Characteristics Detection from Handwriting Using Machine Learning

Submitted in partial fulfillment of the requirements

For the degree of

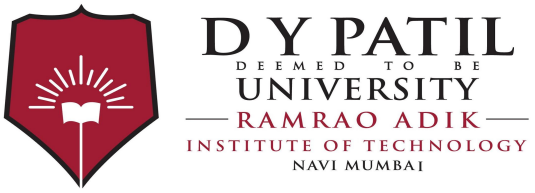
Bachelor of Engineering
in
Information Technology

by

ROHAN KADAM 17IT1009
MAHESH MAHAWARKAR 17IT1006
RUSHABH MAHAWARKAR 17IT1014

Supervisor

Prof. (Mrs. Puja Vakhre)



Department of Information Technology

Dr. D. Y. Patil Group's

Ramrao Adik Institute of Technology

Nerul, Navi Mumbai 400706.
(Affiliated to University of Mumbai)
(2021)

D Y PATIL
DEEMED TO BE
UNIVERSITY
RAMRAO ADIK
INSTITUTE OF TECHNOLOGY
NAVI MUMBAI

Ramrao Adik Institute of Technology

(Affiliated to the University of Mumbai)
Dr. D. Y. Patil Vidyanagar,Sector 7, Nerul, Navi Mumbai 400 706.

CERTIFICATE

This is to certify that, the Project-I titled

"CHARACTERISTICS DETECTION FROM HANDWRITING USING MACHINE LEARNING"

is a bonafide work done by

ROHAN KADAM
MAHESH MAHAWARKAR
RUSHABH MAHAWARKAR

and is submitted in the partial fulfillment of the requirement for the degree of

Bachelor of Engineering
in
Information Technology
to the
University of Mumbai

Supervisor

Prof. (Mrs. Puja Vakhre)

Project Coordinator	Head of Department	Principal
(Mrs. Reshma Gulwani)	(Dr. Ashish Jadhav)	(Dr. Mukesh D. Patil)

Project Report Approval for B.E.

This is to certify that the project entitled *"Characteristics Detection from Handwriting using Machine Learning"* is a bonafide work done by *Rohan Kadam, Mahesh Mahawarkar and Rushabh Mahawarkar* under the supervision of *Prof. Mrs. Puja Vakhre***.** This project has been approved for the award of *Bachelor's Degree in Information Technology, University of Mumbai.*

Examiners:

1. . . . . .

2. . . . . .

Supervisors:

1. . . . . .

2. . . . . .

Principal:

. . . . . .

Date :

Place :

Declaration

I declare that this written submission represents my ideas in my own words and where other's ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

ROHAN KADAM	17IT1009	Signature
MAHESH MAHAWARKAR	17IT1006	Signature
RUSHABH MAHAWARKAR	17IT1014	Signature

Date :

Abstract

Handwritten recognition is basically a technique in which the computer takes input in the form of an image or document and interprets handwritten input. We have used the Support Vector Machine (SVM) algorithm in our model. While writing we can control our conscious mind about what to write, but not our subconscious how to write. From a handwritten paragraph the perspective of the writer's internal mind can be discovered. It is a process of interpreting characteristics from patterns in handwriting. Graphology is the analysis of the physical characteristics and patterns of handwriting with an attempt to find the writer's psychological state at the time of writing and evaluate personality characteristics. The accuracy of handwriting analysis depends on the skills and experience of the graphologists. Handwriting analysis means predicting the behaviour from mannerism in handwriting. But the manual process of handwriting analysis is expensive; sometimes it is not precise and may lead to incorrect results. Thus for reducing the errors and improving accuracy this paper proposes an algorithm to recognize the handwriting which is Support Vector Machine (SVM).

Contents

Abstract **i**
List of Figures
List of Tables

1. Introduction **1**
1.1 Handwriting and Characteristics recognition
1.2 Problem Definition……
1.3 Scope of project……
1.4 Relevance and Motivation of project…….
1.5 Organization of the report…….
2. Review of Literature
3. Planning and Formulation…….
4. Methodology
4.1 Proposed System……
4.2 Proposed Methodology……
4.3 System Requirements……
5. Design of System
5.1 System Design……
5.2 Data Flow Diagrams……..
6. Implementation
7. Experimental Results
8. Conclusion

Literature Cited
Publication
Acknowledgement

List of Figures

Figure no.	Figure Name	Page no.
1	Gantt Chart	13
2	Sample Data	14
3	Count of Age	15
4	Count of Gender	15
5	Characteristic Analysis	15
6	Example of Affine Transformation	17
7	SVM Data flow diagram	24
8	CNN Data flow diagram	25
9,11,13,15	Sample Input	26
10,12,14,16	Sample Output	28
17	CNN Testing Graph	29

List of Tables

Table no.	Table Name	Page no.
1	Baseline Traits	18
2	Letter Size Traits	19
3	Line Spacing Traits	19
4	Word Spacing Traits	20
5	Top Margin Traits	20
6	Pen Pressure Traits	21
7	Slant of Letters Traits	22
8	Accuracy Table	29

Chapter 1

Introduction

1.1 Handwriting and Characteristics Recognition

Handwriting means a writing done by the ball pen or pencil. It is everyones physiological and psychological. Handwriting is totally different as compared to typing using a keyboard or touch screen. When a person writes he uses all his manual skills that he learned throughout his life. It is proven that writing engages our brain more thus improving our ability to recollect something that has been written. As every person has his own unique style of writing it can be used to verify the person's identity. Thus Graphology is the study of handwriting to reveal the personality of the writer. An expert can identify many characteristics of handwriting from a piece of handwriting. In graphology the content of the written text is always irrelevant, the main focus is on the style in which the writer has written the content.Some of the factors that can be identified in the handwritten text are Top Margin, Pen Pressure, Baseline Angle, Letter Size, Line Spacing, Word Spacing and Slant Angle. It checks the personality traits of the writer like Emotional Stability, Mental Energy or Will Power, Modesty, Personal Harmony and Flexibility, Lack of Discipline, Poor Concentration, Non-communicativeness, Social Isolation and lots of others. As technology is developing humans are moving more towards the automation of every field in a vision of decreasing human error and reducing the time consumption. Some of the well-developed areas are Artificial Intelligence and Machine Learning. In this paper we have considered two algorithms i.e. CNN and SVM to automate the process of predicting the character of the writer. By using the CNN algorithm in the handwriting recognition field, 2-D images can be directly input and feature extraction is done. Many experiments with CNN have seen moderately good performance. SVM is easier to understand and also it is effective where the dimensions are greater.
Through graphology the companies can do employee profiling, psychological analysis, Medical diagnosis, Criminal profiling, handwriting recognition while solving cases of homicide that involve suicide notes written by the victim or while trying to identify the sociopathic tendencies of abductors and criminals who leave behind ransom notes and many more.

1.2 Problem Definition

There are graphical features that can be analyzed through graphology like top margin, pen pressure, baseline angle, letter size, line spacing, word spacing and slant angle.[ ] It is a well-known fact that the handwriting of a human can vary from time to time, depending on one’s emotional and psychological state. irony. Emotions are the critical component of human instinct and are generally being examined in the field of neuroscience, psychology, behavior science, health care, communication, computer science, cognitive science, human decision handling, etc.

