School of Electronic Engineering and Computer Science

MSc Computer Science - Software Engineering with Industrial Experience
Project Report 2018

Profiler Application using Sentiment Analysis

Soheil Novinfard

Queen Mary
University of London

August 2018

Disclaimer

This report, with any accompanying documentation and implementation, is submitted as part requirement for the degree of MSc in Software Engineering at the Queen Mary University of London.

It is the product of my own labour except where indicated in the text. The report may be freely copied and distributed provided the source is acknowledged.

Acknowledgements

I would like to express my deep gratitude to Professor Peter McOwan, my research supervisor, for his patient guidance, enthusiastic encouragement and useful critiques throughout the period of this project. My grateful thanks are also extended to; Professor Paul Curzon for teaching me some useful methods in interactive system design and related evaluation methods, Doctor Ioannis Patras for his helpful lectures in machine learning and to my dear friend Pegah Vaheb for pushing me to continue working hard during the writing.

Finally, I wish to thank my parents for their support and encouragement throughout the course of my studies.

Abstract

Social media channels such as Facebook, Instagram and Twitter have become integral parts of our lives. Large number of posts and pictures made by populations across the world, sometimes lead to phenomena of unstructured data streams. However, these data streams are popular with many people as they believe it aids in understanding of their own personality. Additionally, there are times when employers make use of social media to judge potential candidates' personalities with respect to the employment process, which ultimately can lead to the ideal team being built. The idea behind this project is to create a mobile application that helps users in predicting their own, or others, personalities. After collecting enough data from users, we applied sentiment analysis algorithms to this data and classified users' personality traits according to Big Five Factor model. Following that, the result is presented in an interface designed and evaluated during three iterations and evaluation process based on user interface principles.

Table of Contents

1 Introduction

Continuous activity on social networks makes an integral part of many people lives these days. We write a variety of posts on Facebook and Instagram or write tweets on different occasions, sometimes even all day long. This unstructured information has some direct meaning in each word and sentence, but also has an implicit interpretation which can be made. For instance, the following sentence indicates sadness in the speaker's expression:

"I deeply regret what I have said."

Sentiment analysis methods could extract the meaning in the text. This could be categorised by emotions for example sadness in above sentence, happiness, or angriness. After deciding on our target categories (classes), they can be predicted by machine learning approaches that exist for sentiment analysis.

Considering the social posts as a chain of sentences, we can find new meanings such as personality. There are different definition for personality in terms of psychology, but the majority of popular articles in this area accept the Big Five Factor model [1] [2] [3]. This model believes that each person's personality can be measured by extroversion, neuroticism, agreeableness, conscientiousness, and openness. In this project, each of these personality traits is defined as class/category for our machine learning classifiers. The Naive Bayes and Support Vector Machine are used as our classifiers because this classification task is a supervised learning type with "positive" and "negative" label for each personality trait. Then, models are evaluated by using 10-fold cross-validation approach. After that, the results measuring the accuracy of predictions are compared with two other state-of-the-art models.

A mobile application is made by combining the trained model with a proper user interface in the next phase of the project. The user interface was created during three iterations. A low-fidelity prototype is designed in the first two versions of the design, and then it is evaluated by user-based and expert-based methods. After that, two high-fidelity prototypes were prepared by applying the user interface evaluation results in each step. Rigorous expert evaluation method is performed on our high-fidelity by three experts. According to our mobile application characteristics, a suitable colour scheme was decided to use in our design.

For using the trained model as a machine learning engine in the application, two different trending methods are discussed. One revolves around to deploying the models outside the device in a server and communicating to receive the results from an external server over the Internet by a network layer. Whereas, the other method is to implement a Machine Learning engine inside the device using each mobile platform facilities for this purpose. The former method is used in the final design.

2 Literature Review

A lot of information is created by users in social networks, web forums and wikis these days, thanks to the technology. By applying Sentiment Analysis (SA) with Natural Language Processing (NLP) techniques on a dataset of text and essays categorised by different personalities, the personality of people can be predicted from this information. These are like three vertexes of a triangle which this work is based on. In the following sections, each of these vertexes is described in more details.

2.1 Sentiment Analysis using Natural Language Processing

Some of the data found on the Internet is about opinions and reviews of users about different products and services. This information is useful when we extract the sentiments, opinions, and emotions from them. It is where sentiment analysis is used for. The related applications cover a wide range of subjects from information retrieval systems to decision support systems. Some organisations and companies find how people think about their products and services over the time by applying sentiment analysis to product reviews. It is valuable data for their marketing section which helps them to define their strategies. Search engines have provided more accurate results by combining the information retrieval with sentiment analysis towards the user's opinions about the searched terms [4].

Sentiment analysis techniques can be performed on the document level, sentence level or word level [5]. In a vast majority of recent researches, the focus is on document-level where we identify sentiment polarity on the whole document, e.g., finding positive and negative reviews about a specific service. To find the document level sentiment, we can analyse each sentence in the document:

I was *unable*$^{-}$ to install it after 2 hours reading the setup manuals.

Amazing^{+}! I will *recommend*$^{+}$ this book to my friends.

In preceding examples, lexicons help us to identify the sentiment in sentences. There is no significant difference between sentence-level and document-level analysis due to the fact that sentences are just small pieces of documents. However, some documents present different opinions, while most sentences give only one opinion [6].

The data source used for input of sentiment analysis could be fed from customer reviews, political views, marketing and financial data [7], news [8] or social media activities such as Facebook posts [9] and tweets [10].

There are different machine learning approaches to sentiment analysis including supervised and unsupervised learning techniques. Walaa Medhat [11] illustrated and categorised them as shown in Figure 1.

