VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

CSE3001

SOFTWARE ENGINEERING PROJECT

FALL SEMESTER 2020

J COMPONENT PROJECT REPORT

INTELLIGENT GRAPHOLOGIST

Presented By

SAURABH KUMAR JHA (18BCE0197)

NITIN RANJAN (18BCE0272)

TARUSHI JAIN (18BCE0859)

UTKARSH SINGH (18BCE2139)

Under the guidance of

GAYATHRI A

Assistant Professor Sr. Grade 1

Department of Software Systems

School of Computer Science and Engineering

ACKNOWLEDGEMENT

We want to acknowledge our respected professor Gayathri A who helped us in all research and assisted us in all the way she could have done. She has been the backbone by supporting us in all the possible ways and made us this far to achieve our result. We thank our prestigious university for giving us this wonderful opportunity for proving us whom we are and not limiting theories to just books but also in the involvement of various mini projects. We consider this opportunity as a chance to deliver our token of thanks to VIT for providing us immense facilities and equipment required to make our project a great success. We take this time to express all our gratitude towards all the staff members who helped us during the course of the project.

Contents

1. ABSRACT

Data such as the dynamically captured direction, stroke, distance, size, pressure and shape of an individual's signature enable handwriting to be a reliable indicator of an individual's identity. Forensic handwriting examination has a new frontier: the digital signature in biometric modality that uses, for recognition purposes, the anatomic and behavioral characteristics that an individual exhibit when signing her/his name. Handwriting examiners often have to determine if the signature is genuine or simulated, dynamic information such as velocity and pressure are fundamental and can be estimated qualitatively.

Handwriting Analysis is the study of handwriting, especially when employed as a means of analyzing character. Deep learning has been widely used to recognize handwriting. In offline handwriting recognition, text is analyzed after being written. Information, such as pen stroke, pressure and speed of writing are analyzed. It is particularly necessary for historical documents, archives, or mass digitization of hand-filled forms. It is generally considered a pseudoscience. It can be achieved using neural networks. Neural networks are able to learn features from analysing a dataset, and then classify an unseen image based on weights. A neural network is a series of algorithms that endeavours to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Neural networks can adapt to changing input; so, the network generates the best possible result without needing to redesign the output criteria.

2. INTRODUCTION

Handwriting also termed as brain-writing is a useful measure in identifying the characteristic personality traits of an individual. Handwriting analysts also known as graphologists can examine an individual‟s handwriting to predict the personality raits of the writer. Automated handwriting analysis can be used to examine personal traits of candidates during interviews accurately as the accuracy of an analyst highly depends on his skill set. Also hiring a graphologist to analyze hundreds of samples for recruitment purpose will be time consuming and not be feasible economically. This work discusses about a method for analysing real world handwritten text samples with the aid of technology. The analysis is done for specific features of the sample for determining various characteristic behavioural traits of the person.

The first release of this project aims to distinguish between different people's personality through handwriting analysis. This software will be inputting an image of the handwriting of a person and then will tell what their personality is like. To achieve this, simple Neural Network and Back Propagation Algorithm will be used.

3. LITERATURE SURVEY

[1] The authors define graphology and further argues the business needs and impacts of graphology, the previous attempts made at automating the same using various segregation techniques and a brief comparison between the same.

[2] The authors present a brief about graphology and use a sentiment-polarity based model. The authors present their framework where they create a four-phase model with four engines that carry out lexical analysis on words and then give sentiment scores on a normalised polarity scale (i.e. between -1 and 1) and store it in a student specific memory location.

[3] This research article was mostly used for understanding graphology itself. The author tries to explain why it is important to understand graphology, the class features of handwriting analysis, the features that should be looked for and the unique features that professionals try to look for. The author further explains how samples ought to be collected, why handwriting analysis is not treated to be a science and why forgery is always a possibility.

[4] The authors argue that graphology might serve as an excellent mean of introspection and self-growth. They collect 25 random samples and use standard graphological functions to segregate the data and then analyse them. They do mention that it is possible using fuzzy logic functions, however they use image processing. Finally, they compare the data with the one provided by a professional and find a high degree of similarity, thus establishing that it is possible to analyse handwriting and personalities digitally.

[5] One of the founding papers for modern study in graphology, James McNeal provides the idea of how handwriting might serve as a tool to evaluate a person's personality, mood and aid marketeers in establishing better inter-personal relationships. It does not provide any mathematical or experimental basis for it, however it does provide a theoretical basis by arguing that graphologists and psychiatrists stand in agreement with a correlation coefficient of 0.6 to 0.7. However, he does argue that it is too early to use graphology as a tool in the market and it might take a few years before its usability or short-comings are established.

[6] Using seven professional graphologists and seven test subjects, each subject was asked to select 20 traits about himself/herself and the graphologists were asked to determine the traits of the control-subjects using a page of text each subject had written. These datasets were compared. Further, all subjects were subjected to a standard personality test. And this data was then compared to the previous data. This was used to evaluate a coefficient of correlation between the graphologist's study and the actual personality of the control-subjects. The results were not very promising. However, this was on of the earliest experiments in graphology and so the author does not conclusively present any statement for or against it.

[7] This research paper has outlined a behavioral system or a tool to achieve high analysis rate on the nature of the person with the help of artificial neural networks and character recognition and detection systems. The proposed system will automatically detect the traits of the person by letter slants.

[8] In this paper, we propose a study on the complexity measure of an object. It is based on the analysis of different details that may be limited by object contours. They may be holes or convexity evolution along the contour line. We focus in the same way on empty zones and filled zones. This study leads to a novel measure of the topology complexity – vacuity measure – that quantifies the relation between emptiness or space and objects. Based on the vacuity measure, we propose to define a novel shape descriptor and the associated dissimilarity measure. They can be applied in handwritten character analysis and in object recognition in general. The experiments are performed on a handwritten character dataset (ORIFLAMMS) and the object shape dataset.However, human recognition process considers not only the object itself but also considers the elements not belonging to the object, for example, the holes and contour details of the object. Such elements are defined the vacuity part of the objet. Human perception will capture both the vacuity and the object into the process. Therefore, we are working on a vacuity measure and its application on pattern recognition. This measure quantifies the relation of object and its empty parts disposition.

[9] Personality is a fundamental basis of human behaviour. Personality affects the interaction and preferences of an individual. People are required to take a personality test to find out their personality. Social media is a place where users express themselves to the world. Posts made by users of social media can be analysed to obtain their personal information. This experiment uses text classification to predict personality based on text written by Twitter users. The languages used are English and Indonesian. Classification methods implemented are Naive Bayes, K-Nearest Neighbours and Support Vector Machine. Testing results showed Naive Bayes slightly outperformed the other methods. The system will retrieve a collection of traits from users. Text from user then pre-processed into vector data. Classification process will classify user's text into a labelled dataset. The results are predictions for each Big Five traits, primary personality characteristics and secondary personality characteristics which obtained from the combination between two traits. System developed is a web application.

[10] This paper is to specify the handwritten data of 114 students categorised under three classes, the action, person and event. The authors has tried to perform the analysis on the imagination of the people as the subjects have to write the imagination content video displayed to them and highlight the coloured text of the handwritten data.

[11] This paper presents an experimental framework corresponding graphological features and main aim is to build a model to measure the qualities of the writer and provide resourceful manpower for human resource experts to save both time and efforts and make the selection process easy.

[12] This paper put forward the personal feature traits of those in particular between the age group of 20-35 years when they tackle many interviews. Polygonalization method is used to evaluate the baseline while margin will be calculated using the method of vertical scanning. Supervised machine learning algorithm, Feature Vector Matrix and similarity matrix method are key approaches used over data sets. The planned system can be used as a corresponding tool by the graphologist to recover the accuracy of graphological analysis and also makes the process rapid.

[13] The authors try to describe an online handwriting system that can support 102 different languages. They argue that this new system has 40% lesser errors than a conventional

segment-and-decode system. Also, their new algorithm produces 10 times faster results. While most of their data analysis techniques are still OCR based, what is improved is the neural network model used. It is of recurrent nature and bidirectional. These neural networks are fed with Bezier's curves over datasets and the results are finally decoded.

[14] The authors argue that through structural feature extraction followed by established studies on zone, loop, angle and other features, they were able to obtain about 87% accuracy using CNN algorithm. Further, using SVM did not yield any significant improvements. However, using noisy datasets led to a drastic reduction in result accuracy.

[15] The author argues that there already exist several established methodologies for correlating handwriting and the mood of a person. However, they have tried to formulate a new methodology. On a control group of 36 participants and a clinical sample collected from 44 patients, they used Cohen Kappa method to correlate these two datasets with the results produced by 4 professional graphologists. The test was carried out after establishing a global evaluation standardisation for graphical analysis. The coefficient of correlation is high (~0.62) between graphologists and between 0.47 to 0.60 between the psychiatric assessment and graphological assessment. However, Cohen's Kappa yielded better results than Lorusso et al.

[16] Discussing RCNN and how it optimises CNN problems and gives rise to LSTM.

[17] Handwriting Analysts study the handwriting and predict the behavior of a person based on their skills. To make this more accurate, a relatively simpler method has been proposed to anticipate the personality of a person by exploring various handwriting features. The system considers five discriminating features such as breaks, size, space between words, baseline, loop of 'e' and few other features like pressure, margin, slant and dot distance in 'i'. The proposed system can be used as a twin tool by graphologist to improve the accuracy and anticipate the behaviour of a person faster. The estimated weighted accuracy of 93.77 % is achieved.

[18] In this paper, a method has been proposed to predict the personality of a person from the features extracted from his handwriting using Artificial Neural Networks. The personality traits revealed by the baseline, the pen pressure and the letter„t" as found in an individual's handwriting are explored in this paper. Three parameters, the baseline, the pen pressure and the height of the t-bar on the stem of the letter_t' are the inputs to the ANN which outputs the personality trait of the writer. The evaluation of the baseline is using the polygonization method and the evaluation of the pen pressure utilizes the grey-level threshold value. The height of the t-bar on the stem of the letter _t' is calculated using template matching. The performance is measured by examining multiple samples.

[19] This research proposed a tool that can identify small letter 't' character from a set of handwriting and the ambition levels of a person based on the small letter 't' character using back propagation neural network. It will analyze three different angles of t-bars: up-turned, horizontal and down-turned. From there, we will determine the level of ambition of a person. This study has proven that the artificial neural network is able to classify the level of ambition of person based on handwriting. It learns the input patterns and able to classify them. The implementation of the personality analysis using the back

propagation neural network model is the alternate tool for the graphologist to do their personality analysis judgment.

[20] In this paper, an innovative system biometric of specific writers' identification based on technical expert calligraphic and graphology on handwritten script is presented. The system allows for the distinguishing with a success rate of 99.34% among the different writers from our database, using only 5 graphological parameters and integrating them in the automatic biometric system based on NN+MV A for short database or SYM when the number of writers for our database is increased. These graphological parameters are "longitude," "Union of letters," "pressure," "thinningarea," and "infovocalA." Therefore, the use of adequate parameters is a main reason for obtaining stability and efficiency on the implemented system. Finally, the independence of the system is also demonstrated regarding the used classifier, because it provides similar success for 29 writers, and a bit better SYM vs. NN, when the database is increased.The success percentage achieved with five of these characteristics from this database of 29 writers is 99.34%. In new experimentation, with these same parameters and enlarging the database to 70 users, a success rate of 92% was reached.

[21] This paper presents a hybrid KNN-SVM method for cursive character recognition. Specialized Support Vector Machines (SVMs) are introduced to significantly improve the performance of KNN in handwrite recognition. This hybrid approach is based on the observation that when using KNN in the task of handwritten characters recognition, the correct class is almost always one of the two nearest neighbors of the KNN. Specialized local SVMs are introduced to detect the correct class among these two different classification hypotheses. The hybrid KNN-SVM recognizer showed significant improvement in terms of recognition rate compared with MLP, KNN and a hybrid MLP-SVM approach for a task of character recognition.The experiments demonstrated that the combination of kNN with SVMs experts pairs of classes that constitute the greatest confusion of kNN, have improved performance in terms of recognition rate. The results showed a significant improvement from 1.00% to 3.61% in recognition rate for all cases tested (uppercase, lowercase and uppercase + lowercase).

4. PROBLEM STATEMENT

In offline handwriting recognition, text is analyzed after being written. Information, such as pen stroke, pressure and speed of writing are analyzed. Here the text will be analyzed simultaneously and will be interpreted. It can also be digitalized presenting the use of paper and the possibility of losing the record. It will also save time as all the work is done by the software rather than the person. All these benefits are associated with the software. Mainly the benefit is that we are getting an analyzer as accurate as a professional psychologist.

The project is intended for all the persons who are technology savvy. Apart from this it is intended for project managers, examination testers, psychologist, researchers and software engineering professors. The scope for the product includes psychologist, researchers, people who want to understand their personality from handwriting.

Basically, this project is intended for every person who can write and wants to analyse their handwriting. Ultimately, the objective of the project is to take the first steps in this field of human psychology with Machine Learning with the hope that eventually this analyzer becomes as accurate as a professional psychologist.