1.3 Scope of Project

Accuracy of handwriting evaluation relies upon the competencies and experience of the graphologists. But the manual process of handwriting analysis is costly, sometimes it is not accurate and may lead to incorrect results. This project has an algorithm called Support Vector Machine (SVM) which is used to recognize the handwriting. By using it we can find out the characteristics of a person like he/she is emotionally stable, lack of discipline, modesty, poor concentration etc. By using the SVM algorithm we can reduce the errors and improve accuracy of the existing models.

1.4 Relevance and Motivation of project

Digital technologies like Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) are being used to reduce the manpower efforts and time to achieve the maximum accuracy and reduce the minimal errors caused by humans. Several Machine Learning models have been developed to analyze and predict the characteristics of a person. ML is being used in applications like recognition and analysis of the handwriting in various languages like Hindi, Marathi, Urdu and many more. The model can help various NGOs, Schools, Colleges and various companies too. It can be used to know the characteristics of teenagers and school students and further it can be analyzed and the people can know their characteristics.

1.5 Organization of the report

This report is divided into eight chapters and they are as follows. Chapter 1 contains the introduction part, Chapter 2 has the literature review referred for our work, Chapter 3 displays the planning and formulation of work with the help of gantt chart, Chapter 4 elaborates the proposed methodology, Chapter 5 tells about our system design, Chapter 6 has the implementation pictures, Chapter 7 discusses the experimental results, Chapter 8 concludes all our work and some future work.

Chapter 2

2.1 Review of Literature

Characteristics of a person can be recognized on the basis of his/her handwriting. There are various techniques and algorithms used to identify the character of a person by doing the handwriting analysis. The handwriting analysis can be done by both humans and machines but the major difference is the accuracy.

In A Review Optical Character Recognition [19] authors have presented a review in detail about different Optical Character Recognition (OCR) techniques used for recognition of character from handwriting. They proposed a system in which the handwritten scanned document will be converted to a text document. For this they have included methods like neural networks, Artificial Neural Networks (ANN), SVM and various other classifiers. Another paper named Optical Character Recognition using Artificial Intelligence [5] has proposed a method in which the OCR system will train itself and help to extract the text from any image. The neural networks are used for training the machine. Results obtained by using this approach were able to recognize the character up to 100% when the image was noiseless and almost 95% in noisy images[5].

Human Behavioral Analysis Based on Handwriting Recognition and Text Processing [4] proposes an artificial intelligence system approach for recognizing characteristics of humans through text processing and handwriting recognition. MATLAB software is used for image processing and the same for handwriting analysis. Parameters like pressure, slant, baseline, dimension, and writing tone are taken in consideration for recognizing the characteristics of an individual. The accuracy rate of this method is 95%. Handwriting Recognition on Form Document using Convolutional Neural Network and Support Vector Machines (CNN-SVM) [15] proposes CNN as a strong feature extraction method applied to extract the feature of the handwritten characters and linear SVM using L1 loss function and L2 regularization used as end classifier supported the experiment results using data from NIST SD 19 2nd edition, both for training and testing, the proposed method achieves an accuracy rate better than only CNN method. A system for automatic handwriting recognition on form documents has been constructed. The identification rate achieved by the proposed method is 98.85% on numeral characters, 93.05% on uppercase characters, 86.21% on lowercase characters[15].

An Effective Approach for Emotion Detection [20] has proposed a sequence based convolutional neural network (CNN) technique which will recognize the emotion of a person. A different mechanism is applied to basically focus on the words that have more effect on the classification of the features that should be attended more and that mechanism is called attention mechanism. The proposed framework correctly detects the emotion and the accuracy is 92%. EMOTHAW: A Novel Database for Emotional State Recognition From Handwriting and Drawing [16] proposes a database dedicated to a new procedure that will recognize the emotional states on the basis of handwriting and drawing too. By using a digital tablet they have been given seven tasks like drawing pentagons,circles, one sentence in cursive writing and many more that should be completed on the basis of which the emotional state will be recognized. The records consist of main factors which are responsible for the output that are pen position, time stamp, pressure, pen azimuth and altitude which are taken under consideration to

recognize the emotional state. Random Forest Machine Learning Algorithm is used in this procedure which is used to recognize the emotional state of people.
Handwritten Recognition Using SVM, KNN and Neural Network [8] shows KNN and SVM predict the dataset correctly except for MLP Neural Network that's some mistake to predict the number 9. This is because for KNN and SVM it predicts directly from the feature extraction. But for MLP, it is a nonlinear function. So it is more suitable for learning non-linear models. And MLP with hidden layers has a non convex loss function where there exists quite one local minimum. Online Handwriting Recognition using Support Vector Machine [21] proposes a method in which a hybrid of SVM/HMM OHR systems is developed. Some preliminary experimental results of using SVM with the RBF kernel on IRONOFF, UNIPEN and IRONOFF-UNIPEN character databases are provided. The concluded result was that character recognition rate is higher in SVM because of structural risk minimization implementation by maximizing margin of separation in the decision function.