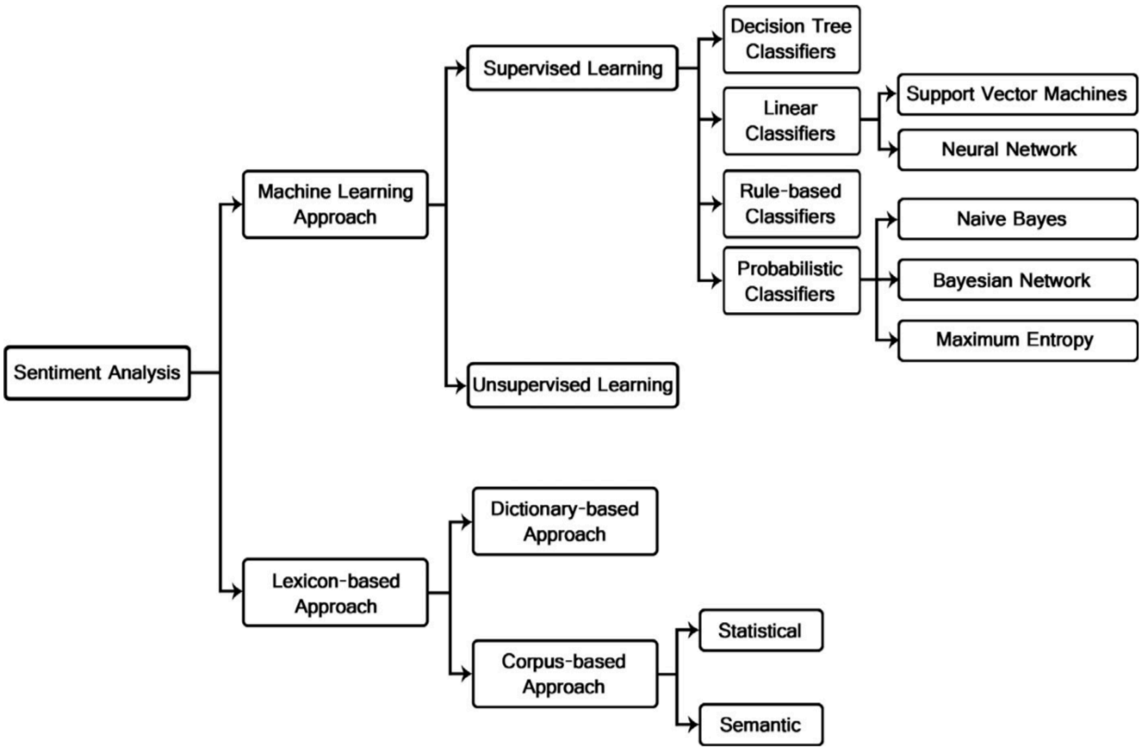


Figure 1 - Sentiment analysis approaches [11]

Personality detection which is discussed in our project (refer to section 2.2) has some specific number of categories; thus, unsupervised approaches are not usable in this work. We have used some supervised learning methods including linear and probabilistic classifiers.

2.2 Personality Detection

There is a wide range of models for recognising different personalities in a person, but the most acceptable one in Big Five model which is used in many researches until present such as in the works of McCrae [2], Digman [1] and Goldberg [3]. Although some of them use different names for different personalities, all of them refer to five personality traits. People in these five personality traits can be described as follows [12]:

1) Extroversion (EXT): This trait is well recognisable in people who talk a lot like someone energetic who tries to broaden their communication skills through networking. An extravert is someone confident and sociable in depth.
2) Neuroticism (NEU): They are more sensitive to emotional situations such as stress and anxiety and can get angry quickly. Their resilience to emotion makes them stable regarding their emotions.
3) Agreeableness (AGR): People in this trait are more compassionate with others. They are generally trustworthy and humble. Less agreeableness can be seen in an argumentative person.
4) Conscientiousness (CON): They tend to act more organised and disciplined manner. They have a more goal-directed lifestyle.
5) Openness (OPN): Their character is best described by their curiosity, and tendency to art, new experiences and ideas. They are more involved in risky actions. In addition, they try to avoid the routines as much as they can.

These are not mutually-exclusive traits; thus, moderate people are supposed to belong to more than one personality trait.

2.3 Competitor Analysis

There are some mobile applications for personality detection in both iOS and Android platforms. These applications use questionnaires to find the personality of people. One of the drawbacks of such approaches is the time the user should spend to answer the massive amount of questions and fill in the forms. However, in our proposed model we do not ask any questions from the user. We rely on the sentiment analysis techniques to auto-detect the personality traits based on the social channel activities of the user. Although the prediction accuracy may be a little lower than one which is achieved by direct questionnaire, the corresponding processing time is relatively lower. Moreover, in our application, there is an option for the user to enter a text or essay relating to them, in case they do not have activities on social networks.

3 Requirement Capture and Analysis

The first step to creating any software program including a mobile application is to define the problem being undertaken. Then, based on the problem definition, the specification and requirement of the software are made clear and comprehensible. There are different approaches to generate requirements, some of which rely more on users or client by using interviews and surveys and others rely on designer's knowledge and experience. In the document level sentiment, the simplest software engineering process leads to a bunch of software features. Simple yet not a suitable approach, because designing based on the requirements of all groups of users would create software which is not usable for anyone. To tackle this issue, we can describe the user, their needs and goals in a scrutinizing manner and drive the design based on that. This is what Cooper [13] suggested in his "Persona" concept:

"Develop a precise description of our user and what he wishes to accomplish" [13].

The main advantage of persona comparing with average user and aggregate user is that although persona is not a real person existing in the world, we do not develop our design according to a factious unacceptable user with an unrealistic group of requirements.

We created the main persona for our design named "Alex" as follows:

Alex *is an electrical engineer. He graduated 5 years ago and now he is working for a telecommunication company. Reading psychology books is one of his hobbies. He is curious about celebrities and follows their social channels. He can use simple smartphone applications, but any complex mobile apps are removed from his phone just after installing. Recently the company makes him responsible for creating a new team for power dense machines sector. He wants to balance the team of different* *attitudes and character to create a better team, but he does not know how to find applicants' approximate personality.*

Analysis of Alex's persona (and therefore any group of people with similar lifestyle, character, and life) points out that we can create a mobile application to satisfy his needs. This mobile application has the specifications below:

a) It predicts the user's personality and shows the result.
b) The input source for the user's data used for prediction could be a direct sample of essay or text of the user or any other person.
c) The user can log-in to their social network account, for example, Facebook, Twitter, or Instagram. Then the application fetches some posts (tweets) and uses them to predict their personality.
d) Some sample personality analyses performed by the application on celebrity posts are available in a "Celebrity's" section.

The flowchart of the application's work-flow is shown in Figure 2.