5. REQUIREMENTS

5.1. FUNCTIONAL REQUIREMENTS

F1<Image Uploading>: The system shall be able to upload images from the user's device and be able to use it for further image analysis. From the end user's point of view, thumbnails are shown in the place of the file name in the list of attachments.
F2<Image Processing>: The system shall be able to process the image and perform Feature Extraction to obtain useful information like dynamically captured direction, stroke, distance, size, pressure and shape of an individual's signature to assess his personality.
F3<Progress Bar>: The system shall be able to display the progress of the feature extraction to the user. From the end user's perspective, the progress shall be displayed in the form of percentage on a progress bar.
F4<Output Result>: The system shall be able to display the final output to the user in the form of a report and with an estimate of the system's confidence in the output.

5.2. NON-FUNCTIONAL REQUIREMENTS

PERFORMANCE REQUIREMENTS

P1 <Image Uploading>: The software should be able to upload images within 30 seconds
P2<Image Processing>: The software should be able to process the image within maximum of 5 minutes

SAFETY AND SECURITY REQUIREMENTS

Authentication: The software should have a login/signup page to ensure user authentication
Data Protection: To ensure data protection the website should use an https protocol. Also, the

site should have an option for users ‘right to be forgotten’ so that they can have their details removed from a website if they request it.

SOFTWARE QUALITY ATTRIBUTES

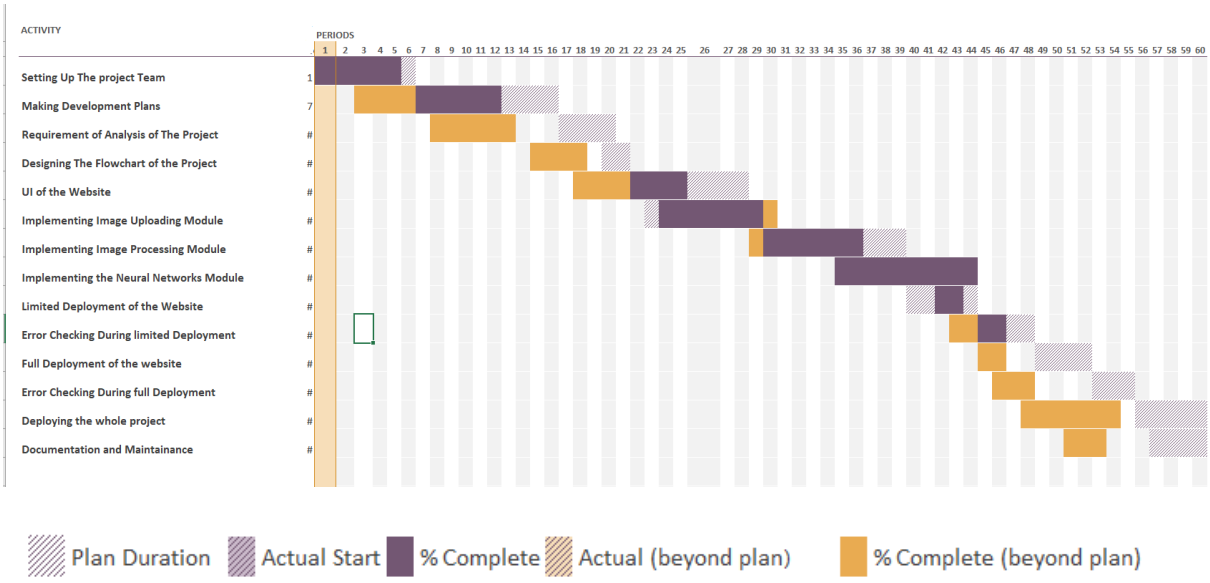
Usability: The software requires no advanced technical knowledge from the user’s perspective to run the product. With its easy to use UI the user is only required to upload his signature and with a click of a button is able to obtain his personality results without any hassles.

Modifiability: The software is built in terms of different modules working in tandem to achieve a cohesive whole. Since each module can be run independently it is quite easy to test different components.

6. PROJECT SCHEDULING CHARTS

6.1. Gantt Chart

Gantt Charts are bar charts that illustrate a project schedule. They may also show dependency relationship between activities and current schedule dates.

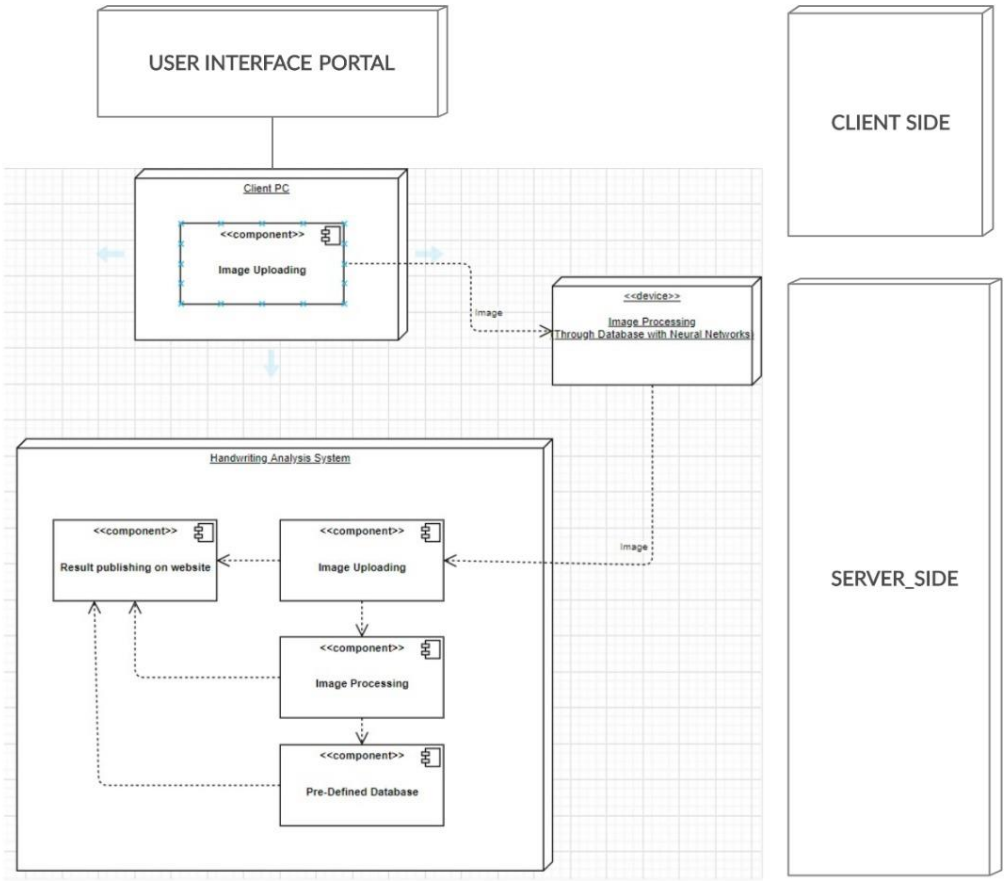


ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Setting Up The project Team	1	6	1	5	100%
Making Development Plans	7	10	3	10	100%
Requirement of Analysis of The Project	17	4	8	6	100%
Designing The Flowchart of the Project	20	2	15	4	100%
UI of the Website	22	7	18	8	100%
Implementing Image Uploading Module	23	7	24	7	100%
Implementing Image Processing Module	30	10	29	8	100%
Implementing the Neural Networks Module	35	10	35	10	100%
Limited Deployment of the Website	40	5	42	2	100%
Error Checking During limited Deployment	45	4	43	4	100%
Full Deployment of the website	49	4	45	2	100%
Error Checking During full Deployment	53	3	46	3	100%
Deploying the whole project	56	6	48	7	100%
Documentation and Maintainance	57	4	51	3	100%

7. DESIGN DIAGRAMS

7.1. ARCHITECTURE DIAGRAM

Architecture diagram is a diagram of a system that is used to abstract the overall outline of the software system and the relationships, constraints and boundaries between components.

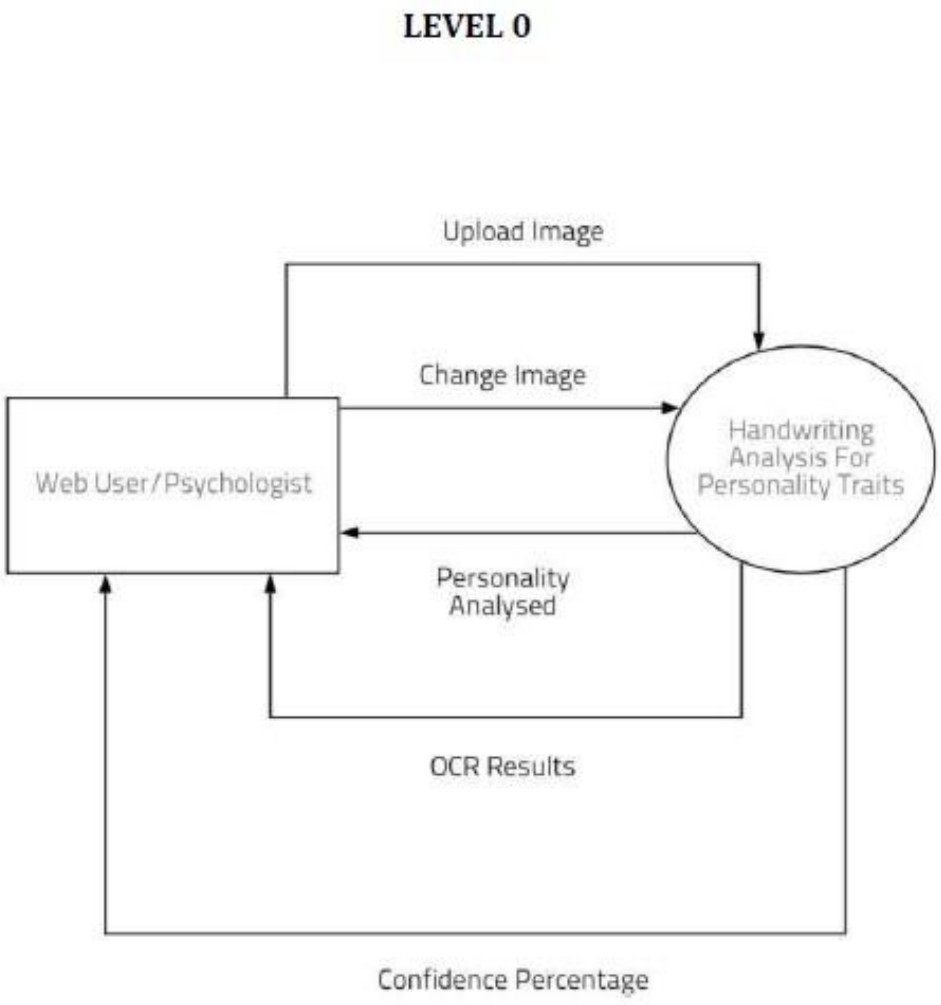


7.2. DATA FLOW DIAGRAM

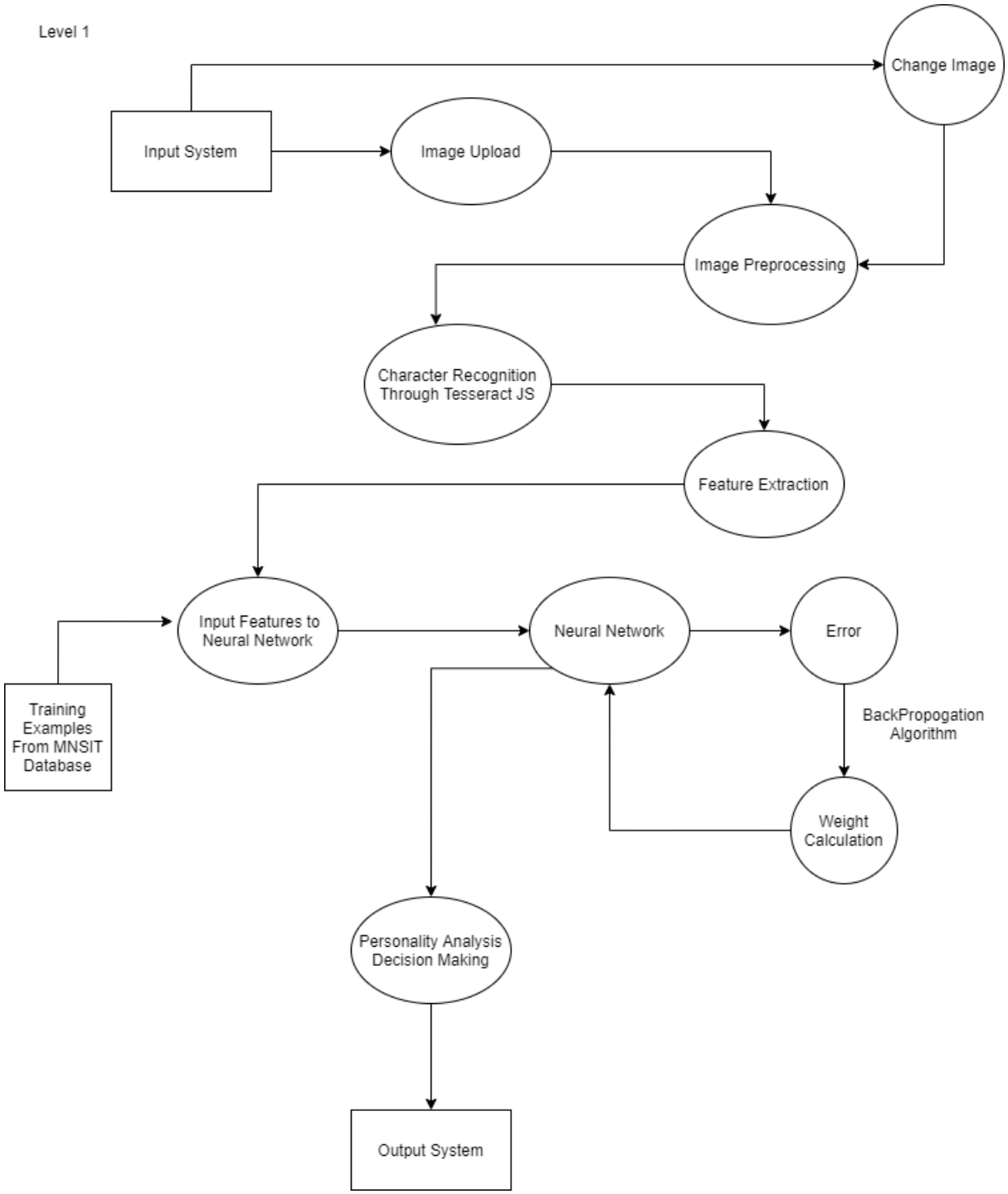
A data-flow diagram is a way of representing a flow of data through a process or a system. The DFD also provides information about the outputs and inputs of each entity and the process itself. A data-flow diagram has no control flow, there are no decision rules and no loops.