Chapter 3

3.1 Planning and Formulation

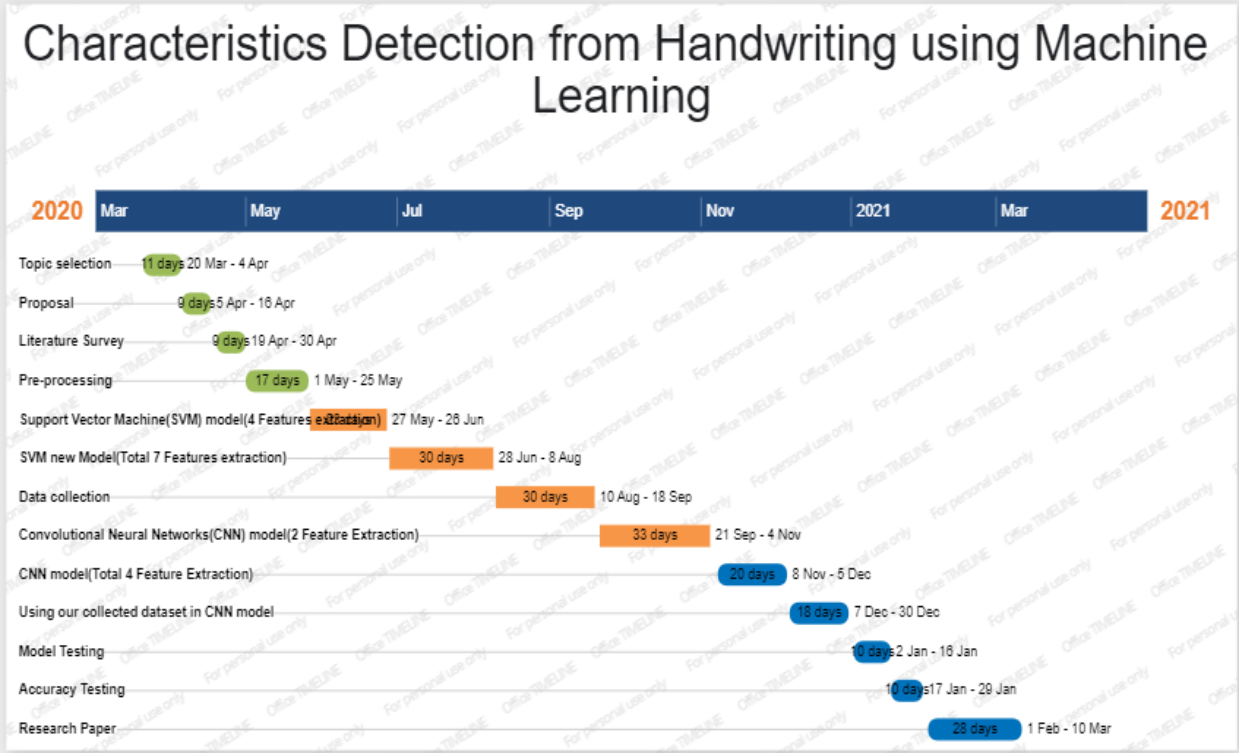


Figure 1 : Gantt Chart

Chapter 4

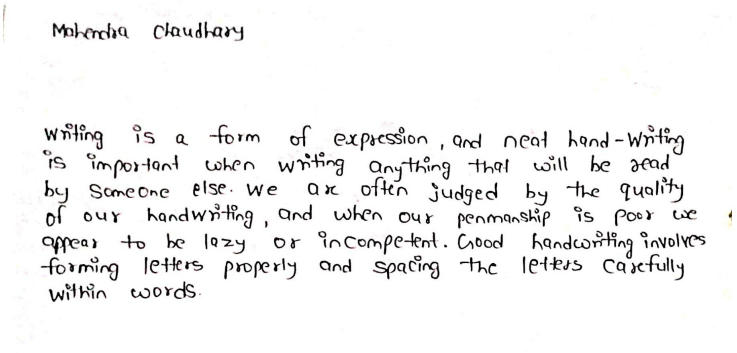
Methodology

4.1 Proposed system

For SVM algorithm features extracted are base line, pen pressure, top margin, word spacing, letter size, slant angle and line spacing. The algorithm is trained to predict the writer's personality with the help of the combination of features extracted.
For CNN algorithm features extracted are ascending slant angle, descending slant angle, left margin, right margin. The algorithm is then trained to predict the writer's personality with the help of extracted features.

4.2 Collecting Dataset

We created a data set with the help of google form. Where we have asked the Name, Age and Gender of the person and given a paragraph that should be written on a blank paper and he/she has to upload it. Then we have asked eight basic questions to the person and he has to attempt that questionnaire. The questions are related to the features that we have extracted using our model on the written paragraph. The person has to rate himself in each question out of five. By using the rating system we can find out the person's inclination towards each characteristic and further this result will be compared with our model's output so that the accuracy can be found out and necessary changes can be done in future. The paragraph has to be written on a blank paper so we can measure the baseline and slant of the handwriting. We have circulated this form to our friends, relatives and colleagues. We have collected 200 peoples data by using the google form till now. The sample of data collected are shown below.


Mahendra Chaudhary

Writing is a form of expression, and neat hand-writing is important when writing anything that will be read by someone else. We are often judged by the quality of our handwriting, and when our penmanship is poor we appear to be lazy or incompetent. Good handwriting involves forming letters properly and spacing the letters carefully within words.

Figure 2 : Sample Data

Also we have data set from the IAM Handwriting Database of Research Group on Computer Vision and Artificial Intelligence INF, University of Bern, Switzerland. The data set contains 1538 pages of scanned text for which 657 writers contributed samples

of their handwriting. Each handwriting sample is labelled with the corresponding psychological traits by manually studying each document.

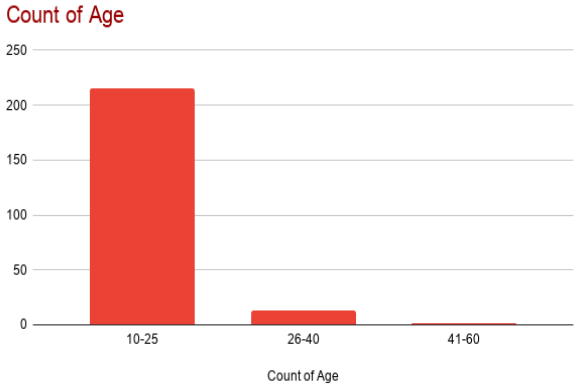


Figure 4 : Count of Age

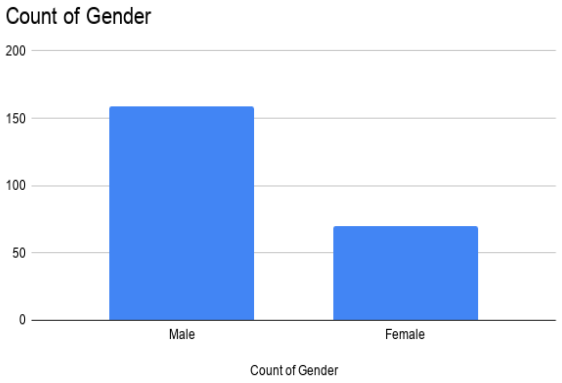


Figure 5: Count of Gender

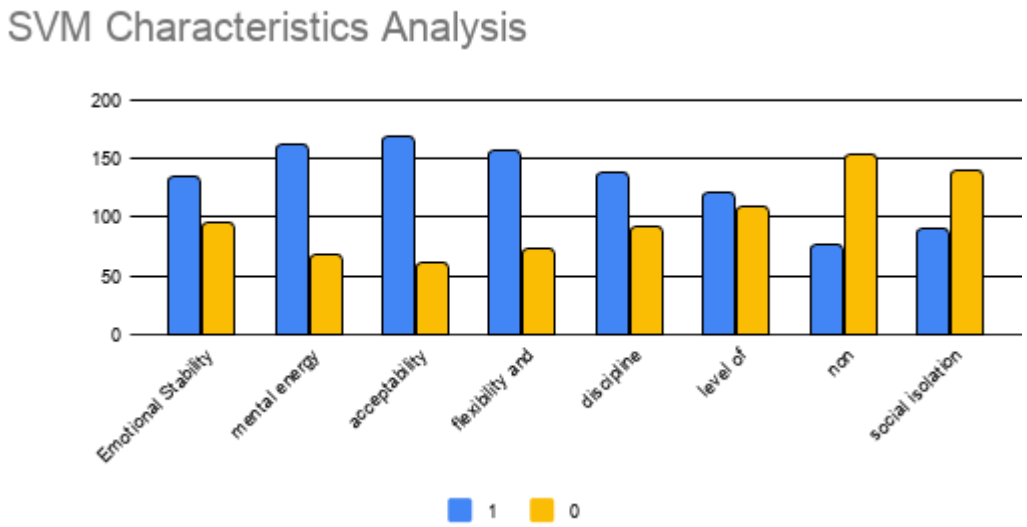


Figure 6 : Characteristic Analysis

4.3 Proposed Methodology

Graphology is the study of all graphic movements. In addition to handwriting, the graphologists study doodles, drawings and sculptures. These studies give insight into the physical, mental and emotional states of the writer. Our handwriting is unique. Whether you write it together with your hand or feet or mouth, it's all an equivalent. The handwriting isn't done by your hand or feet but by your brain. Therefore, some call handwriting ‘brain writing'. Your psychological functions are revealed in your handwriting.