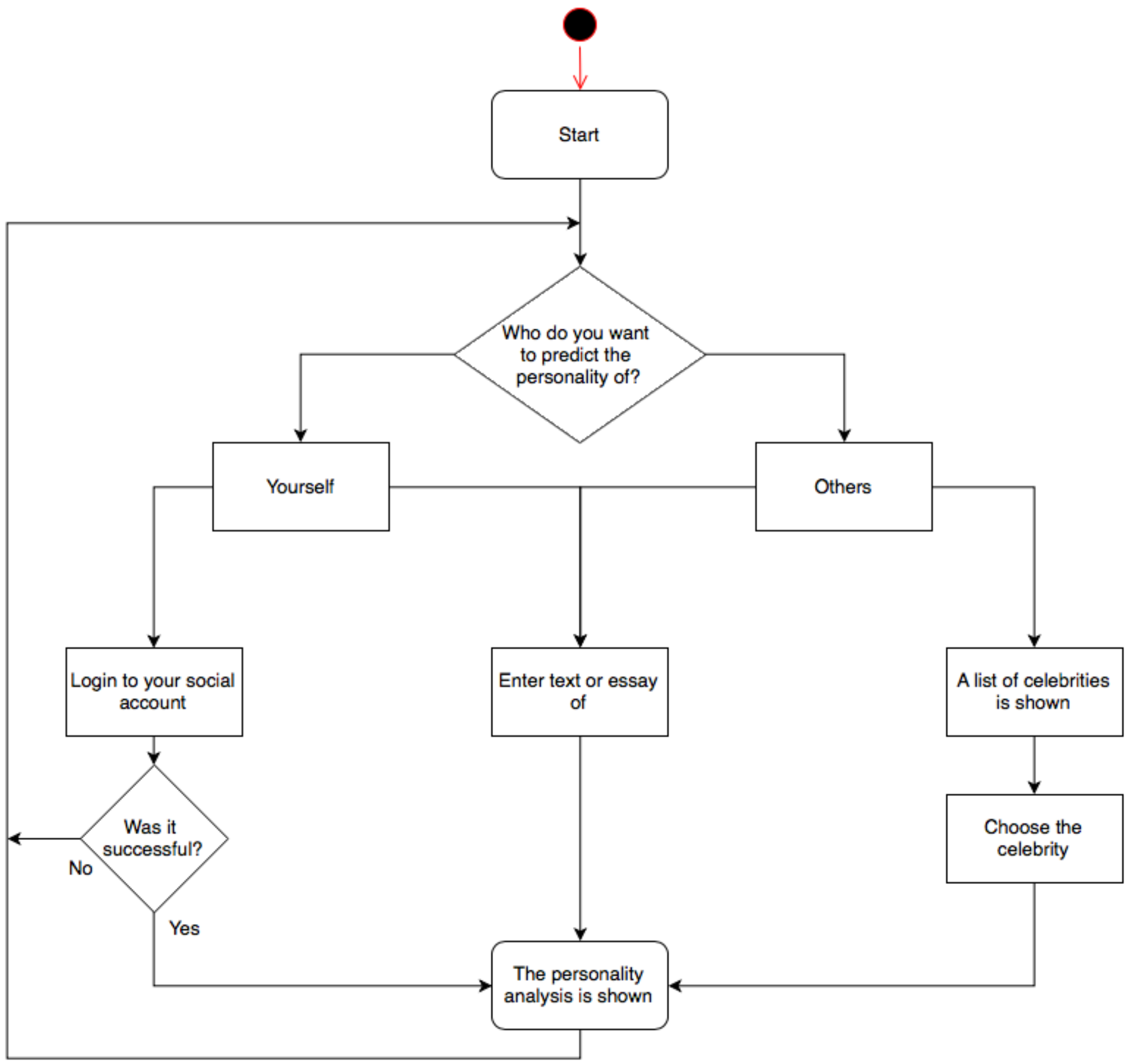


Figure 2 - Flowchart of application's working flow

To reach this goal and have better insight on each part of our solution, we divided the project into three following scopes:

1) Algorithm and Model: Methods and models for analysing the user's personality
2) Interface: An interactive interface in which the user's personality can be easily reviewed
3) Mobile Application: An application that integrates the model and interface to use in the user's smartphone

4 Design

The proposed solution to the problem with the previously stated requirements is described in this chapter, categorised by the scopes.

4.1 Algorithm and Model

Our method starts with pre-processing the data. Next, features extraction is performed in the model and finally, we classified the document for each personality category.

4.1.1Input Dataset

To automate personality recognition, Pennebaker and King provided a dataset of essays which is tagged with different personalities [14]. The dataset includes 2,465 stream-of-consciousness essays. The personality categories are based on the Big Five Factor model overviewed in 2.2 with EXT, NEU, AGR, CON and OPN labels.

4.1.2 Pre-processing

At the first stage, we pre-processed the dataset by removing the accent letters such as “é”. We also excluded digits, question and stop marks. Additionally, all words in documents are converted to lower-case to generalise the words.

After cleaning, we transform each document in the dataset to a series of sentences.

4.1.3 Feature Extraction

We have some documents, streams of strings. However, we need some numerical representation of the text to use it as features in the classifier.

In our method, we used the "bags of n-grams" for feature extraction in the first step. Text is first tokenised and separated by giving an identification for each token. White-spaces and textual marks are some of the separator delimiters which are applied in the tokenizer. The occurrences of each token are counted in the whole document. Then these numbers were normalised regarding how many times they have been seen in documents.

For tokenisation, n-grams are used by unigram and bigrams characterisation of documents. Although we tokenised text up to a 5-gram sequence in a limited number of models, it does not show significant improvement to our model's accuracy. Therefore, we skipped more than 2-gram sequences due to the matter of model speed-up.

For the second step in feature extraction phase, the "term-frequency" times "inverse document-frequency" $(tf - idf)$ have been applied to feature which conveys the importance of words comparing with document [15]. It is called "weighting factor" in some contexts. The term-frequency (tf) is row count of a token in the document. The inverse document-frequency (idf) is defined as below:

$$idf(t) = \log\frac{1 + n_d}{1 + df(d,t)} + 1$$

In this formula, given the token t, $df(d,t)$ is the frequency of t among all documents and n_d is the number of documents.

The output at this stage is a list of tokens with corresponding numerical representation called feature vector.

4.1.4 Classification

We have two classes, “positive” and “negative” for each personality. We train each classifier 5 times for 5 different personalities. As mentioned in 2.1, we used some supervised learning approaches for this work.

Given our vector of a tokenised document as $\overline{X}$, and related vector of coefficient s(weights) related to the X with the same dimensionality as $\overline{A}$ and scalar bias as B, then $P = \overline{A}.\overline{X} + B$ acts as the predictor of the “linear classifier”. For our model, this classifier creates a hyperplane on which we have positively labelled points on one side and negatively labelled points on the other side of it. “Support Vector Machine” is one of the linear classifiers.

4.1.4.1 Support Vector Machine

The first classifier used in this project is a linear classifier named “Support Vector Machine” (SVM) (or support vector networks) with stochastic gradient descent learning rate. For simplification, suppose $X = \{x_1, x_2, \cdots, x_n\}$ is a set of input points and $Y = \{Y_1, Y_2, \cdots, Y_d\}$ is the corresponding output. Y_i can be either 1 if it belongs to a class (positive in our model) or -1 if it is not a member of it (negative in our model). The hyperplane which divides the points in our classifier can be written as follows:

$W.X - b = 0$

where W is called a normal vector. By division of b over the length of W we can calculate the offset of the hyperplane from the origin which is shown in Figure 3:

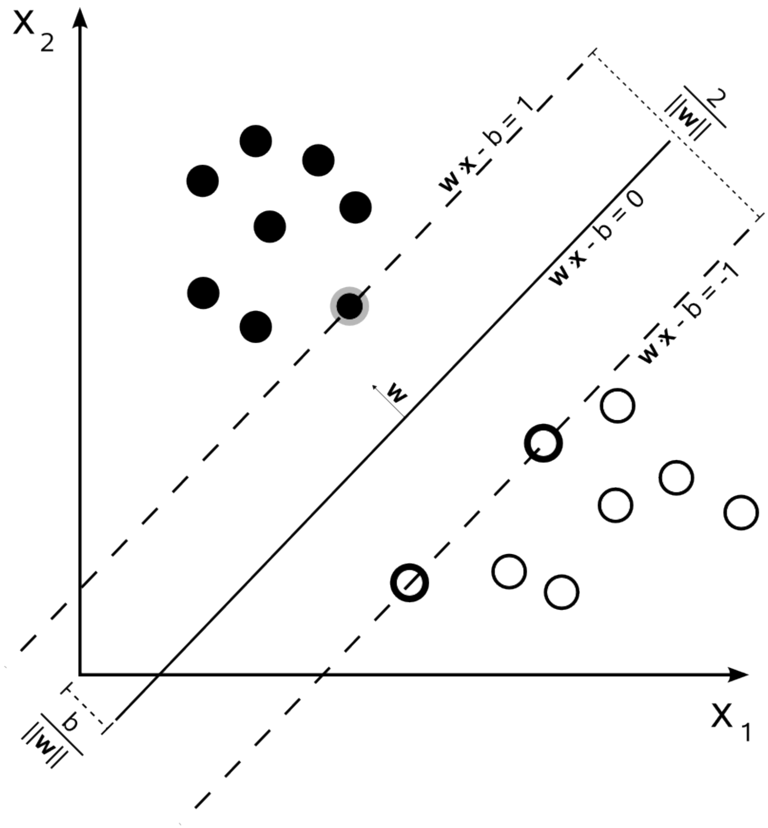


Figure 3 - support vector machine representation for classification [16]

Text classification is fitted well with SVM due to the fact that we normally have a large sparse vector of features of text which they are correlated with each other indirectly.