DFD levels are numbered 0, 1 or 2, and occasionally go to even Level 3 or beyond. The necessary level of detail depends on the scope of what you are trying to accomplish. DFD Level 0 is also called a Context Diagram. It's a basic overview of the whole system or process being analyzed or modeled.

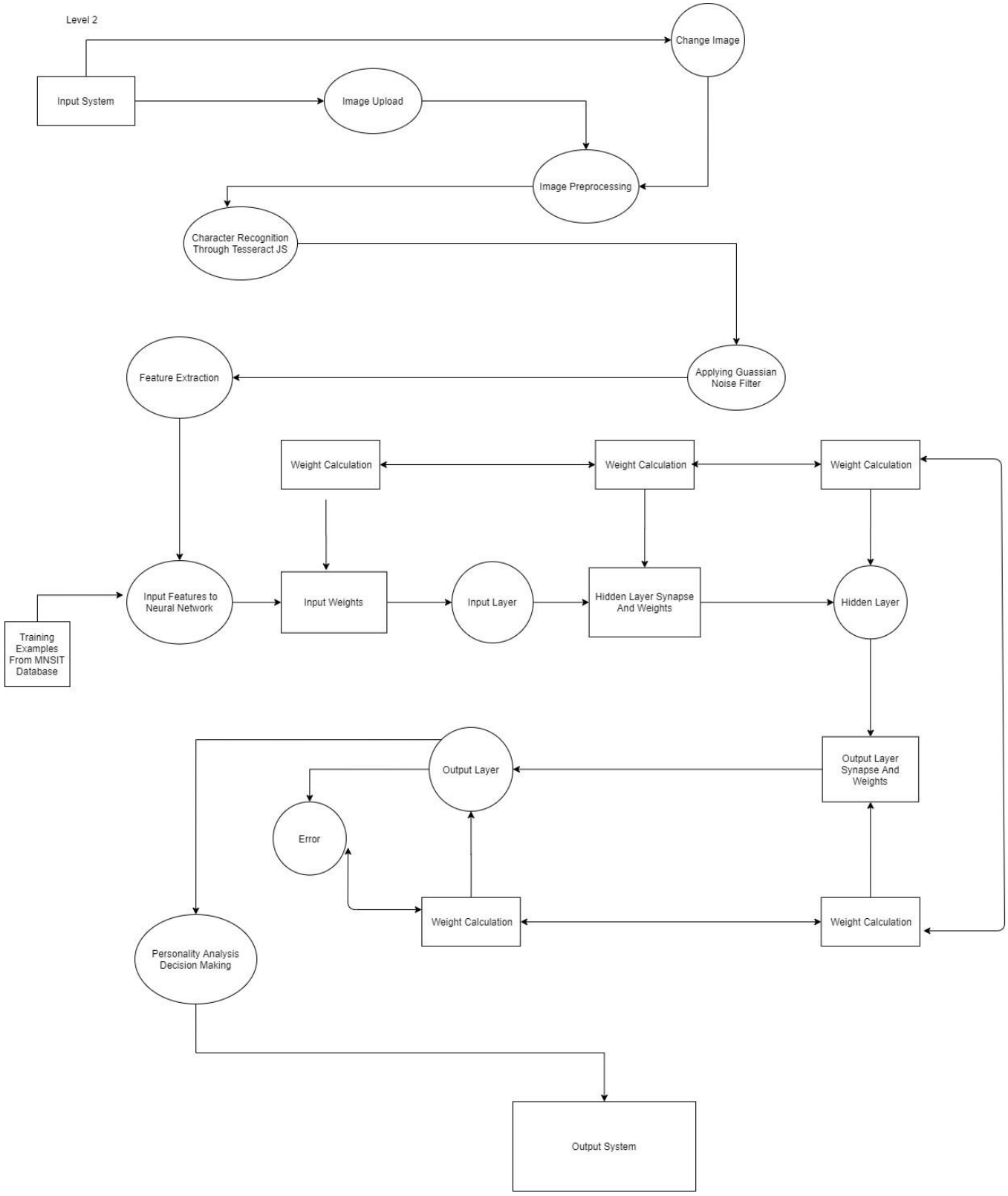
LEVEL 0



LEVEL 1


Level 1
Change Image
Input System
Image Upload
Image Preprocessing
Character Recognition Through Tesseract JS
Feature Extraction
Input Features to Neural Network
Neural Network
Error
Training Examples From MNSIT Database
BackPropogation Algorithm
Weight Calculation
Personality Analysis Decision Making
Output System

LEVEL 2

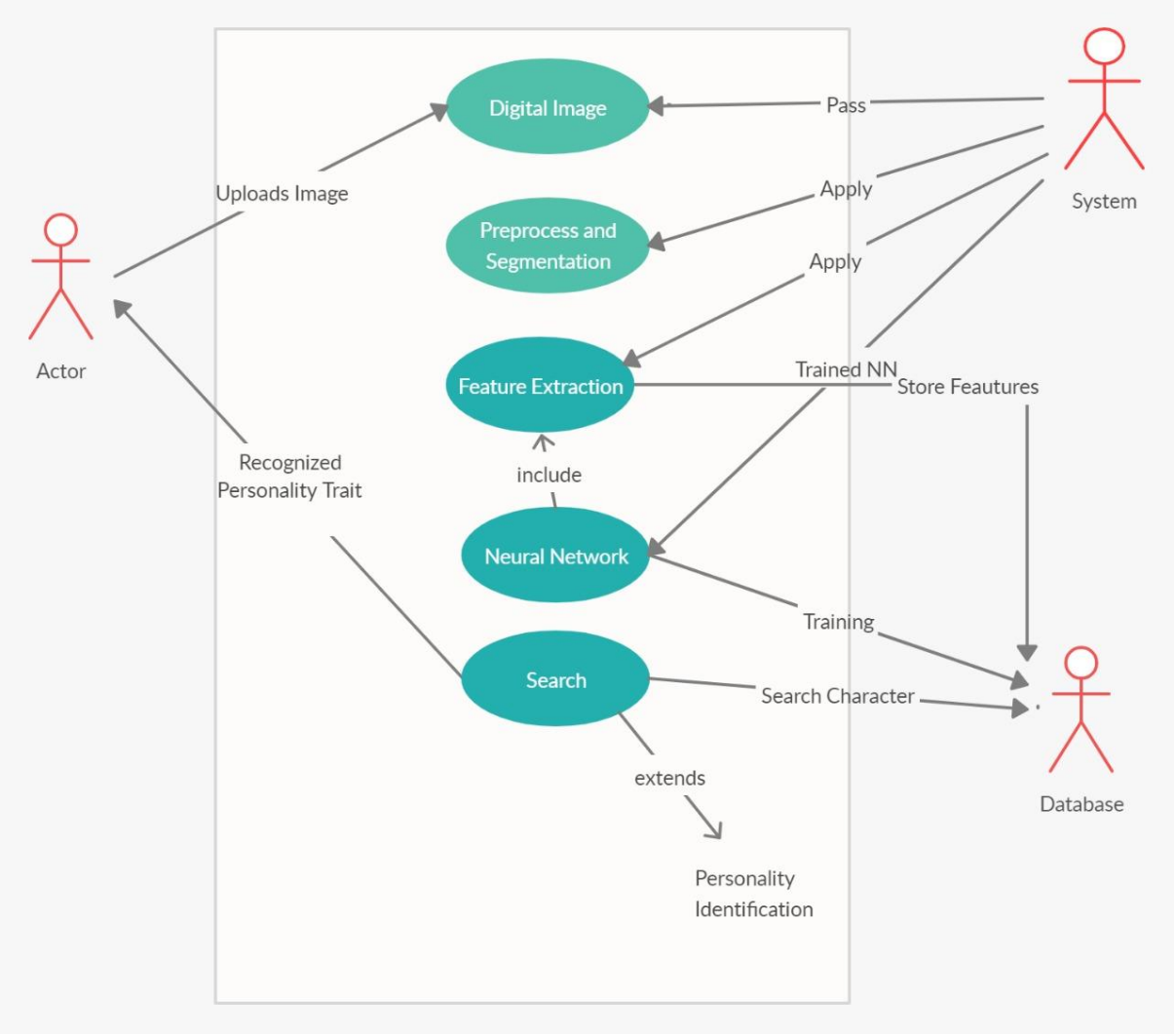

Level 2
Change Image
Input System
Image Upload
Image Preprocessing
Character Recognition Through Tesseract JS
Feature Extraction
Applying Guassian Noise Filter
Weight Calculation
Weight Calculation
Weight Calculation
Input Features to Neural Network
Input Weights
Input Layer
Hidden Layer Synapse And Weights
Hidden Layer
Training Examples From MNSIT Database
Output Layer
Output Layer Synapse And Weights
Error
Weight Calculation
Weight Calculation
Personality Analysis Decision Making
Output System

7.3. USE CASE DIAGRAM

A use case diagram at its simplest is a representation of a user's interaction with the system that shows the relationship between the user and the different use cases in which the user is involved. The purpose of use case diagram is to capture the dynamic aspect of a system.

A use case diagram is usually simple. It does not show the detail of the use cases. It only summarizes some of the relationships between use cases, actors, and systems. It does not show the order in which steps are performed to achieve the goals of each use case.

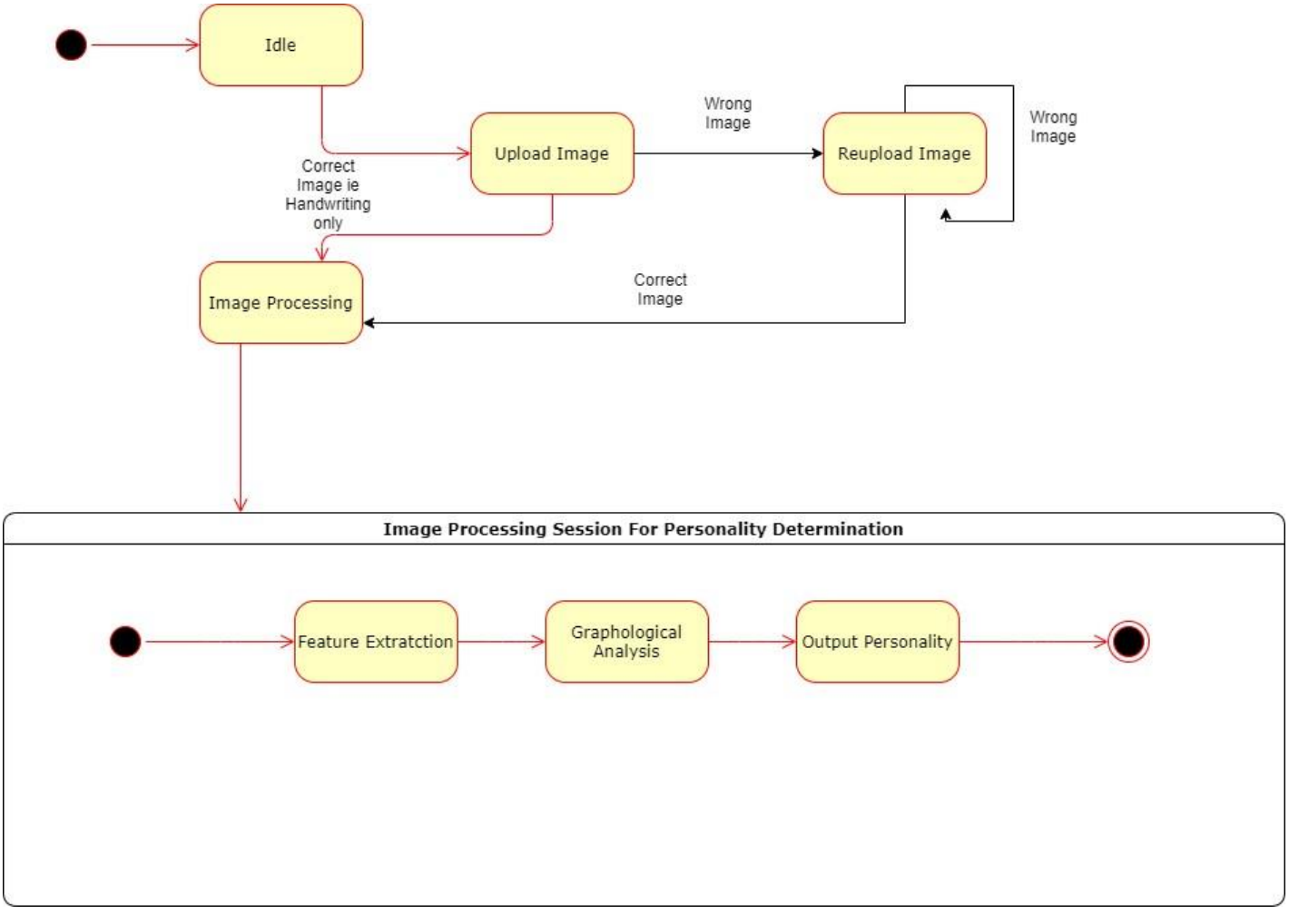
Below Use cases specify the expected behavior (what), and not the exact method of making it happen (how). It has helped us design a system from the end user's perspective and also for communicating system behavior in the user's terms by specifying all externally visible system behavior.