This chapter will firstly present a detailed explanation of the image processing methods which are used in the project. Support Vector Machines and Convolutional Neural Network will be used for classification and are discussed briefly afterward.

4.3.1 Image Resolution and Cropping

The original images also have a very big resolution by default. The aim of pre-processing is to make the image data suitable for feature extraction by filtering unwanted attributes, enhancing the quality, and performing transformations. Adobe Photoshop is used to automatically crop out the left and right margins, resize all the images with 850 pixels width and perspective height, and save the images in PNG format. PNG format is used instead of JPEG because the format is a lossless format as compared to JPEG image format and is more suitable for storing text images, printed or handwriting.

4.3.2 Thresholding:

Thresholding is the approach of image segmentation. From a Grayscale image, thresholding may be used to create binary images. A colour image is a combination of three colours Red, Green, and Blue (RGB). The conversion of a colour image into a grayscale image means converting the RGB values (24 bit) into grayscale values from a range of 0 (black) to 255 (white). Binary images are produced from colour images by segmentation.

In image processing, thresholding is used to split an image into smaller segments, using at least one colour or Grayscale value to define their boundary. The benefit of acquiring a binary image is that it reduces the complexity of the statistics and simplifies the procedure of recognition and classification.

4.3.3 Noise Removal:

Image noise is described as a random variant of brightness or coloration statistics in images and is typically an issue of digital noise. It may be produced via way of means of the sensor and circuitry of a scanner or virtual camera. There are different types of noise like,

- Gaussian Noise
- Gamma Noise
- Salt and Pepper Noise

Some unwanted noises are present in the original images. These noises are required to be removed from the images for efficient feature extraction. A bilateral filter is a non-linear, edge-preserving, and noise-reducing smoothing filter for images Therefore Bilateral filter is used to remove these noises because it preserves the edges of the elements in the image. This weight will be based on a Gaussian distribution.

4.3.4 Grayscale and Binarization:

For extracting the handwriting features, conversion to grayscale and binarization are important parts. The image instances are converted to grayscale and binarized using inverted global thresholding.

A binary picture is an image that has only two possible values for every pixel that are black and white. The binarization is performed on the basis of a threshold value given by the user. Binary images are obtained in digital image processing as masks or as the result of certain operations such as segmentation, thresholding, and dithering. In our module, the image is primarily converted into grayscale and then converted into binary.

4.3.5 Contour and Warp Affine Transformation:

After noise removal and the image conversion to grayscale and inversely binarized, the lines of the handwriting are straightened using dilation, contour and warp affine transformation of OpenCV library. This will give better results with further operations using horizontal projection of the image to extract these handwriting lines.

Dilation is a mandatory operation before finding contours in the image

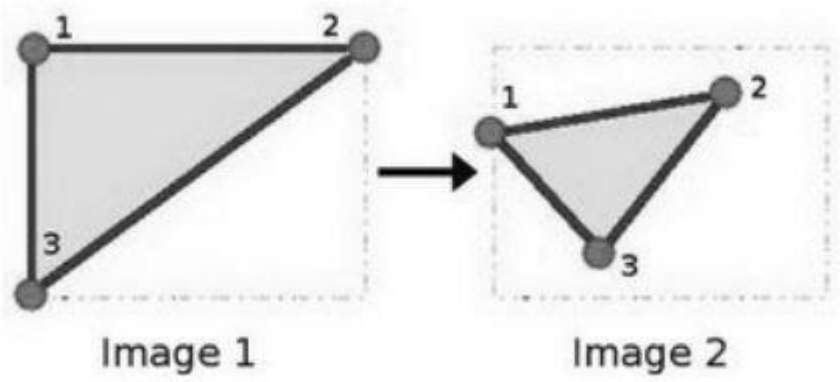


Figure 6 : Example of Affine Transformation

4.3.6 Horizontal and Vertical Projections:

The horizontal projection of an image is a python list of sum of all the pixel values of each row of the image, while vertical projection is a python list of sum of all the pixel values of each column of the image. Both of these operations are performed on grayscale images.

4.4 Handwriting Features

4.4.1 Baseline:

The baseline of handwriting forms an invisible line between the center and upper zones above and therefore the lower zone below. The direction of the baseline could also be hooked into a short lived condition, like psychological state or fitness . The three commonest baselines found in writing are - slanting downwards, rising upwards and balanced.

Table 1 shows each type reveals a particular personality trait of the writer.

Baseline		Personality Traits
	Straight Lines	Emotions are being disciplined by the writer. Self-control, Self Command, Calmness, Emotional Stability, Dependability.
	Straighten up	Happy, Carefree, Ambition, Confidence, Restlessness.The author needs to avoid routine demands.
	Straighten Down	Unhappiness, Tiredness, Depressions, Disappointment. Also, inflexibility and will power.

Table 1: Baseline Traits

4.4.2 Letter Size:

The size of the letters in a document represents the writer's desire to be noticed. In handwriting, large and bold writing indicates that the writers want to say “notice me”. While in the case of small writing writers are less desired to be noticed. Middle writing is the normal writing indicating the writer to be fit in the running world.

Table 2 shows the variations of the size of the letter and the personality associated with it.

Letter Size	Personality Traits

	Normal Letter Size	The writers are practical in nature and realistic.
	Large letter Size	The authors are capable of behaving with braveness, excitement, and confidence, but they are also capable of loudmouth, and lack of focus and discipline.
	Smaller letter Size	The authors are humble,Often to the point of feelings of lowness.

Table 2 : Letter Size Traits

4.4.3 Line Spacing :

Line spacing is the space kept by the writer between words. Table 3 shows the traits of the writers' different types of spacing.

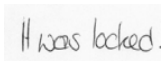
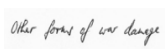
Line Spacing		**Personality Traits**
	Normal line spacing	The author has his own individual harmony and multiskill.
	Small line spacing	The writer's ideas and emotions are puzzled.
	Large line spacing	The writer isolates himself, socially, mentally, from his world.

Table 3 : Line Spacing Traits

4.4.4 Word Spacing:

It indicates the disposition towards criticism, and towards argument. The space left between the written words represents the distance that the writer would like to maintain between himself and society at large. Table 4 elaborates more about it.

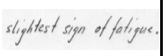
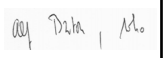
Word Spacing		Personality Traits
	Normal word spacing	Provides proof of the social maturity of the writer.
	Narrow word spacing	For attention, the author will crowd others, craving continuous presence and closeness.
	Wide word spacing	It suggests that authors, mostly because of an inner need for privacy, need to maintain a distance from human form

Table 4 : Word Spacing Traits

4.4.5 Top Margin:

Top margin is the space between the top of the paper and the first word written by the writer. Table 5 shows the traits of the person with different top margins.

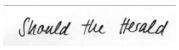
Top Margin		Personality Traits
	Small top margin	The writer shows lack of respect.
	Broad top margin	Respect for humility and formality are shown to the person being written to.