4.1.4.2 Naive Bayes classifier

The second classifier used in this project is Naive Bayes. The Naive Bayes can often outperform more complex classification methods regardless of its simplicity, and it is particularly suitable when the dimensionality of the inputs is high.

The Naive Bayes classifier is a robust algorithm that is commonly used for Sentiment Analysis. Due to its great result in multiclass problems and independence rule, Naive Bayes classifiers used in text classification have a high success rate comparing to other algorithms.

The Naive Bayes classifier assumes the features are independent. It first calculates the probabilities for every part. Then, it selects the one with the highest probability. This classifier method is based on the posterior probability which is also known as Bayes Theorem.

The Naive Bayes model that is used in this project is the Multinomial type, and its math is elaborated upon hereunder:

Given a set of variables, $X = \{x_1, x_2, \cdots, x_d\}$, the aim is to create the posterior probability for the event C_j among a set of probable outcomes $C = \{c_1, c_2, \cdots, c_d\}$. In a simpler word, X is the predictor and C is the set of categorical levels that are presented in the dependent variable. Using the Bayes' rule:

$$\mathrm{P}(C_j|x_1, x_2, \cdots, x_d) \propto \mathrm{P}(x_1, x_2, \cdots, x_d|C_j)\mathrm{P}(C_j)$$

where $\mathrm{P}(C_j|x_1, x_2, \cdots, x_d)$ is the probability that X belongs to C_j known as the posterior probability of class membership. Considering Naive Bayes' assumption that the conditional probabilities of independent variables are statistically independent, decomposing the likelihood to a product of terms is allowed:

$$\mathrm{P}(X|C_j) \propto \prod_{k=1}^{d} \mathrm{P}(x_k|C_j)$$

and then, rewrite the posterior as:

$$\mathrm{P}(C_j|X) \propto \mathrm{P}(C_j) \prod_{k=1}^{d} \mathrm{P}(x_k|C_j)$$

With the usage of the above Bayes' rule, a new case X is labelled with a class level C_j that has the highest posterior probability.

Although the assumption that the predictor variables are independent is not always reliable, it dramatically simplifies classification problems, because it provides separate (for each variable) calculations of the class conditional densities $\mathrm{P}(x_k|C_j)$, in the way that it reduces a multidimensional task to one-dimensional task. As a result, Naive Bayes reduces a high-dimensional density estimation task to a one-dimensional kernel density estimation. Moreover, the assumption does not seem to significantly affect the posterior probabilities, particularly in areas near decision boundaries; therefore, it leaves the classification task unaffected.

The training process is demonstrated in Figure 4:

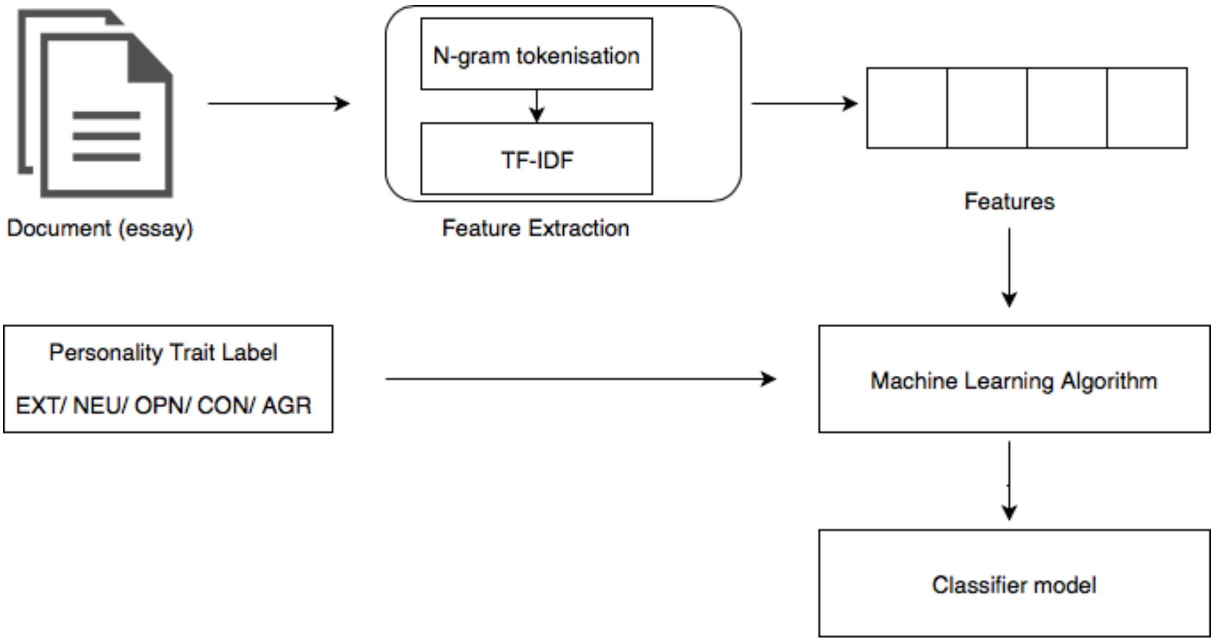


Figure 4 - The model's training diagram

4.2 Interface

One of the best techniques for designing an interactive user interface for anything, not only software, is using prototypes. Houde and Hill [17] introduced the prototypes as a way of defining design problems. They created a baseline by which we can evaluate the design in the cycle of interactive interface design.

The iterative process was used for a gradual improvement of the design, since it is hard to implement everything correctly at first. Each iteration and evaluation of the results allows improving the quality. Quality could be a function of the number of iterations and refinements a design undergoes. It is also useful reaching a right balance between client and design requirements.

Three iterations were done during the design process and are described below:

4.2.1 Iteration 1: First low-fidelity prototype

In our first iteration, we created a low-fidelity prototype of application. A low-fidelity prototype is a brief representation of the concept. Quick feedback and improvement of design are two of its benefits. Generally, it is not technology-oriented; thus, any substances for example papers, foams, wood or cardboard can be used to create them. We created our prototype on paper because it was easy to modify, portable and suitable for mobile application presentation.