7.4. STATE CHART DIAGRAM

State chart diagram is used to model the dynamic nature of a system. They define different states of an object during its lifetime and these states are changed by events. They are useful to model the reactive systems. They typically are used to describe state-dependent behavior for an object.

An object responds differently to the same event depending on what state it is in. Hence this diagram is drawn to show the different states of an entity and also show how an entity responds to various events by changing from one state to another. Their main purpose is to model the dynamic nature of our project.

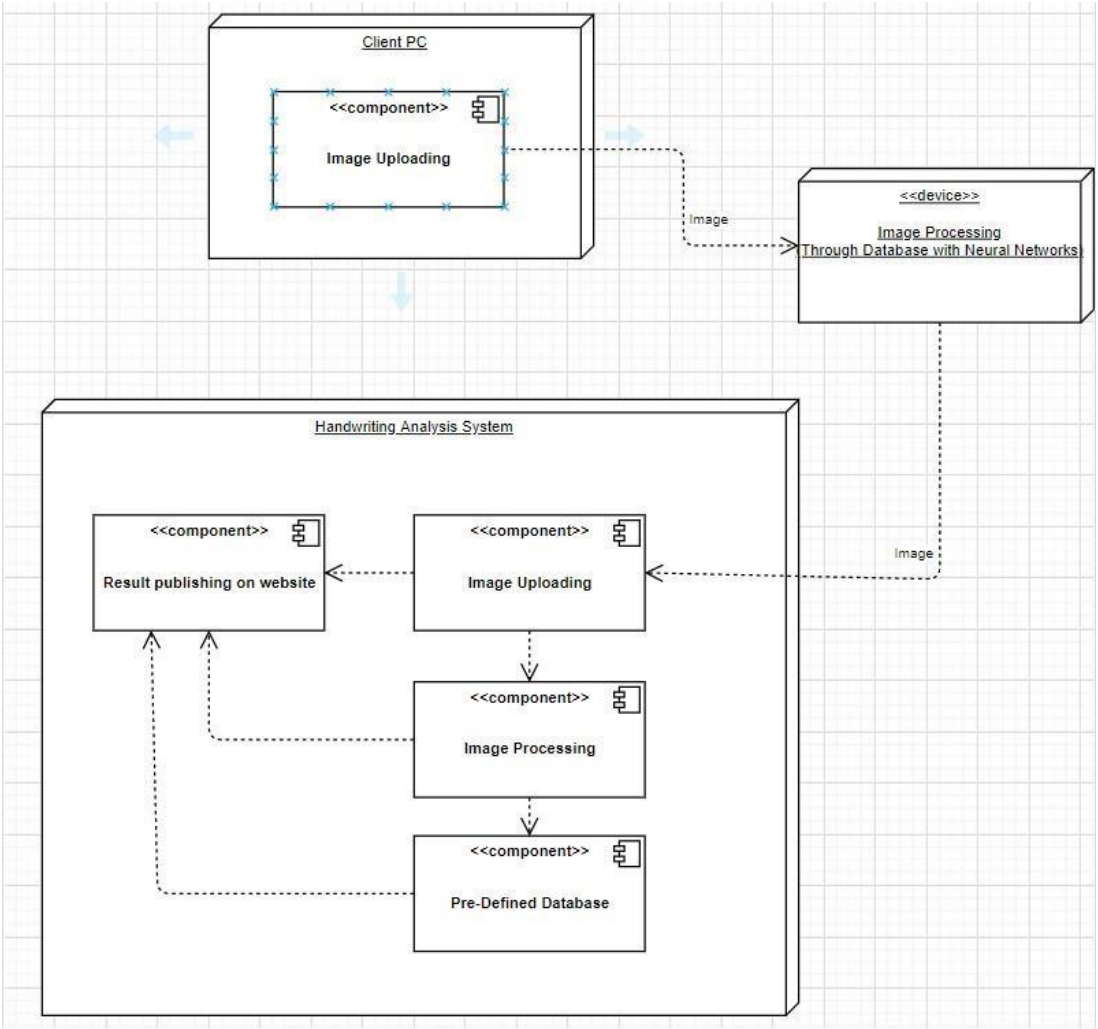


7.5. COMPONENT DIAGRAM

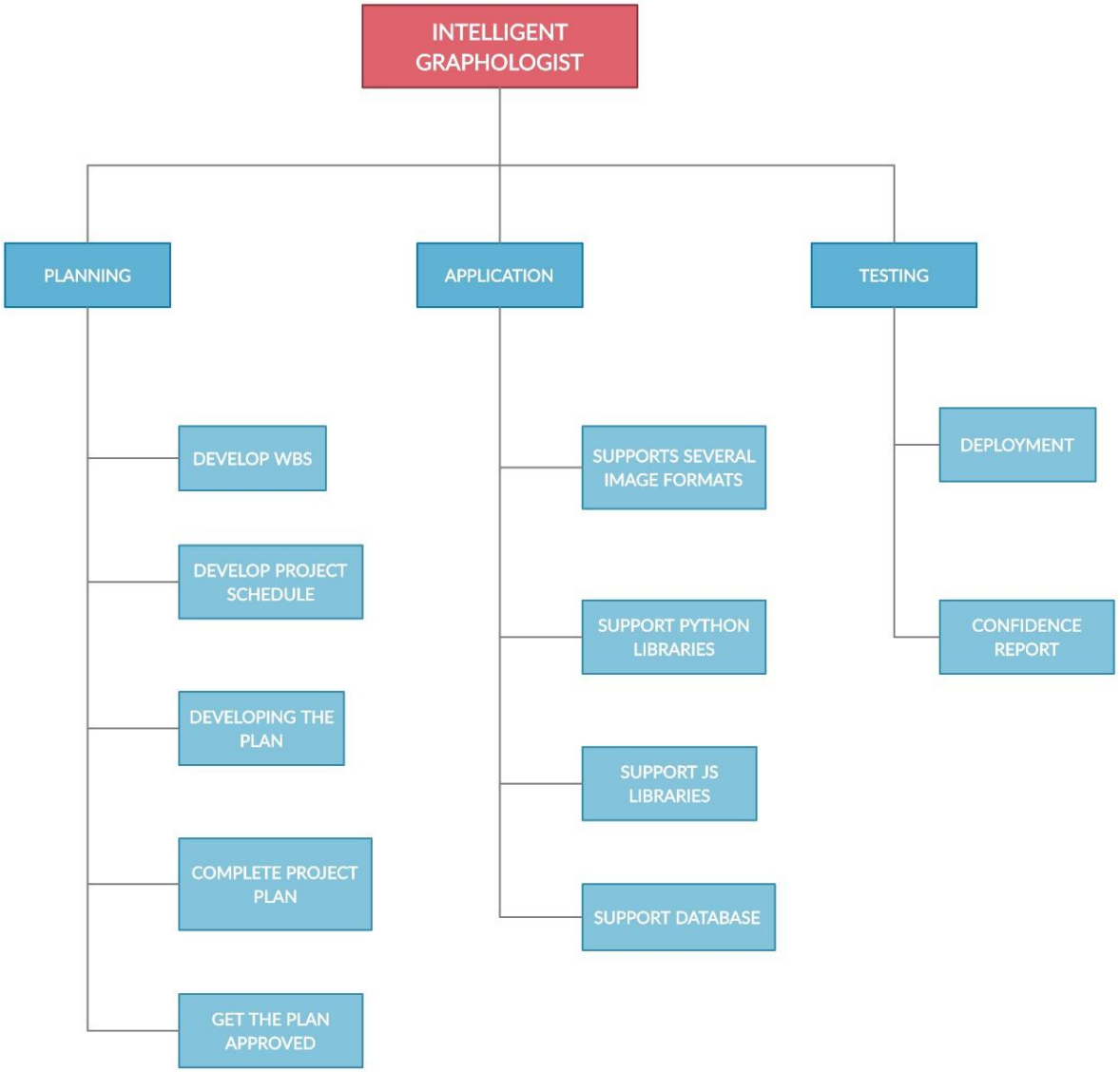
Deployment diagrams are used for describing the hardware components, where software components are deployed. Component diagrams and deployment diagrams are closely related. Component diagrams are used to describe the components and deployment diagrams shows how they are deployed in hardware.

Component diagrams are used to describe the components and deployment diagrams shows how they are deployed in hardware. All other diagrams portrayed above focus on software but these two diagrams focus on hardware components making them special.

Their purpose is to Visualize the hardware topology of a system, describe the hardware components used to deploy software components and describe the runtime processing nodes.



7.6. WORK BREAKDOWN STRUCTURE

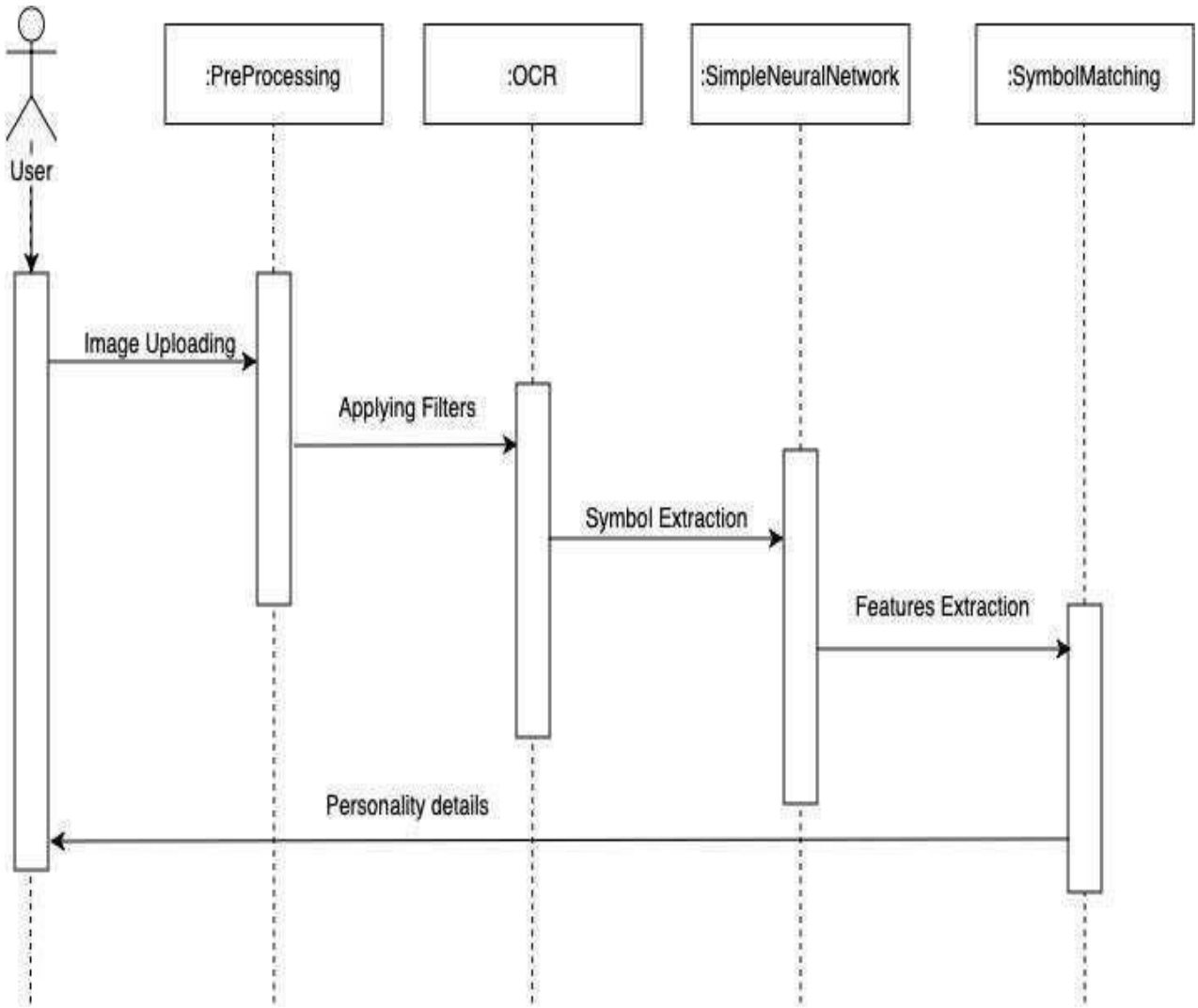


7.7. SEQUENCE DIAGRAM

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario.