Table 5 : Top Margin Traits

4.4.6 Pen Pressure:

Pressure is the amount of pressure applied by the author to the writing surface. Table 6 explains more about the different characters of different pen pressure.

Pen Pressure	Personality Traits

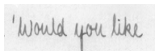
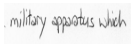
	Heavy pen Pressure	These types of writers make an impression.
	Light pen Pressure	These individual writers have a certain sense of delicacy.
	Medium pen Pressure	The writer indicates healthy life and willpower.

Table 6 : Pen Pressure Traits

4.4.7 Slant of Letters:

The slant of a writing refers to the direction of letter slope and is determined by the angle formed between the downstroke and the baseline.There are basically three types of slant: vertical, inclined and declined.

- Vertical - this sort features a head-over-heart emotional attitude.
- Inclined - The writer is normally sensitive and emotionally healthy.
- Declined - These people cry and laugh readily, give vent to their feelings.

Table 7 shows the various traits of different styles of writing letters.

Slant of Letters		**Personality Traits**
	Perpendicular	This kind of writer has an emotional attitude that is head-over-heart.
	A little tilted	This kind of author is typically sensitive and emotionally stable, but the responses are fair..
	Very tilted	These types of writers readily weep and laugh, give opening to their emotions, are future-oriented and goal-oriented and have an emotional nature that is eager, love, care and sensitive.

	Extremely tilted	These kinds of writers are very abnormal, and show a strong indication of wild behaviour and emotional illness.
	Lean Back	These authors refuse to consider sudden improvement or change in their routine.

Table 7 : Slant of Letters traits

4.5 System Requirements

1. Python 3.1: Python is an interpreted high-level programming language for general-purpose programming. Python is chosen over Matlab because it is easier to install and readily available. Python has extensive free libraries for image processing and machine learning, which makes it an ideal language for the project.
2. OpenCV library: This open source library for Python 3.1 provides various implementations of image processing algorithms used extensively in the project.
3. Sci-kit Learn library: This free library for Python 3.1 provides implementations of support vector machine algorithms and various other machine learning algorithms**.**

Chapter 5

Design of System

5.1 System Design

5.1.1 Support Vector Machine (SVM)

Our SVM model uses a scikit-learn library for implementation and eight classifiers are trained using the Radial Basis Function (RBF) kernel. With the help of this library we have extracted seven features from the handwritten text.

Contours are used to find the baseline of the written text. With the help of horizontal projection the individual lines are extracted. After the extraction of individual lines height of every line is calculated and the average is considered as the letter size. Total number of rows with horizontal projection and number of rows having value less than threshold are added and divided with the total number of lines to find out the line spacing. Vertical projection is used to find out word spacing. To calculate top margin the distance above the first line of text is calculated and it is divided with the letter size. Pen Pressure is calculated by an inverted binary threshold where the average of all non-zero pixels is the final value. Shear transformation is applied on the image and every angle is calculated, the highest value of the angle is considered as the slant of the text.

To predict the character of the writer various combinations from above extracted features are used.

5.1.2 Convolutional Neural Network (CNN)

The deep learning technique called Convolutional Neural Networks (CNNs) for this experiment has a model that combines both feature extraction with classification. In this model, the emotional state of a writer is recognized from his/her handwritten text sample and psychological assessment.
First of all we have mounted google drive with google colab for the path. We declare model json path, model path and file directory. One of the most important advantages of JSON is that it can be used to interchange data between computers and processes. JSON, like XML, afterward we declare a path for images and label index as LM, RM, SA, SD. Afterwards the required packages like keras, openCv, pickle, etc are installed. Then epochs are set and batch_size for_number of training examples are utilized in one iteration. The path for training images is set and preprocessing is performed on the images where resizing the image, thresholding, binarization, etc is done.
We created 4 label indexes for image data training, left margin, right margin, ascending, descending. The images are dilated to get a desired rectangular kernel size of 2*2 array and append the image.
The keras model is saved to a file and can be loaded up again to make predictions. The weights are saved directly from the model and later loaded using the function.

Emotion recognition using psychological assessment which shows the anxiety, stress and depression values.