The first iteration result is shown in Figure 5.

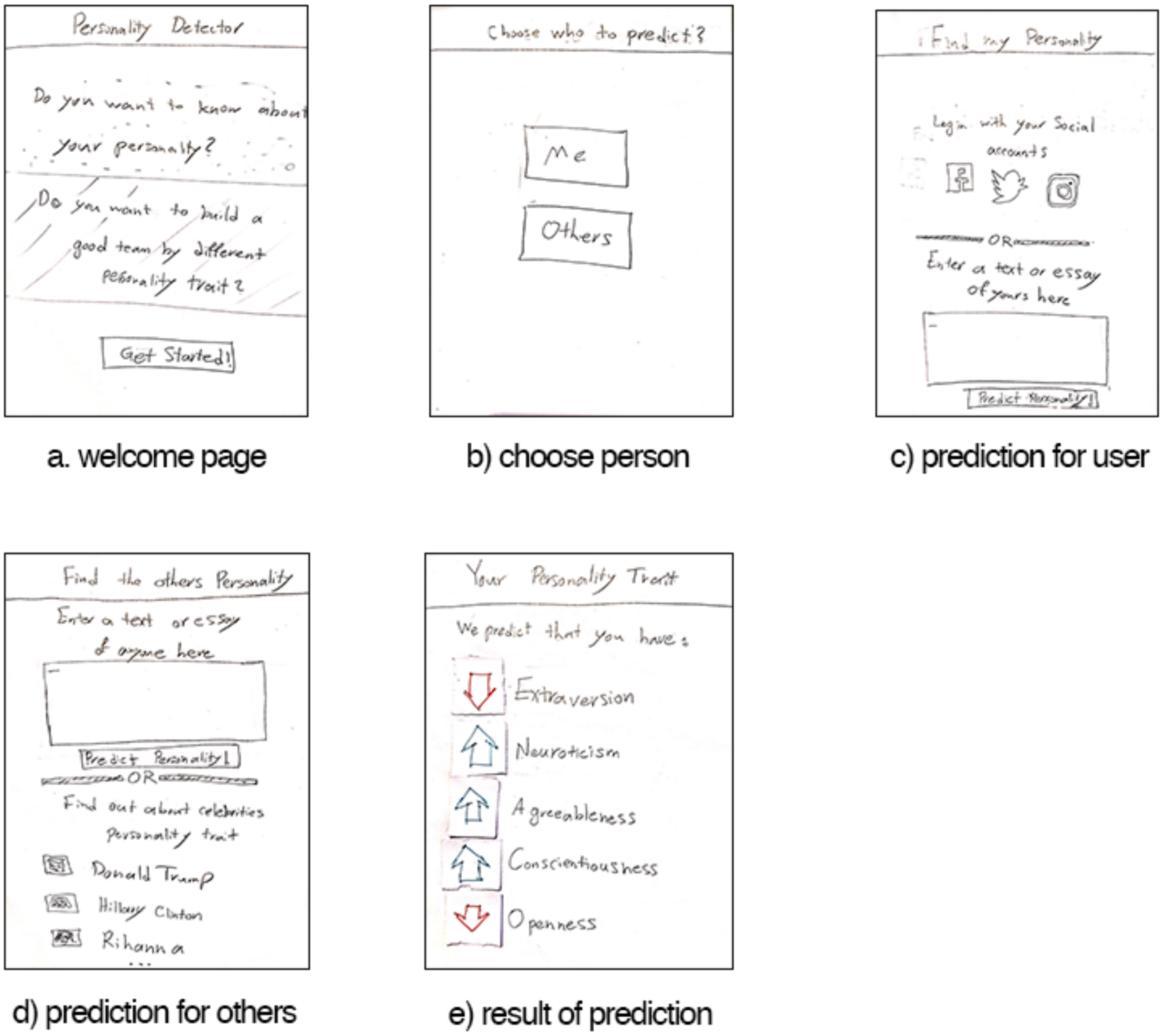


Figure 5 - First version of low-fidelity prototype

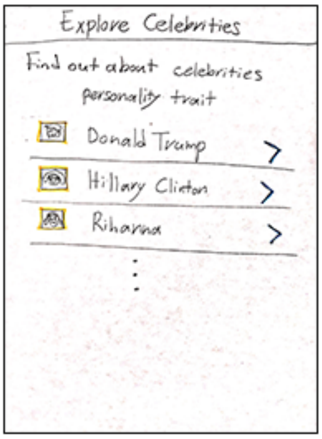
4.2.2Iteration 2: Second low-fidelity prototype

During this iteration, issues identified at the end of the first iteration (section 6.1.1) were fixed. As a result, a more accurate and clearer low-fidelity design was created, which can be considered as a medium-fidelity prototype. The pages of the created design are presented in Figure 6.

a) welcome

b) home

c) predict by text



d) celebrities

e) result of prediction

Figure 6 - Second version of lo-fi prototype

In the new design, all action buttons are marked by colours to make user interface elements more recognisable. We have the following pages after the second iteration:

Welcome page

This is the first page users can see after application installation. A short description of the application placed here to encourage users to use this application and tell them 'why' they should use the application. After clicking on the "Get Started" button, the user is directed to the homepage and then this page is not shown to the user again.

Homepage

The homepage is the start page of application which the user sees every time they open the application. To predict the personality, the application navigates the users to different sections of the application from here. They can log in to their social accounts such as Facebook, Twitter or Instagram from this page and get the result of their personality after successful login.

Predict by text page

By entering text on this page and clicking on the enter button, the user can get their personality details.

Celebrities page

Some celebrities whose social activities have already been examined by the algorithm used in the application's model are listed on this page. The user can click on each of them to see the result of personality detection.

Result of prediction page

The result of prediction for each personality trait is displayed on this page. Users are directed to this page from all other pages to see the positive result indicated by green upwards arrow and negative result displayed by red downwards arrow.

4.2.3Iteration 3: High-fidelity prototype

Before converting the developed medium-fidelity prototype to hi-fidelity version, we decided about colour schemes of the design.

The first consideration about colours was that its combination should enable the users to read easily and distinguish between background and text colour. We required some range of colours which make users trust the application such as blue and black. In addition, attractive colours, such as orange or yellow, could motivate users to start using the application after installation.

The chosen colour palette is shown in Figure 7.