Below's Sequence Diagram is an interaction diagrams that details how operations are carried out. It captures the interaction between objects in the context of a collaboration. It shows the order of the interaction visually by using the vertical axis of the diagram to represent time what messages are sent and when.

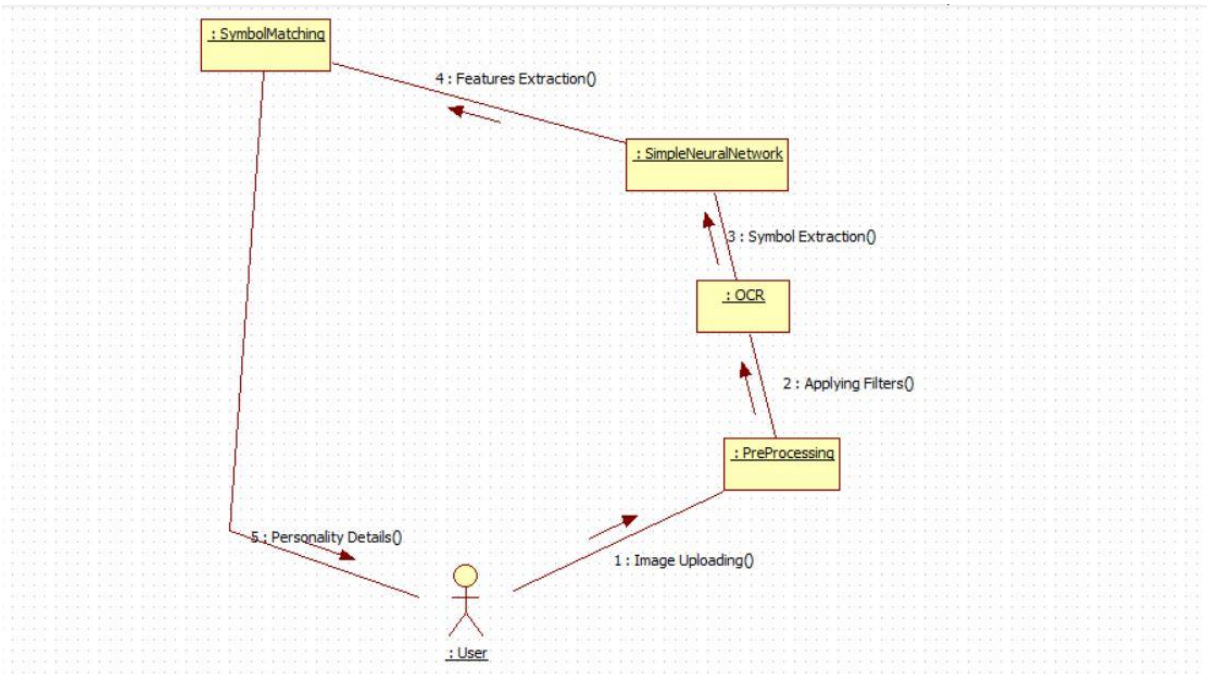
SEQUENCE DIAGRAM



7.8. COLLABORATION DIAGRAM

A collaboration diagram, also known as a communication diagram, is an illustration of the relationships and interactions among software objects in the Unified Modeling Language (UML). These diagrams can be used to portray the dynamic behavior of a particular use case and define the role of each object.

Below Collaboration diagram is used to show how objects interact to perform the behavior of a particular use case, or a part of a use case. They define and clarify the roles of the objects that perform a particular flow of events of a use case. They provide information used to determining class responsibilities and interfaces.



7.9. ACTIVITY DIAGRAM

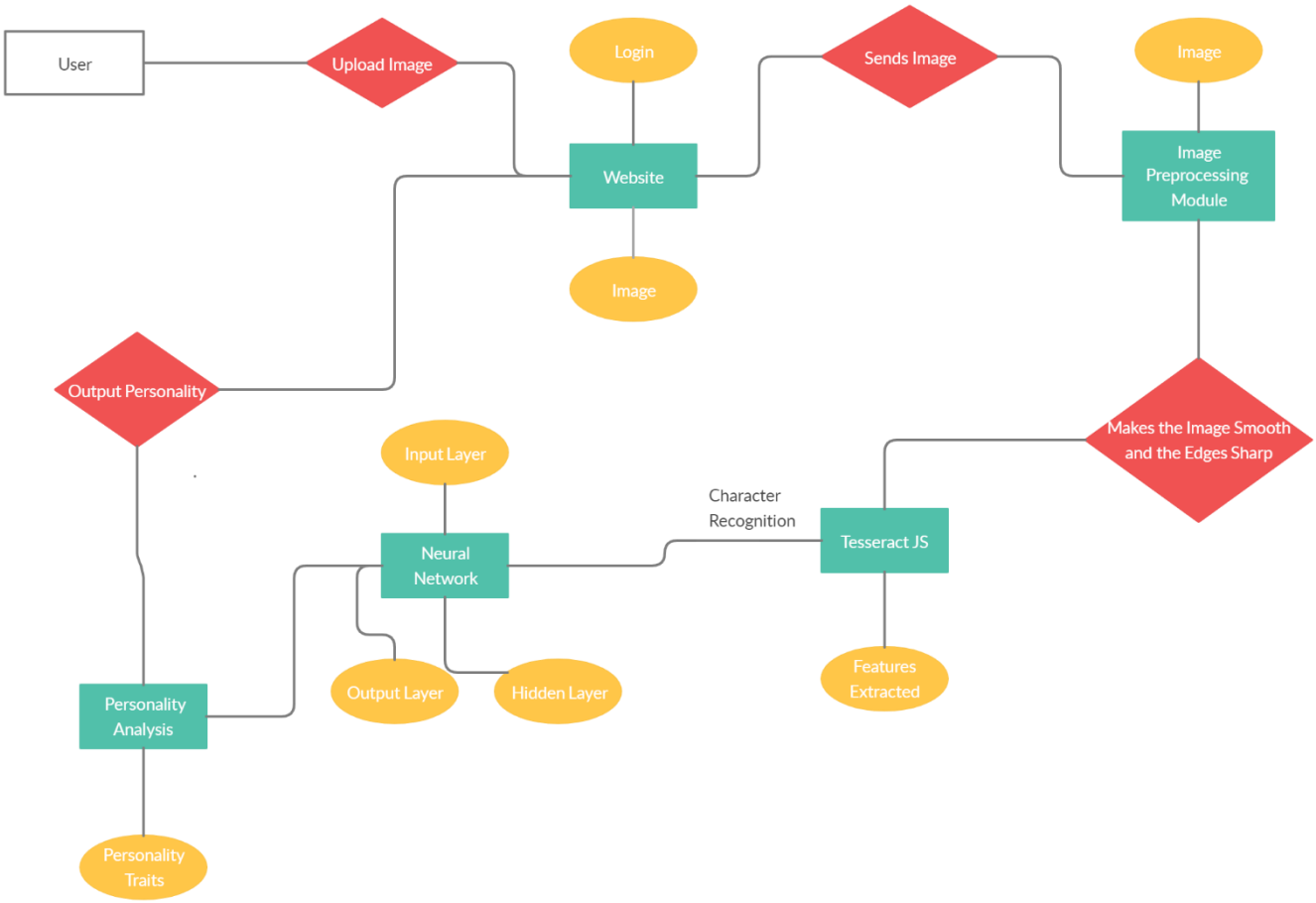
Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. It is used to describe the dynamic aspects of the system. Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another.

Below diagram is used to describe how activities are coordinated to provide a service which can be at different levels of abstraction. It is shows modeling of how a collection of use cases coordinate to represent business workflows.

7.10. ENTITY–RELATIONSHIP DIAGRAM

An entity–relationship model describes interrelated things of interest in a specific domain of knowledge. A basic ER model is composed of entity types and specifies relationships that can exist between entities. An ERD contains different symbols and connectors that visualize two important information: The major entities within the system scope, and the inter-relationships among these entities.

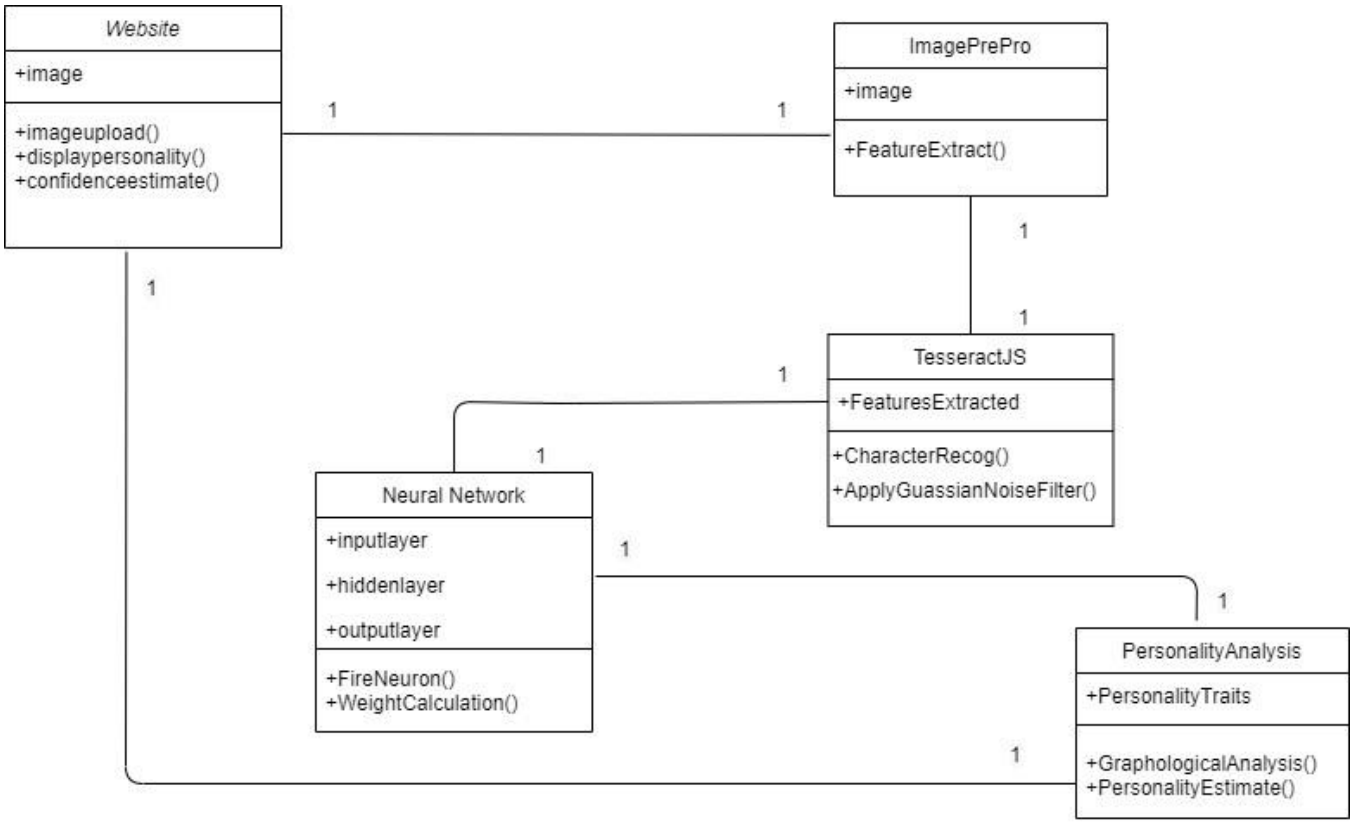
The main purpose of drawing an ER model is to present a design or blueprint of a database of the project.



7.11. CLASS DIAGRAM

The class diagram is the main building block of object-oriented modeling. It is used for general conceptual modeling of the structure of the application, and for detailed modeling translating the models into programming code. It is also used for data modeling. It is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations, and the relationships among objects.

The purpose of drawing class diagram for our project is to Show static structure of classifiers in the system, provide a basic notation for other structure diagrams prescribed by UML, to help developers and other team members and so that business analysts can use class diagrams to model systems from a business perspective.



8. TEST CASE TEMPLATE

1) Image Upload and OCR

Project Title: Graphologist	
TEST CASE TEMPLATE	
Test Case ID: GRAPH001 Test Priority (Low/Medium/High): High	Test Designed by: Team members Test Designed date: 09-10-2020
Module Name: Image Upload & symbol recognition System	Test Executed by: Team Members
Test Title: Test image upload and use OCR	Test Execution Date: 09-10-20
Description: Check if image upload and check if the OCR assessment works.	
Precondition: The user has an account on the page.	
Dependencies:	

Steps	Test Stage	Test Data	Expected Result	Actual Result	Status (Pass/ Fail)	Notes
1	The web portal opens					
2	Upload Image	User has an image of a written text	Image is uploaded	Image is successfully uploaded	Pass	
3	Click on submit					
4	Image symbols recognized from plain text		Application recognizes the symbols correctly	Application recognizes most of the symbols	Partially successful	
5	Upload noisy image	Noisy image				
6	Click on submit button		Image submitted	Image submitted	Pass	

7	Symbol recognition		Symbols recognised	Symbols not recognised	Fail	

Post-conditions:

The symbols obtained are sent to the Neural Network for further processing.