5.2 Data Flow Diagrams

5.2.1 SVM data flow diagram

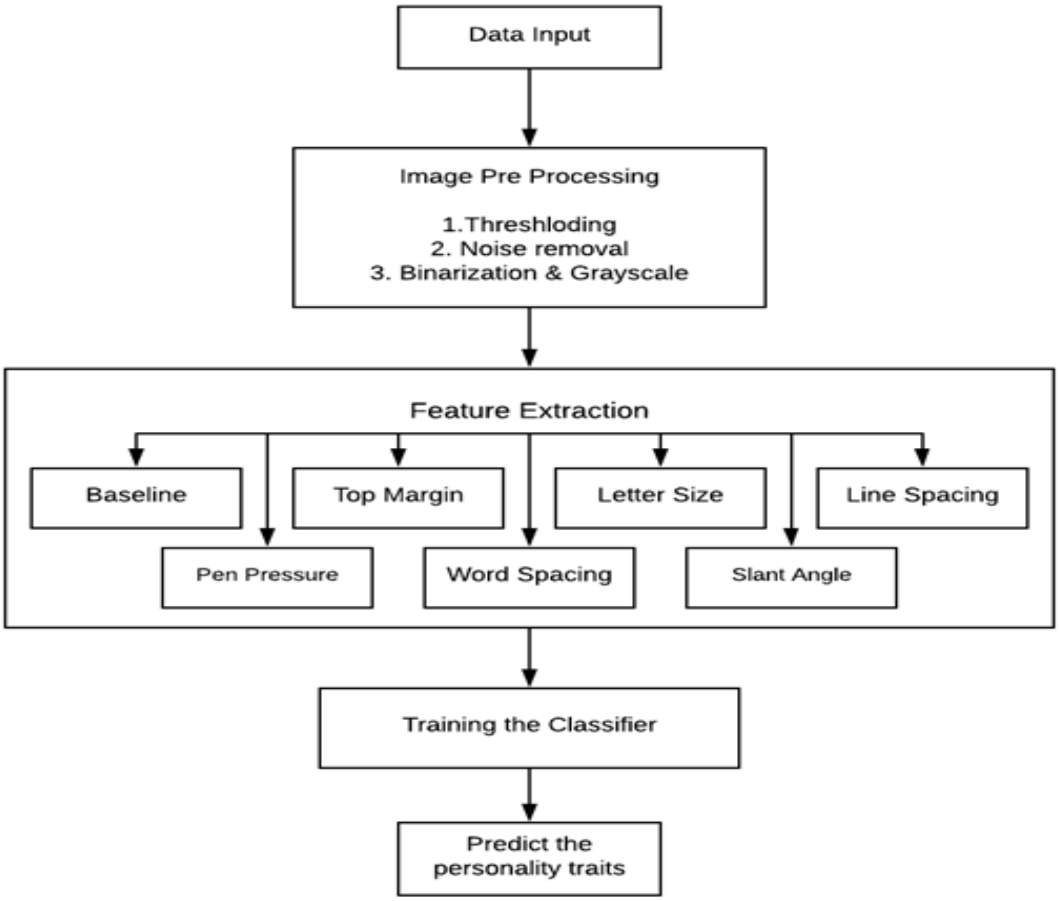


Figure 7 : SVM data flow diagram

5.2.2 CNN data flow diagram

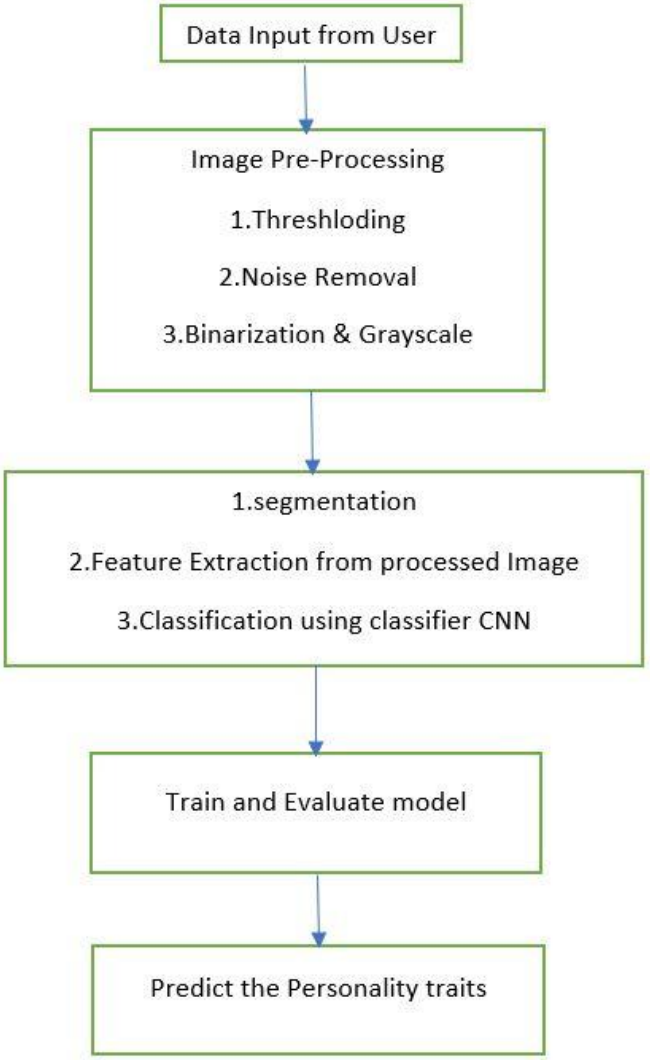


Figure 8 : CNN data flow diagram

Chapter 6

Implementation

6.1 SVM Model

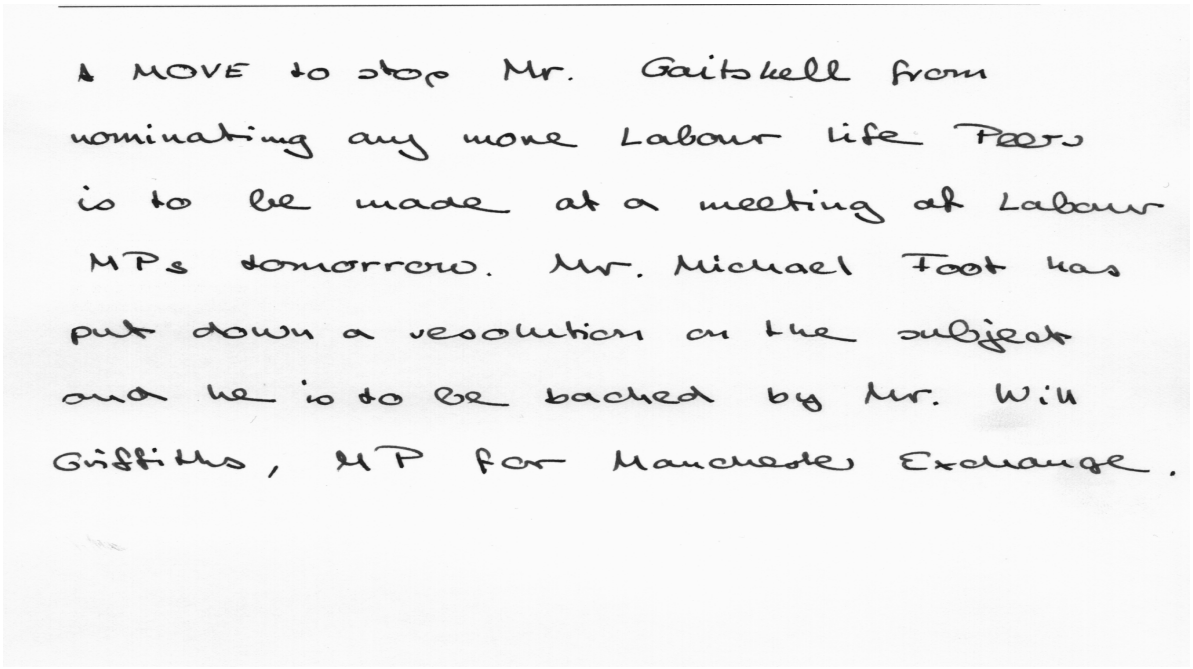


Figure 9 : Sample Handwriting 01 for SVM

```
Enter file name to predict or z to exit: m2.png
Baseline Angle: STRAIGHT
Top Margin: MEDIUM OR BIGGER
Letter Size: BIG
Line Spacing: SMALL
Word Spacing: SMALL
Pen Pressure: MEDIUM
Slant: EXTREMELY INCLINED
--------------------------------------------------
Your character assessments as per our system:
You are less stable Emotionally,
also your Mental Energy/Will Power is High,
generally you are a humble person,
You are short tempered person and take time to adapt changes,
Discipline is not a strong point for you,
and you are able Concentrate ,
Communication of information/thoughts to next person is clear,
You feel comfortable around more people.
--------------------------------------------------
```