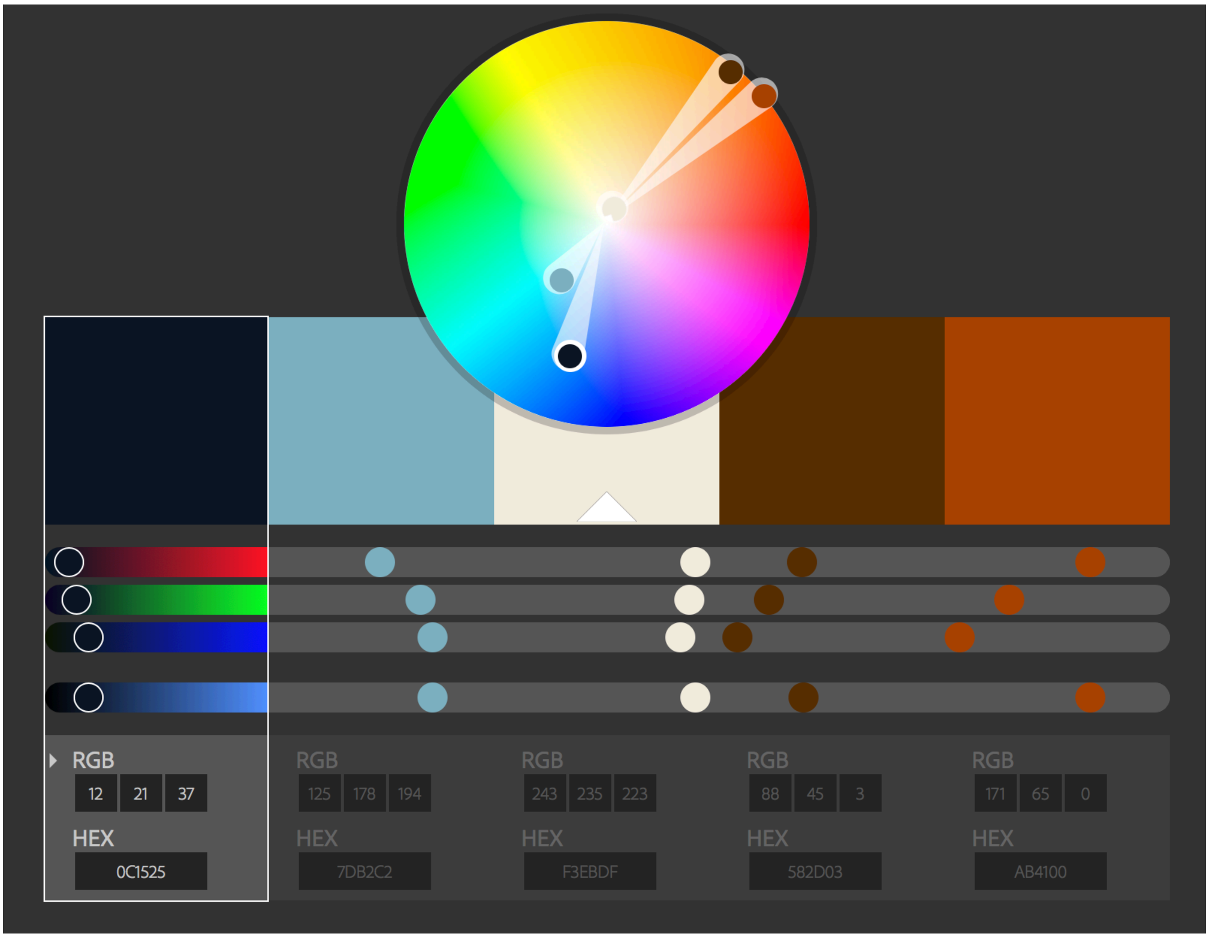


Figure 7 - Colour scheme of design

We named these five colours for our design as below:

- Dark Blue #0C1525
- Light Blue #7DB2C2
- Light Brown #F3EBDF
- Dark Brown #582D03
- Orange #AB4100

In this iteration we created a high-fidelity prototype because this design resembles a lot to the actual product. We focused more on details and functionalities become clearer in this iteration. Generally, the level of fidelity is related to how comprehensive is the design which enables us to examine usability and user experience questions. The user flow of the first version of hi-fidelity design is shown in Figure 8.

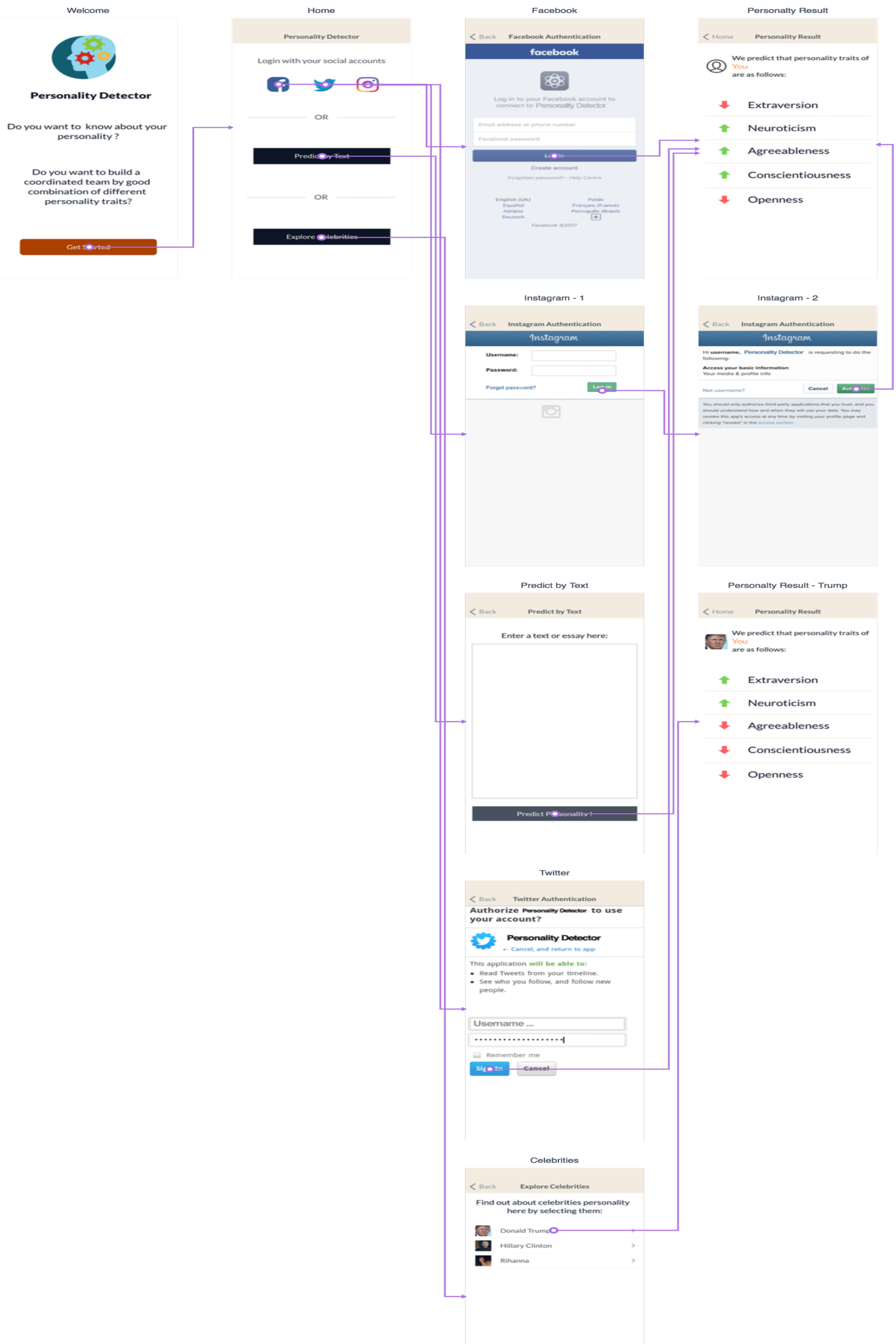


Figure 8 - User-flow of first version of high-fidelity prototype

This prototype is accessible at the following address:

https://marvelapp.com/5f1375b/

4.2.4Final Version

All problems identified at the last iteration were fixed, and an improved version provided as the final prototype. The prototype was developed during three iterations and was evaluated using three different evaluation methods. The numbers of identified issues at the last iteration were relatively low which makes it evident of eliminating all crucial usability problems and high quality of the design. Some essential pages of final design are pictured in Figure 9.