2) Graphological Analysis

Project Title: Intelligent Graphologist

TEST CASE TEMPLATE

Test Case ID: GRAPH002 Test Priority (Low/Medium/High): High	Test Designed by: Team Members Test Designed date: 09-10-2020
Module Name: Graphological Analysis	Test Executed by: Team Members
Test Title: Analysis of the CNN engine	Test Execution Date: 11-10-20

Description: To check if the module for graphological analysis works and to what degree of confidence.	
Precondition: The OCR engine test is successful.	
Dependencies: 1. Python 3.6 or higher.	

Steps	Test Stage	Test Data	Expected Result	Actual Result	Status (Pass/ Fail)	Notes
1	Upload Image		Image uploaded.	Image uploaded successfully.	Pass	
2	Run the program	The OCR test report is pass				
3	Graphological results		Graphology results are displayed	Graphological results are displayed with limited confidence.	Pass	

Post-conditions: None.

9. OUTPUT AND RESULTS

9.1. USER INTERFACE

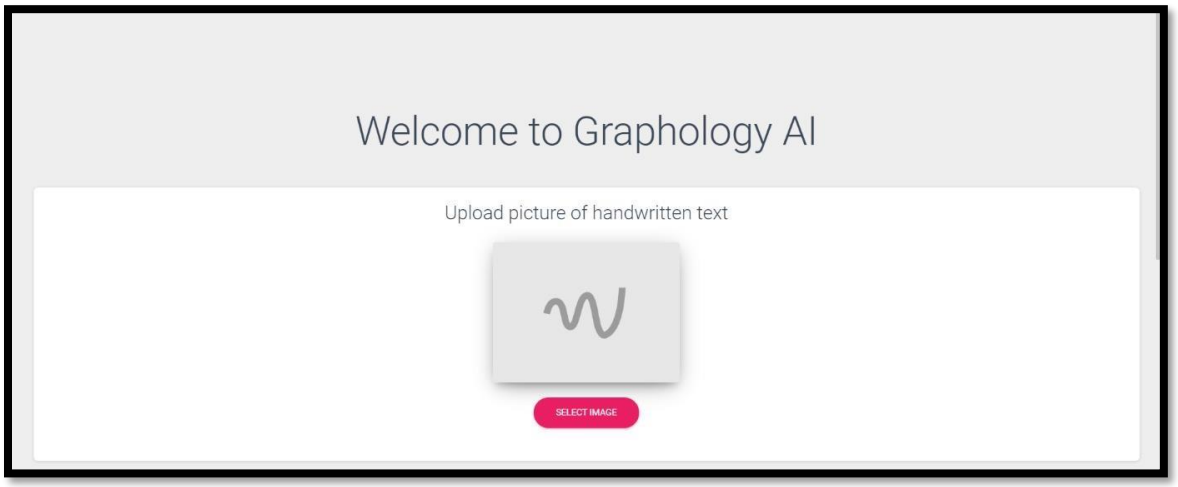
The user interface is minimalistic, simple and obeys the 4 golden rules of UI design which are:

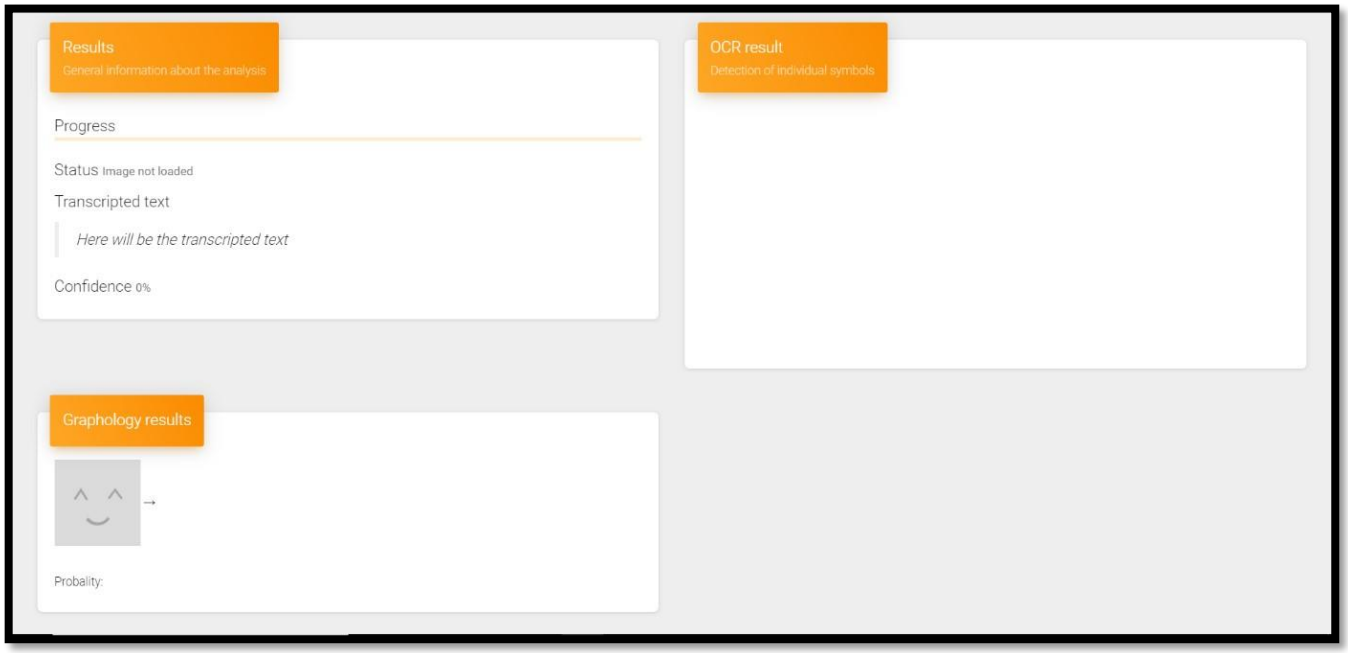
1. The user will find himself/herself/themselves in control of the UI.
2. The user shall not find any difficulty navigating through the project.
3. There is minimal cognitive load on the page.
4. The UI is consistent and pleasing.

9.2. IMPLEMENTATION

STEP1

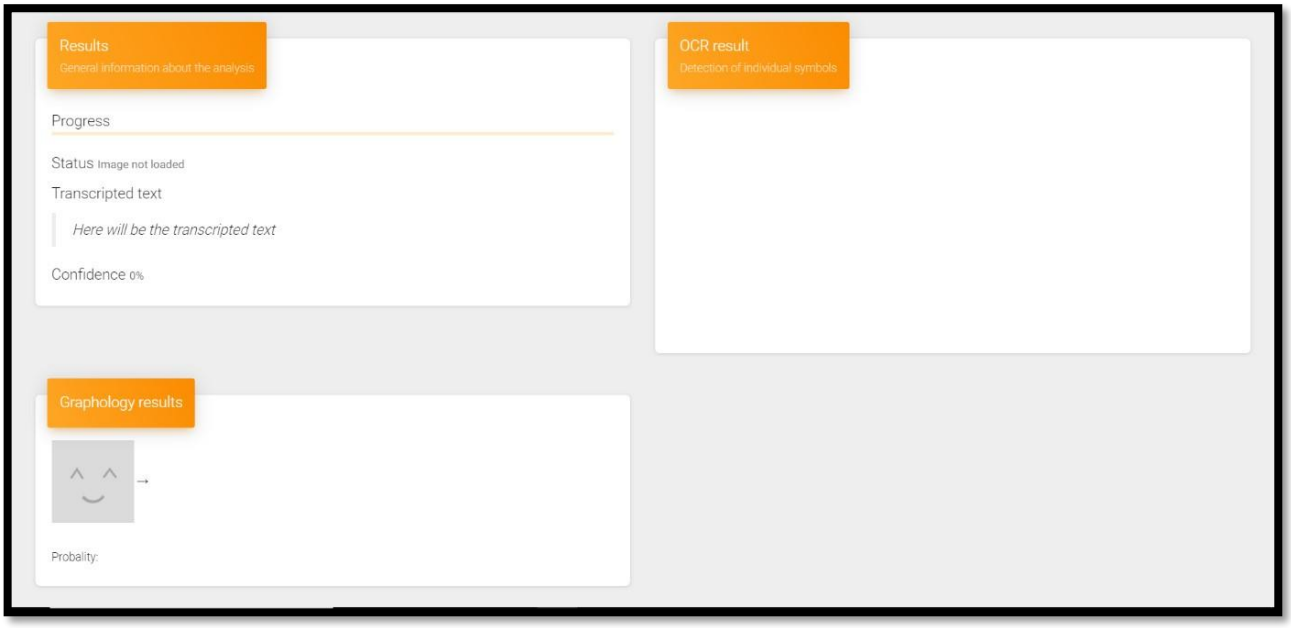
The analyzer opens waiting for the user to input the handwritten image for which graphological analysis has to be done.





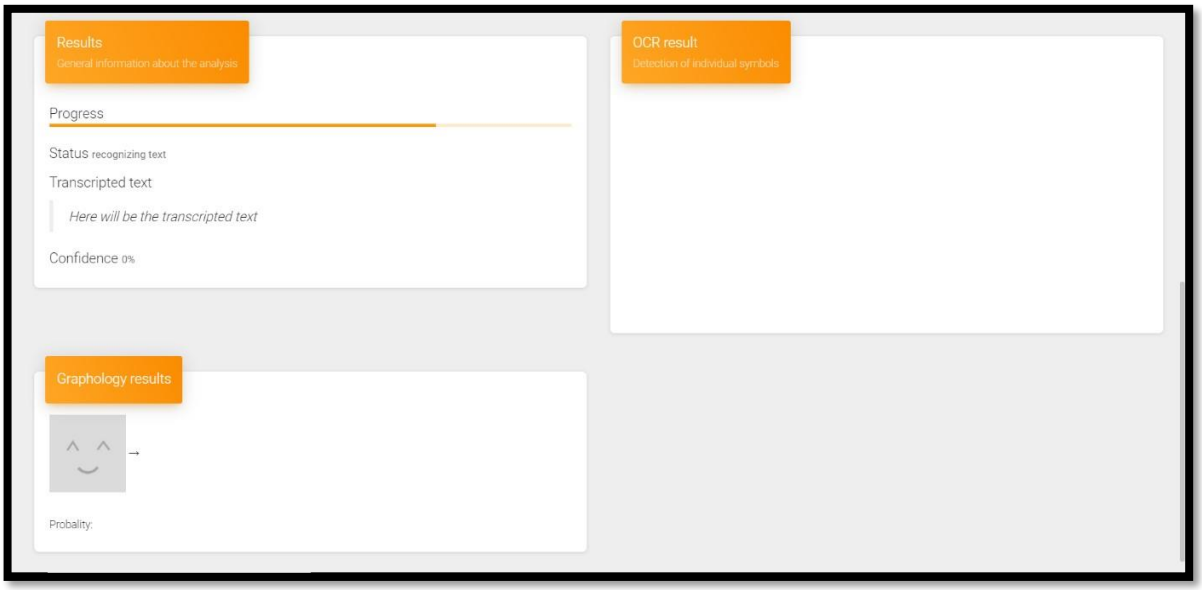
STEP 2

User uploads the handwritten input in the analyzer for which graphological analysis has to be done.



STEP 3

Analyzer starts processing of image by recognizing the character entered by the user. The progress bar shows how much the image has been processed.