Figure 10 : Sample Handwriting 01 Result

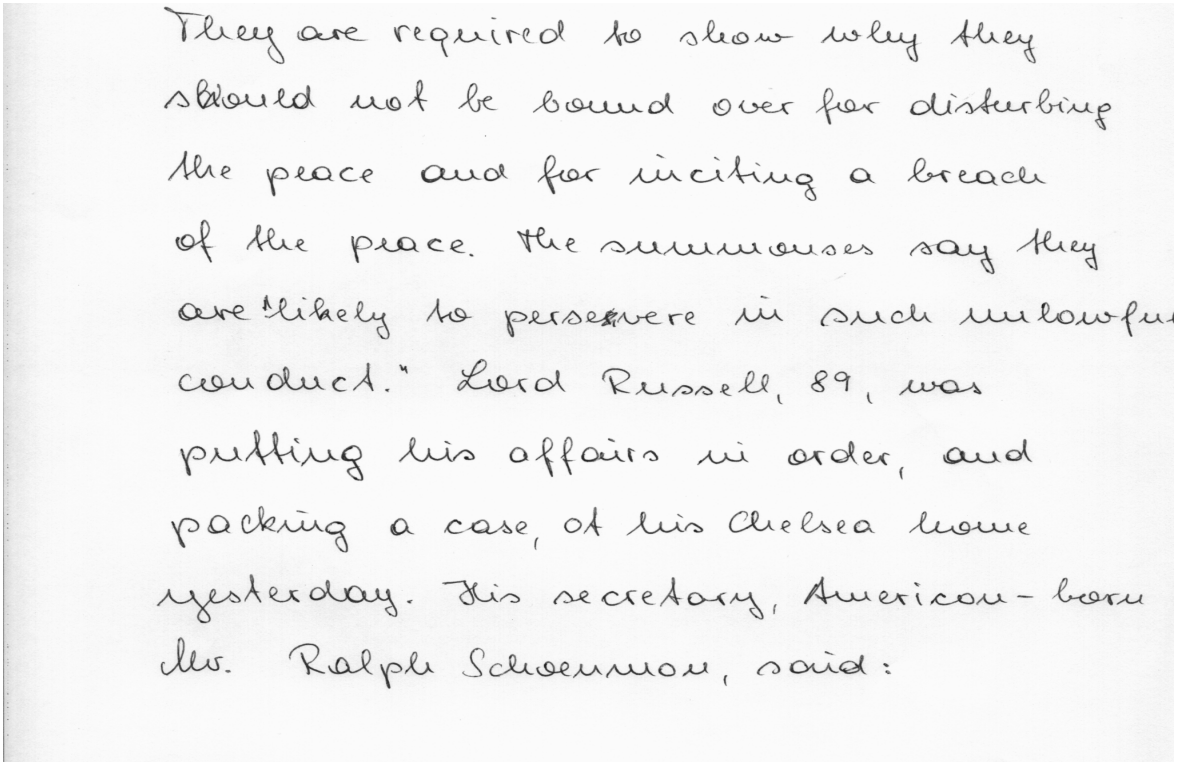


Figure 11 : Sample Handwriting 02 for SVM

```
Enter file name to predict or z to exit: m7.png
Baseline Angle: DESCENDING
Top Margin: NARROW
Letter Size: BIG
Line Spacing: MEDIUM
Word Spacing: BIG
Pen Pressure: MEDIUM
Slant: A LITTLE OR MODERATELY RECLINED
--------------------------------------------------
Your character assessments as per our system:
You are less stable Emotionally,
also your Mental Energy/Will Power is High,
Sometimes you like to boast about yourself,
You are short tempered person and take time to adapt changes,
Discipline is not a strong point for you,
also you have Poor Concentration,
Communication of information/thoughts to next person is clear,
You feel less comfortable around more people.
--------------------------------------------------
```

Figure 12 : Sample Handwriting 02 Result

6.2 CNN Model

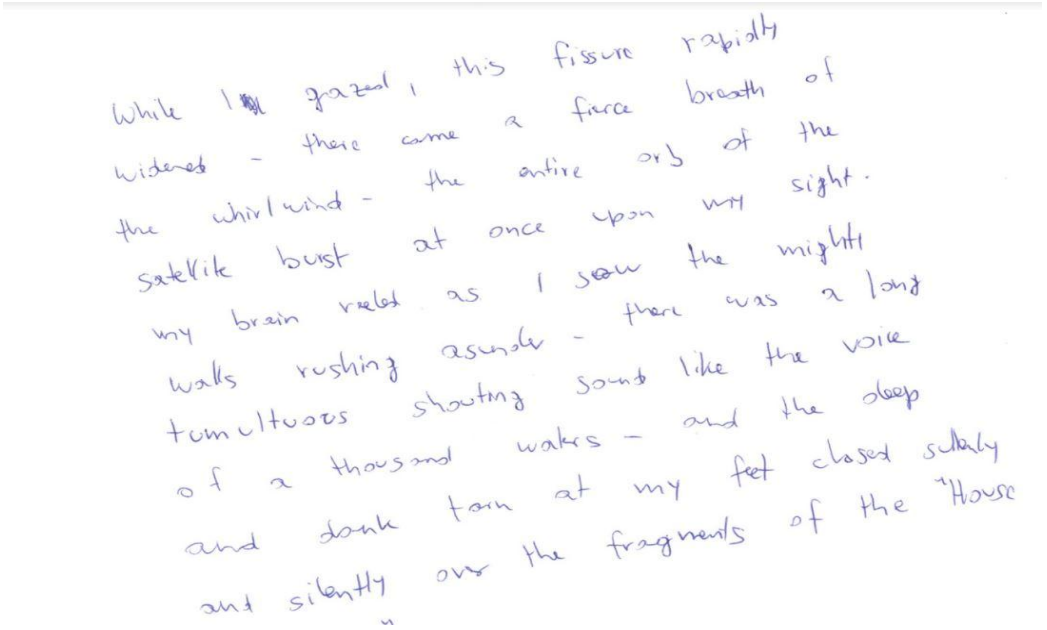


Figure 13 : Sample Handwriting 01 for CNN

```
*****Loaded Model from disk******
WARNING:tensorflow:9 out of the last 9 calls to <function Model.make_predict_function.<

 > Mental Energy :
```

Figure 14 : Sample Handwriting 01 Result

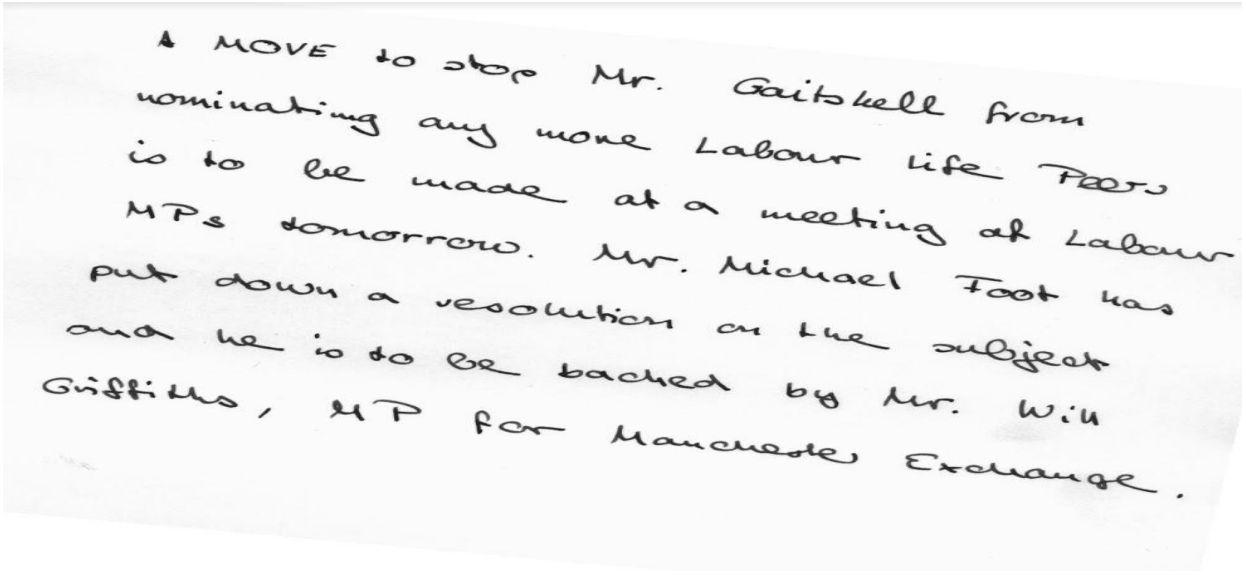


Figure 15 : Sample Handwriting 02 for CNN

```
*****Loaded Model from disk******
WARNING:tensorflow:10 out of the last 10 calls to <function Model.make_predict_function.<locals>.

 > Non communicative :
```

Figure 16 : Sample Handwriting 02 Result

Chapter 7

Experimental Results

Our SVM model uses a scikit-learn library for implementation and eight classifiers are trained using the Radial Basis Function (RBF) kernel. With the help of this library we have extracted seven features from the handwritten text. The eight SVM classifiers are trained with randomly chosen two thirds of all the images. The remaining images are used to test the accuracy score.

