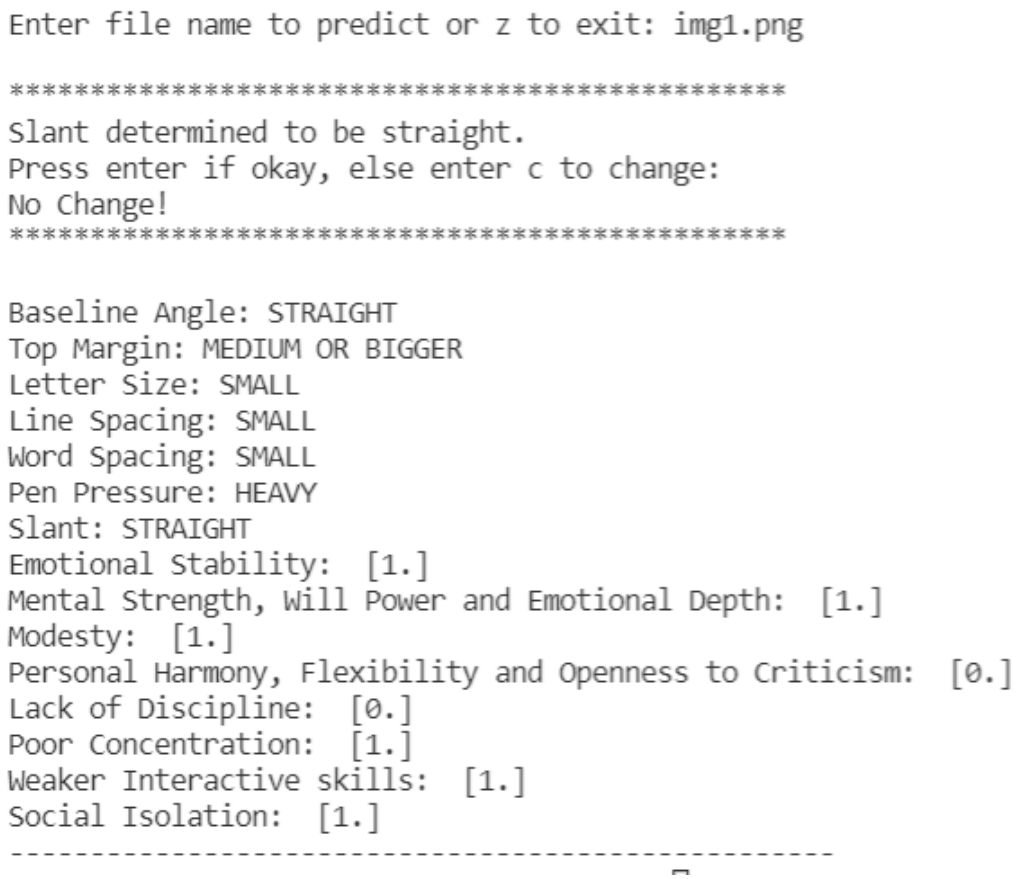

Figure 3: Automated Handwriting Analysis: Implementation Result

8 CONCLUSION

Handwriting is a unique pattern which expresses our mental and emotional states. Handwriting Analysis is used in different domains ranging from Career guidance to medical diagnosis and hence its implications must be realized carefully considering its sensitive nature. On input of test handwriting sample image, seven different features, baseline angle, top margin, pen pressure, slant angle, letter size, word and line spacing, are extracted for handwriting analysis. These features are passed to the trained model to determine human behavior.

The proposed model can be used to complement the work of handwriting analysis by a trained graphologist for faster and accurate results.

References

[1] H. N. Champa and K. R. AnandaKumar, "Automated Human Behavior Prediction through Handwriting Analysis," 2010 First International Conference on Integrated Intelligent Computing, 2010, pp. 160-165, doi: 10.1109/ICIIC.2010.29.

[2] 1. P. Joshi, A. Agarwal, A. Dhavale, R. Suryavansi, and S. Kodolikar, "Handwriting Analysis for Detection of Personality Traits using Machine Learning Approach," Int. J. Comput. Appl., vol. 130, no. 15, pp. 40-45, 2015, doi: 10.5120/ijca2015907189

[3] A. Remaida, A. Moumen, Y. E. Bouzekri El Idrissi and B. Abdellaoui, "Handwriting Personality Recognition with Machine Learning: A Comparative Study," 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), 2020, pp. 1-6, doi: 10.1109/ICECOCS50124.2020.9314529.

[4] N. Lemos, K. Shah, R. Rade and D. Shah, "Personality Prediction based on Handwriting using Machine Learning," 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 2018, pp. 110-113, doi: 10.1109/CTEMS.2018.8769221.

[5] A. Anand, D. Patil, S. Bhaawat, S. Karanje and V. Mangalvedhekar, "Automated Career Guidance Using Graphology, Aptitude Test and Personality Test," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697642.