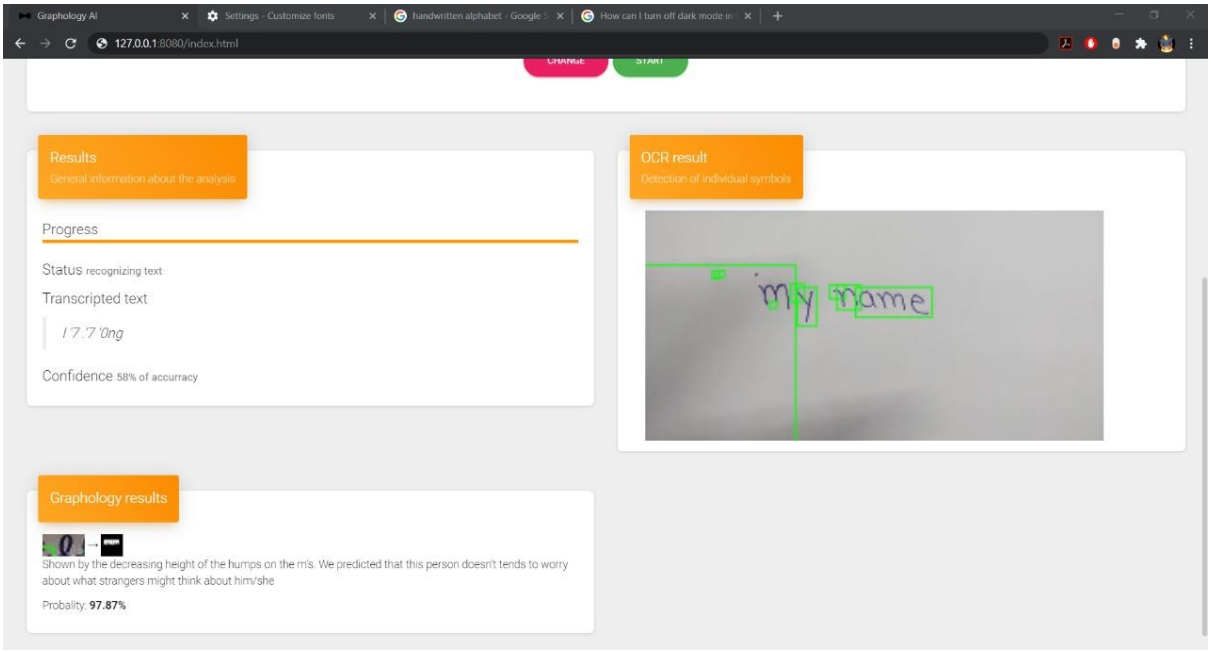


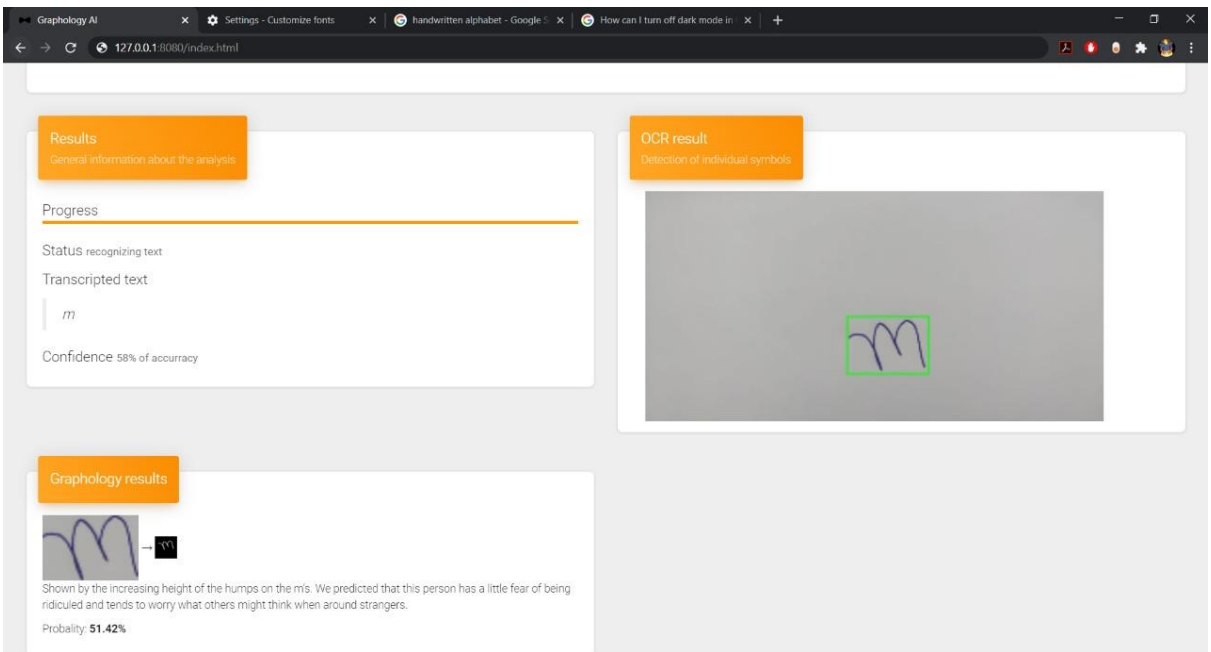
STEP 4

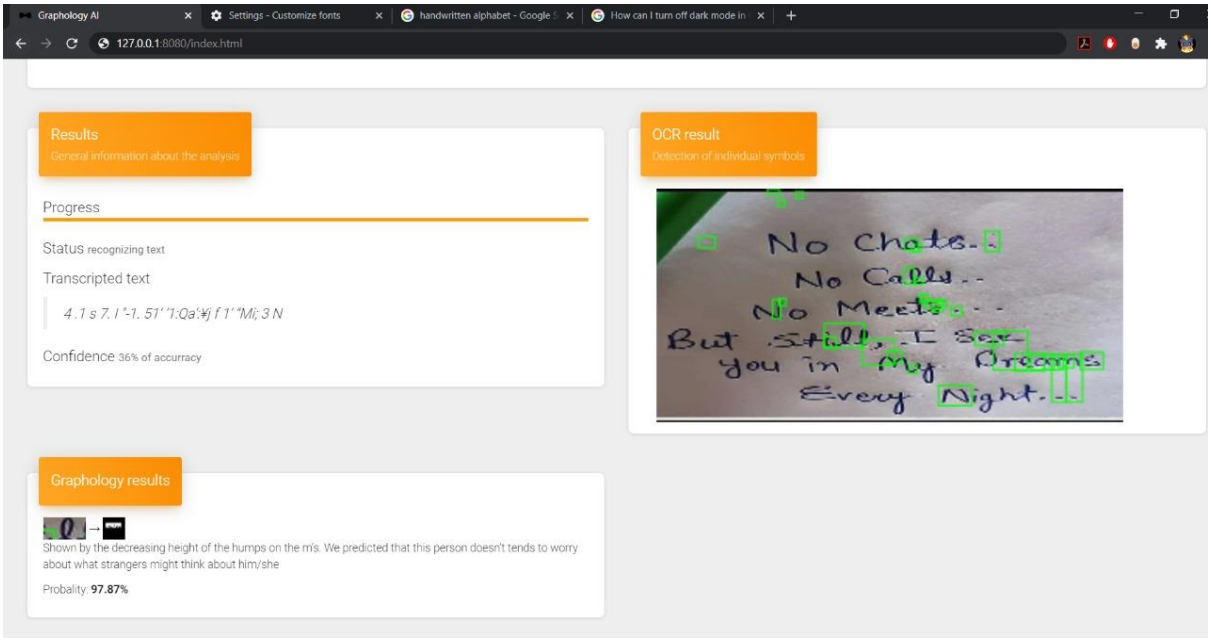
After the image is processed and the character is recognized by OCR, then the analyzer starts personality character check by the way the character has been drawn and comparing its width, height and other attributes from the dataset using neural networks.

After it finds the accurate match to the character, it displays the graphology results corresponding to the attributes found for the uploaded character.

Depending upon how close the attributes match with the uploaded character, it also shows the probability that the personality trait displayed is true along with the accuracy percentage.

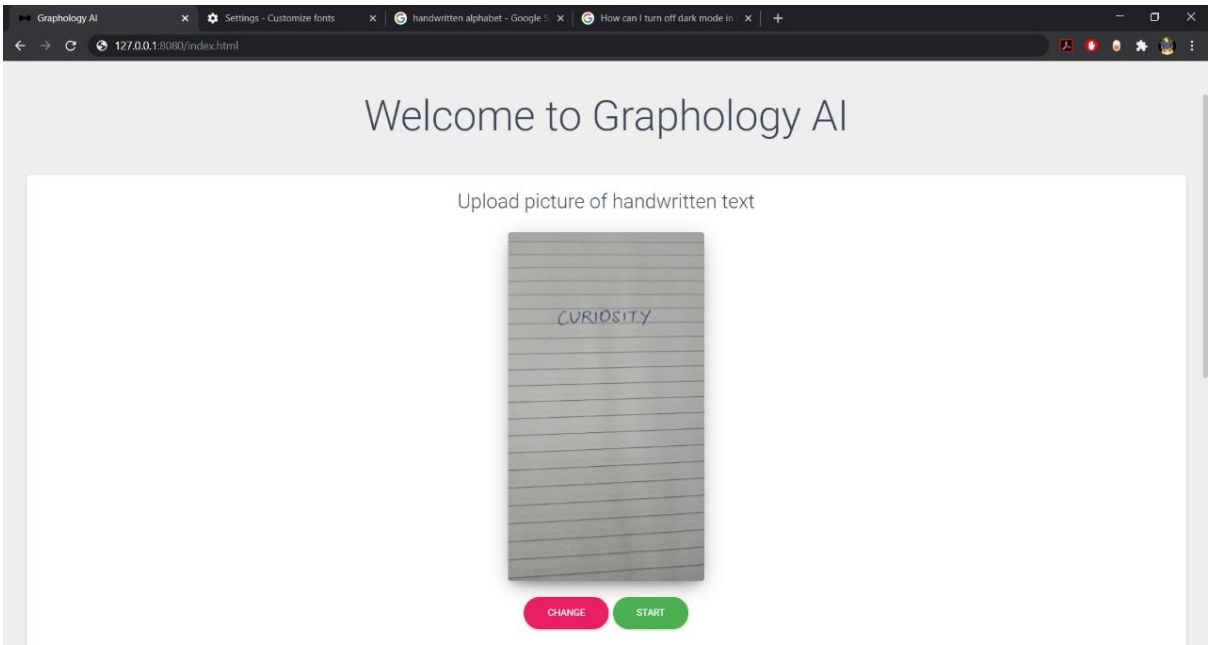


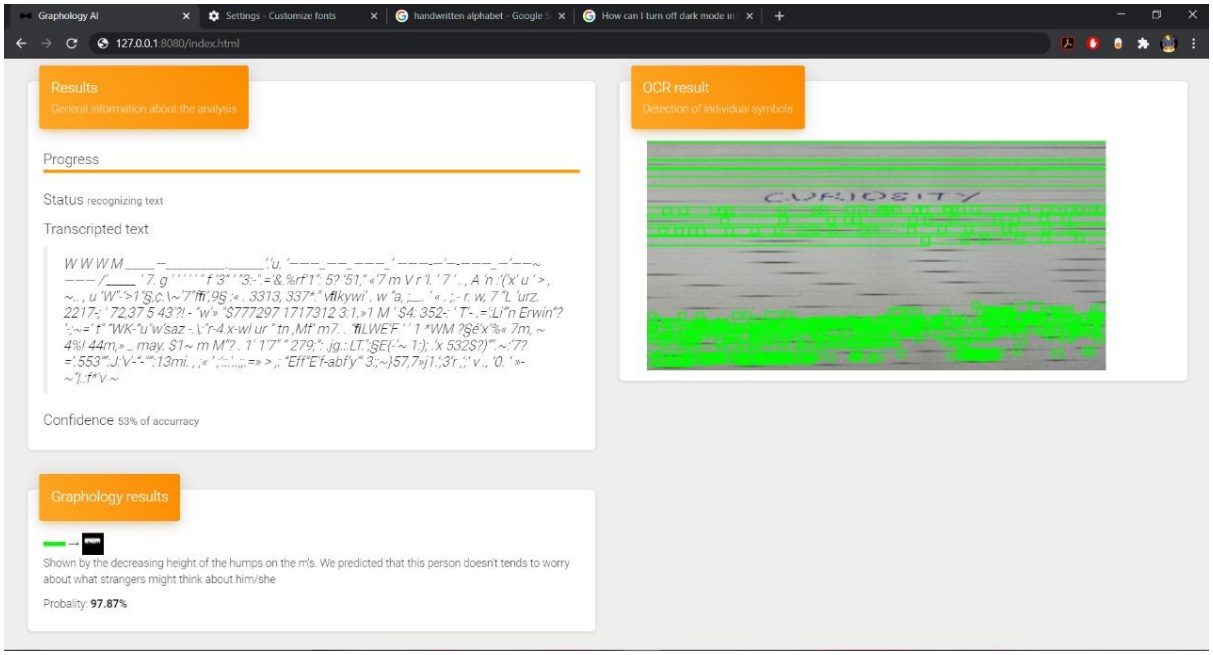




Noisy Outputs-

Noisy outputs are a problem with the project. The project fails to identify the characters partially or completely and hence even though an output maybe derived by the software, the output is clearly not trust-worthy.