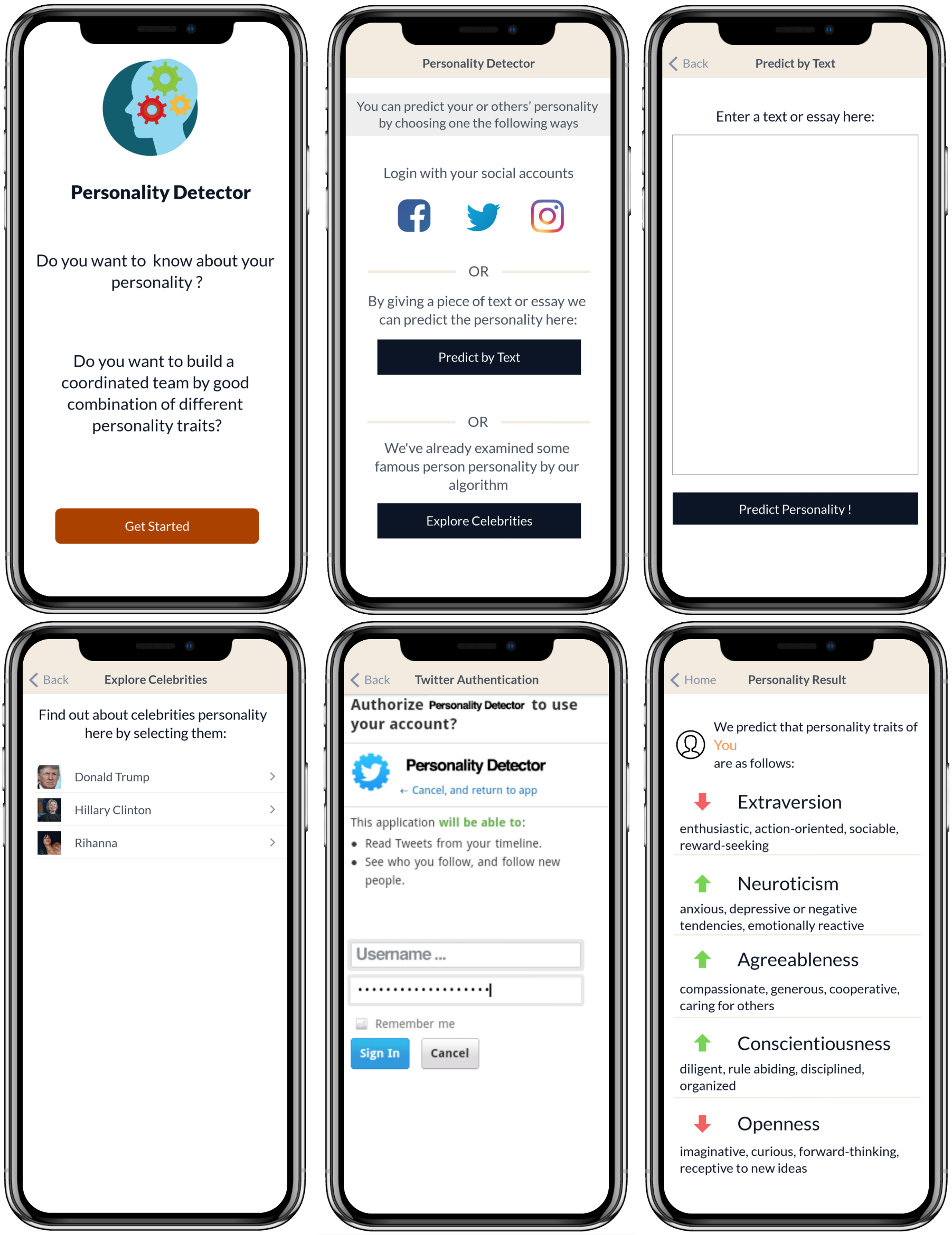


Figure 9 - Final version of prototype

The prototype is available at https://marvelapp.com/18gg7h6. The link can be opened on a web browser or a mobile phone. In the case of mobile phone, it can be saved on the home screen and then opened as a real application. Since it is the prototype, on some mobile phones an additional scroll may appear because of the different screen resolutions. In the real application, this problem will not occur because screen resolutions are considered during the mobile development process.

The prototype has limitations because it contains static pages without any programming of the functional logic. In addition, some user interface elements are not interactive in the prototype. For instance, it is not possible to enter the text in the fields presented in forms.

4.3 Mobile Application

There are two major parts on which the design of the mobile application is based on: input data and machine learning engine.

4.3.1 Input Data

To feed the machine learning (ML) classifiers, we need to collect the user data and send it to the model. The simplest way is to allow the user to enter the text manually. In this way, the user fills in the form in “Predict by Text” section with a piece of text or essay of themselves or the others’. There is no limitation in the length of the input text, however longer text would lead to a more accurate result.

The other method is to use the social activities of the user. To perform this, the user should first be authorised to access their social accounts. After a successful sign-in, the application grants access to read some posts, blogs, or tweets from the user’s social channel. Maximum 100 samples of user’s posts or tweets are collected as a chain of data. The posts with only images are ignored at this level.

In both input methods, the text is pre-processed by the application to make it suitable and compatible with ML model's input form.

4.3.2Machine Learning Engine

The machine learning models use the input data prepared for them in the previous stage described above, and then the classification result for each personality trait is returned to the application. The returned result is displayed to the user on the designated page of the user interface. The trained model could be implemented in either a mobile device or an external server to serve the application.

Nowadays, we can place the trained model inside the application for both iOS and Android operating systems. One of the advantages of this approach is using offline data, which doesn't require sending user data to any third-party APIs through the network and it is examined in the embedded model on the user's phone. Therefore, it keeps the user's privacy safe. However, it should be considered that it can increase the app size significantly.

Another approach is to implement the model on an external server and make a proper API such as RESTful one to respond to the requests. The application could request for prediction by the collected data and receive the prediction result. Since the networking level of the mobile application is used in this way and it needs to connect to the Internet, the prediction process should be performed online, and this is one of the main limitations of this approach. People are concerned about their privacy more these days, so they prefer keeping data on their devices instead of sending it over the network and analysing it on external computers. On the contrary, we can execute heavy machine learning algorithms such as deep models using a server-side approach. However, heavy processing is not that much operational by using the mobile device itself, because firstly it has limited resources of processing units and memory and secondly, it can consume much energy which is not acceptable for mobile devices' weak batteries. Another disadvantage of internal ML model is that it supports a limited type of machine learning algorithms at time of writing this report, for instance, reinforcement learning is supported neither in iOS nor Android platform. One cannot have continuous learning procedure by receiving new data from the user. Moreover, if it is observed that the machine learning models are not accurate enough, one is not able to refine them when they are implemented inside the device unless we update the application.