Samsuryadi & Kurniawan, Rudi & Mohamad, Fatma proposed a model based on neural networks that achieved an accuracy of 76%. Lala Septem Riza, Aldi Zainafif, Rasim Rasim, Shah Nazir proposed a model based on graphology techniques and implementing Fuzzy Rule-Based Classification Systems using Chi's Algorithm and achieved the accuracy of 76%.

Sr. No.	Researchers	Methodology	Accuracy
1	Behnam Fallah, Hassan Khotanlou	Neural Network	76%
2	L. S. Riza, A. Zainafif, Rasim, and S. Nazir	Fuzzy Rule Base Classification	76%
3	Proposed System	SVM	80%
4	Proposed System	CNN	84.84%

Table : Accuracy Table

The deep learning technique called Convolutional Neural Networks (CNNs) for this experiment has a model that combines both feature extraction with classification. In this model, the emotional state of a writer is recognized from his/her handwritten text sample and psychological assessment.

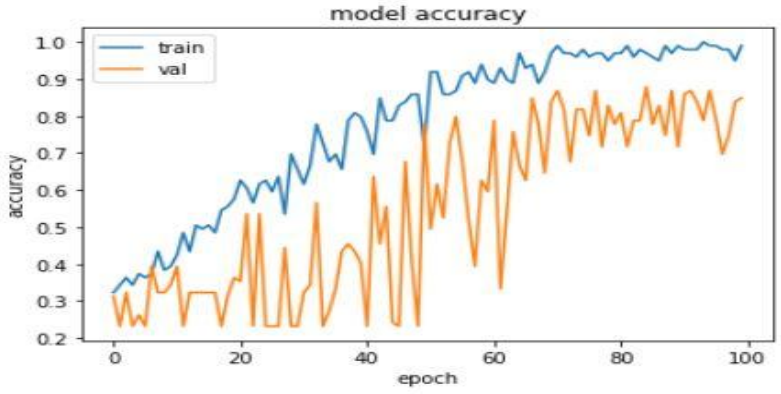


Figure 17 : CNN Accuracy Graph

Chapter 8

Conclusion

In this project, we proposed an approach that uses Machine Learning to recognize the personality traits of people with the help of SVM and CNN. A study was done to automate the process of identifying personality traits through handwriting analysis. The SVM model is trained to extract features like base line, letter size, line spacing, word spacing, top margin, pen pressure and slant of letter, using these features our model predicted the personality of the author. The svm model achieved the accuracy of 80%. CNN model is trained to extract features like left margin, right margin, ascending angle and descending angle, using these features our model achieved the accuracy of 84.84%. Using our research the model can be used to reduce the amount of time required to find the personality of a writer and also reduce human error.

References

[1]Devesh Sony, Rakshita Sawant,”Identifying Human Behavior Characteristics using Handwriting Analysis”, IRJET(2019)

[2]U. Munir and M. Ozturk,”Automatic Character Extraction from Handwritten Scanned Documents to Build Large Scale Database”, IEEE(2019)

[3]U.Karthikeyan, Dr. M. Vanitha,”A Study on Text Recognition using Image Processing with Data Mining Techniques”, IJCSE(2019)

[4]Poorna Banerjee Dasgupta, ”Human Behavioral Analysis Based on Handwriting Recognition and Text Processing”, IJCTT(2018)

[5]Diksha Khurana, Aditya Koli, Kiran Khatter, Sukhdev Singh, ”Natural Language Processing: State of The Art, Current Trends and Challenges”, Research Gate(2017)

[6]Zhaorong Zong, Changchun Hong, ”On Application of Natural Language Processing in Machine Translation”, IEEE(2018)

[7]Noman Islam, Zeeshan Islam, Nazia Noor, ”A Survey on Optical Character Recognition System”, ISSN(2016)

[8]Karez Abdulwahhab Hamad, Mehmet Kaya,”A Detailed Analysis of Optical Character Recognition Technology”, IJAMEC(2016)

[9]Rashi Agarwal, Shiwali Alone, Aaditi Dhadwal, Dr. D. S. Bormane, Prof. S.V. Kedar, ”Automatic Emotion Recognition through Handwriting Analysis: A Review”, IEEE(2015)

[10]Somayeh Hashemi, Behrouz Vaseghi, Fatemeh Torgheh, ”Graphology for Farsi Handwriting Using Image Processing Techniques”, ISSN(2015)

[11]Mr. Shashank H. Yadav, Mr. Pratik Madhukar Manwatkar, ”Text Recognition from Images”, IEEE(2015)

[12]Parmeet Kaur Grewal, Deepak Prashar, ”Behavior Prediction Through Handwriting Analysis”, IJCST(2012)

[13]Subodh L. Wasankar, Harshad Mahajan, Deovrat Deshmukh, Hemant Munot, ”Machine Learning with Text Recognition”, IEEE(2010)

[14]Ayatullah Faruk Mollah, Nabamita Majumder, Subhadip Basu, Mita Nasipuri, ”Design of an Optical Character Recognition System for Camera based Handheld Devices”, IJCSI(2011)

[15]Xiaojun Li, Weiqiang Wang, Shuqiang Jiang, Qingming Huang, Wen Gao, ”FAST AND EFFECTIVE TEXT DETECTION”, IEEE(2008)

[16]Sofianita Mutalib, Roslina Ramli, Shuzlina Abdul Rahman, Marina Yusoff, Azlinah Mohamed, ”Towards Emotional Control Recognition through Handwriting useing Fuzzy Inference”, IEEE(2008)

[17]Vivek Bansal, Sunil Goyal, and Kalpana Srivastava, ”Study of prevalence of depression in adolescent students of a public school”,Ind Psychiatry J.(2009)

[18]Shitala Prasad, Vivek Kumar Singh, Akshay Sapre,”Handwriting Analysis based on Segmentation Method for Prediction of Human Personality using Support Vector Machine“, IJCA(2010)

[19]Lala Septem Riza, Aldi Zainafif, Rasim Rasim, Shah Nazir,”Fuzzy Rule-based Classification Systems for the Gender Prediction from Handwriting”,ISSN(2018)

[20]Samsuryadi, Samsuryadi & Kurniawan, Rudi & Mohamad, Fatma. (2021). Automated handwriting analysis based on pattern recognition: A survey. Indonesian Journal of Electrical Engineering and Computer Science. 22. 10.11591/ijeecs.v22.i1.pp196-206.