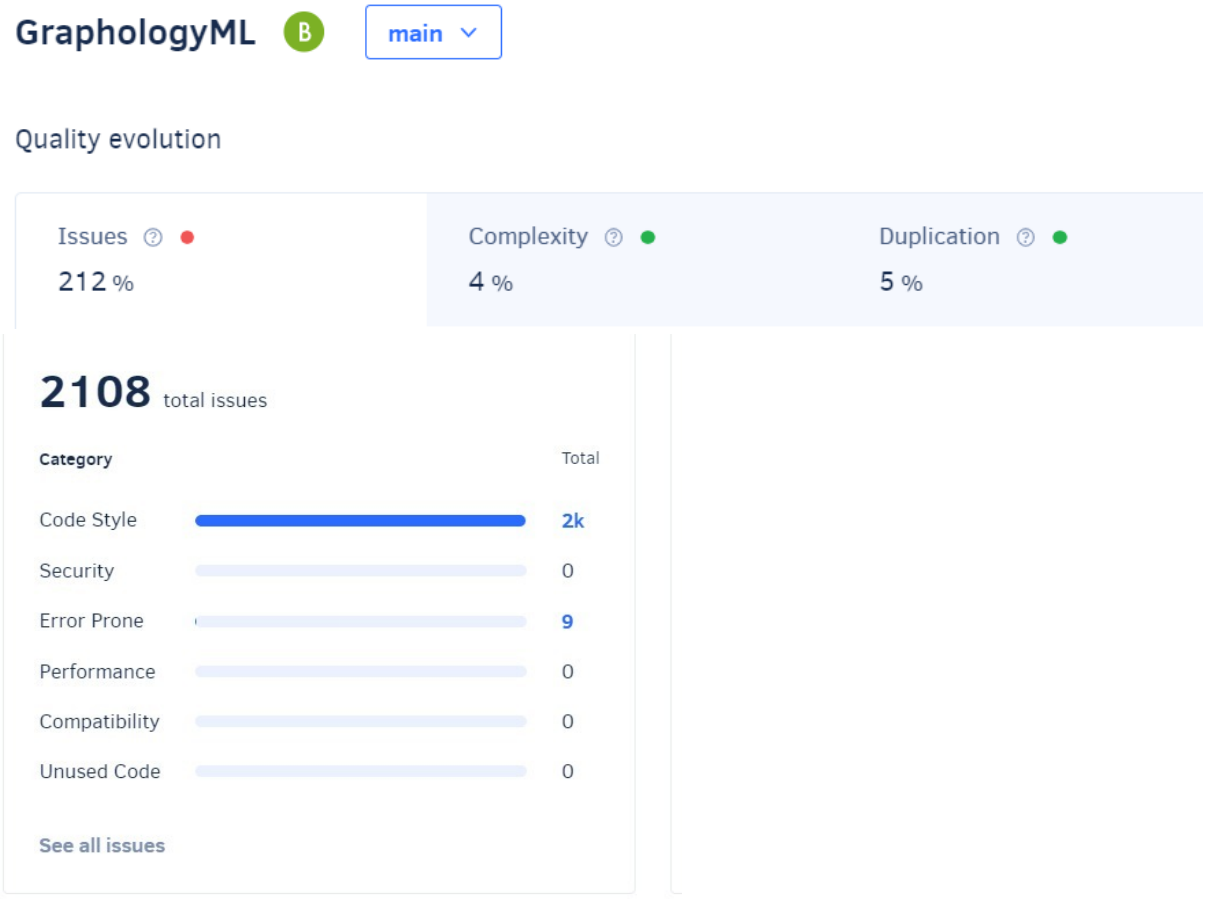




10. SOFTWARE – TESTING

Tools used:

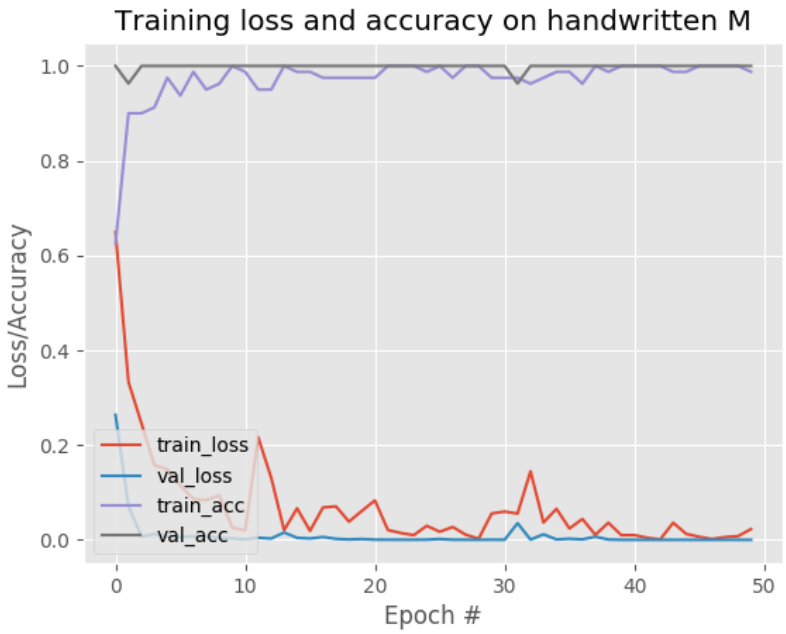
Codacy is a static program analyser. It is a standard tool available on GitHub. Version 1.0 of Intelligent Gaphologist showed the following results:



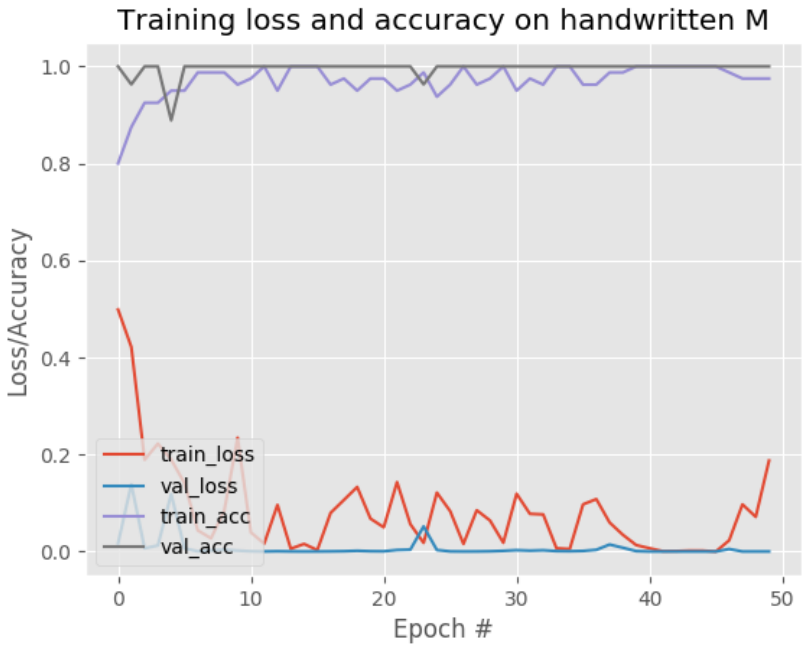
Most of these errors belonging to the category of code styling with very few errors pertaining to security vulnerabilities. This might be chiefly because we do not have any authorisation mechanism for the first version of the software.
Further, there was no need for a load testing as for version 1.0, the project maintains no records of the user or the images that the user has input.

11. PROBLEMS FACED

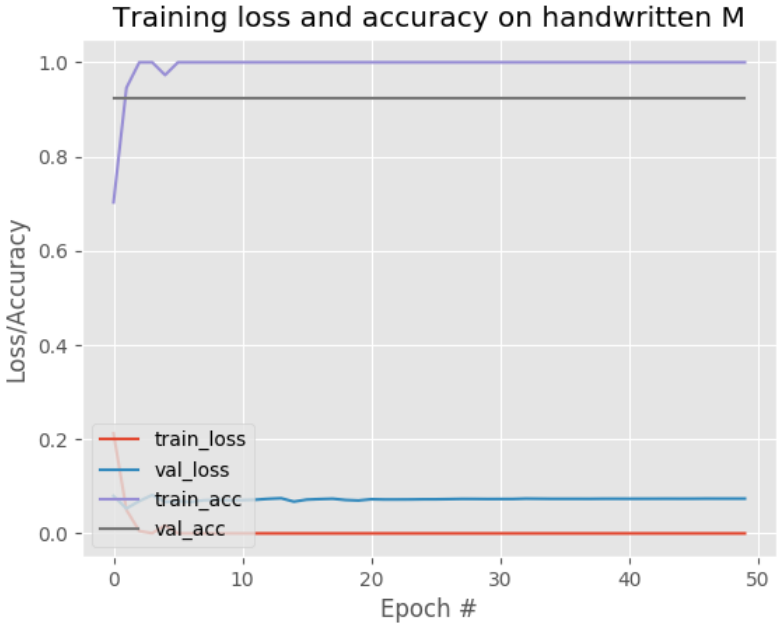
The following was the graph generated by python scripts during training. With an epoch of around 50, the training loss on handwritten data is pretty high when the epochs are increased from 0 to 50 and training is accelarated. A possible reason is that our dataset had only about 2000 data-values which is not a very large number for establishing a CNN based processing network.



Similarly, on an epoch of 50, when we tried to establish a neural network, the usage of CNN with a poor number of data-sets result in increasing loss value as epochs increase. The loss is prominent when the training is accelerated.



We observed similar results in accuracy vs loss in the training of a backend of the neural network which will be established over the OCR engine.



These three observations led us to creation of a simple neural network instead of a more complex one and training it under the system constraints instead of accelerating it.

12. CONCLUSION

Graphology - the study of handwriting and handwriting analysis - is now an accepted and increasingly used technique for assessment of people in organizations. Handwriting analysis is an effective and reliable indicator of personality and behaviour, and so is a useful tool for many organizational processes, for example: recruitment, interviewing and selection, team-building, counselling, and career-planning.

These features and interpretations provide a small but useful guide as to the way people behave, and particularly how they handle their social requirements. Understanding the personality through handwriting is a valuable way of making the best of both personal awareness and interpersonal situations for the benefit of all concerned.

This interpretation should enable people analysed to use the understanding gained, to help them live their lives to the highest level of satisfaction that they choose. In a professional or organizational context, graphology can play an important part in enabling working relationships to be forged that will enhance the quality of the group or team performance.

As a child we are taught to write, but it's not likely that we still write in the way we were taught. This fact itself helps to explain the reason graphology exists and why graphology can be used to interpret personality.

This software shall felicitate these developments at minimalist cost.

13. REFERENCES FOR LITERATURE SURVEY

[1] *Handwriting Analysis through Graphology: A review* by Laxmi Durga and Deepu R, IEEE, 2018

[2] *Emotion Predictor Using Social Media Text and Graphology* by Chhanda Roy & others, IEEE

[3] *Forensic Handwriting Analysis* by Dial, Jackie, Salem Encyclopaedia of Science, 2020

[4] *Mobile Application for Identifying Personality of Person Using Graphology* by Kukuh Adi Prasetyo and others, IES-KCIC, 2017, IEEE

[5] *Graphology: A New Marketing Research Technique* by JAMES U. McNeal, Journal of Market Research, 1967

[6] *MARKETING NOTES AND COMMUNICATIONS, Graphology and Marketing Research: A Pilot Experiment in Validity and Inter-judge Reliability* by PAUL E. GREEN VITHALA R. RAO and DIANA E. ARMANI, Journal of Marketing, 1971.

[7] Akshita Chanchlani, Pratima Kharade, Rutuja Kapase, Sonal Janvalka, Aakanksha Jaitly, "*Predicting Human Behavior through Handwriting*" International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887, Volume 5 Issue X, October 2017.

[8] Van-C uong Kieu,Dominique Stutzmann,Nicole Vincent "*Vacuity Measure for Handwritten Character Analysis*" 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR),IEEE(2017)

[9] Bayu Yudha Pratama ; Riyanarto Sarno "*Personality classification based on Twitter text using Naive Bayes, KNN and SVM*" 2015 International Conference on Data and Software Engineering (ICoDSE) , IEEE,2015

[10] Bayu Yudha Pratama ; Riyanarto Sarno " Personality classification based on Twitter text using Naive Bayes, KNN and SVM" 2015 International Conference on Data and Software Engineering (ICoDSE), IEEE,2015.

[11] Ricard Coll, Alicia Forn´es, Josep Llad´os, "Graphological Analysis of Handwritten Text Documents for Human Resources Recruitment," Computer Vision Center - Computer Science Department Universitat Autonoma de Barcelona Spain {rcoll, afornes, josep}

[12] Prachi Joshi, Aayush Agarwal, Ajinkya Dhavale, Rajani Suryavanshi, Shreya Kodolikar, "*Handwriting Analysis for Detection of Personality Traits using Machine Learning Approach*", International Journal of Computer Applications, (0975 – 8887) Volume 130 – No.15, November 2015.

[13] *Fast multi-language LSTM-based online handwriting Recognition,* by Victor Carbune and others, International Journal on Document Analysis and Recognition (IJDAR) (2020).

[14] *Graphology based handwritten character analysis for human behaviour identification,* ITET, Research Article, January 2020, Springer by Subhankar Ghosh and others.

[15] *Is graphology Useful in assessing Major Depression?* By Marco Giannini and others, Sage Magazine Psychological Reports, 2019.

[16] *Recurrent Convolutional Neural Network for Object Recognition* by Ming Liang and Xiaolin Hu, IEEE.

[17] *Automated Handwriting Analysis for Human Behaviour Prediction* by Manimala S and others, IJCSE,2016

[18] *Artificial Neural Network for Human Behavior Prediction through Handwriting Analysis* by Champa H N & Dr. K R AnandaKumar, International Journal of Computer Applications, 2010
[19] *Personality Analysis Based On Letter 't' Using Back Propagation Neural Network* by Sofianita Mutalib & others, Proceedings of the International Conference on Electrical Engineering and Informatics Institut Teknologi Bandung, Indonesia June 17-19, 2007
[20] *Writer Identification Based on Graphology Techniques* by Omar Santana & others, IEEE A&E SYSTEMS MAGAZINE, JUNE 2010.
[21] *A KNN-SVM Hybrid Model for Cursive Handwriting Recognition* by Cleber Zanchettin, Byron Leite Dantas Bezerra and Washington W. Azevedo, WCCI 2012 IEEE World Congress on Computational Intelligence, 2012