There are some serverless cloud-based machine learning approaches that can be used instead of implementing the ML server-side such as Google cloud machine learning engine and Amazon ML services. We do not have to maintain the infrastructures in this method and the price of the service is paid according to the number of requests.

5 Implementation

Based on the design, there are some implementation notes of the application in algorithm and model, interface and mobile application. Each of them is described in detail in the following chapters.

It is to be noted that the project's code is available at the following web address, and it is free to use for academic purposes and could be set up by the instructions described in the repository: https://github.com/novinfard/profiler-sentiment-analysis

5.1 Algorithm and Model

The algorithm and models are implemented in python language (2.7 version). The reason for this choice is that python is highly used for machine learning algorithms; thus, it is mature enough for stable developments. We had first used Natural Language Toolkit (https://www.nltk.org) for implementation of Naive Bayes and Support Vector Machines classifiers. However, it was not efficient in terms of n-grams tokenisation and the process was too slow (more than 2 hours for learning each model). In addition, it was not compatible with neither cloud-based machine learning models nor smart phone's in-device machine learning SDKs such as Android and iOS. As a result, we changed our libraries to Scikit-learn [18]. Moreover, this python library has some well-designed tools for cross-validation that we used for the evaluation of our trained models.

5.2 Interface

The implementation of interface started with creating a low-fidelity prototype. We just used papers and markers to create this paper-based interface.

Before starting work on a high-fidelity prototype, we should have created our colour scheme. We used “Adobe Colour CC” (https://color.adobe.com), an online tool for creating our colour set for this purpose. Although it has a complex interface, it offers professional toolset for colour mixing.

In the next stage, we used an online tool named “MarvelApp” (https://marvelapp.com) for creating a high-fidelity prototype. MarvelApp provides core user interface tools and functionality needed for creating a suitable prototype, wireframes or mock-ups. The final implementation of the user interface is accessible at the following link:

First hi-fidelity version (iteration 3) https://marvelapp.com/5f1375b

Second hi-fidelity version (final version) https://marvelapp.com/18gg77h6/

5.3 Mobile Application

To implement the mobile application, we chose the iOS platform. We could implement the designed solution on Android platform as well, but because of the time limitation, we decided to put it for further works. We also defined our target devices as iPhone and set the minimum version of iOS as 10 which covers 95 percent of this platform devices at the time of writing the report as shown in Figure 10.

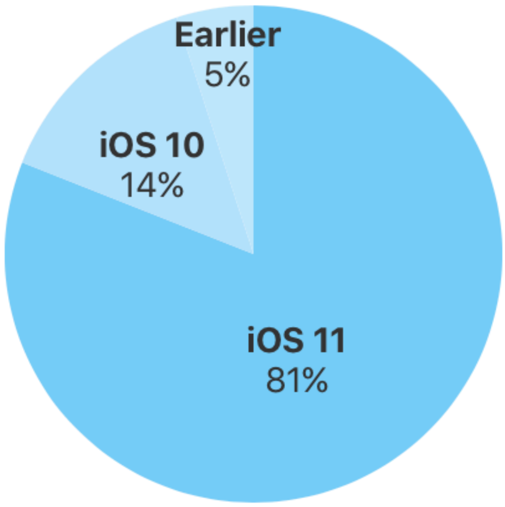


Figure 10 - iOS distribution measured by App Store in May 2018 [19]

Both logic and presentation layers of the applications are developed inside Xcode IDE (the IDE designed by Apple for developing iOS applications).

We make the solution by working on the interface, input data and machine learning sections as described in the following sections.

5.3.1 Interface

The interface of the mobile application is designed according to the final prototype (section 4.2.4). To build the interface we used the apple Storyboards which gives us not only the required user interface elements, but it also demonstrates the relationship between different pages of the application.

We applied the Auto Layout techniques for all objects in the interface to preserve the suitable location and size of UI elements in different resolutions and device sizes. Auto Layout's main idea which gives this flexibility lies in the fact that each object defines a controlling constraint based on its parent view.

5.3.2Input Data

The input data to use in the ML engine in “predict by text” is just a simple input text for the user. However, the process for social accounts login is different. Because of the shortage of time, we implemented the Twitter login, but the process is almost similar for other social networks.

To provide Twitter login, we had first applied for developer account by filling the application form about the reasons that we need to access the Tweets of the user and how we would use this data. After the approval of the developer account, we created the tweet collection functionality on the application by first authorising the user to Twitter using OAuth [20] method. OAuth or Open Authorisation is standard token-based authentication method on the internet. After login, the last 100 tweets of the user are collected via the Twitter API. The data is pre-processed and passed to the ML engine.

5.3.3Machine Learning Engine

For the implementation of ML engine, we had first tried to use our trained model (described in 5.1 and 4.3.2) inside the mobile device.

Apple in WWDC 2017 introduced CoreML framework for machine learning usage in iOS applications [21]. This framework brings mechanism to use pre-trained models in mobile applications.

To convert our written model in python to CoreML model we used CoreMLTools which supports a wide range of machine learning frameworks such as Keras, Caffe, Scikit-learn [18], Libsvm, and XGBoost. However, when we tried to adapt our model into CoreML framework, we concluded that it only supports some basic models in these frameworks. As a result, we had to convert our model by ourselves to CoreML which took extensive time and effort. Therefore, because of the limited amount of time for implementation we decided to use external-server approach for deploying our ML engine there. The Google Cloud ML engine and local server are used for deployment. After that, to predict the personality traits, we send the input data to an endpoint connected to ML engine. In the end, the responded result from ML engine on the cloud is displayed to the user.

6 Testing and Evaluation

For the evaluation part, we focused on two main regions of application: interface design and algorithm. Here we explain each testing part in detail.

6.1 Interface Design evaluation

We used iterative process by binding design and evaluation together at each stage of interactive interface design. The evaluation of each stage described below:

6.1.1 Evaluation of Iteration 1: Expert evaluation

At this stage, a quick and dirty expert evaluation is done by usability experts. This method of evaluation is built on past experience and can suggest potential problems. It shows the severity of design issues which helps that critical issues to be solved first. In addition, as there are no groups of users recruited to do evaluation and just an expert performs the evaluation, it is quicker and cheaper compared with user evaluation methods. However, it may not necessarily point out the real problems of users which should be considered.

To address some of the issues discovered in this stage, following actions are performed.

Replace “choose person” with new “home” page

The “choose person” page gives two options for users, the user itself or others. However, it refers to the same functionality in the next page which predicts the personality by entering the text of the user or others. This causes redundancy in the interface. It also makes confusion for users to which page they have been navigated because of the similarity in their look between “prediction for users” and “prediction for others” pages. On the other hand, to log in with social accounts or view prediction for celebrities' users have to pass two pages. It can be simplified if users can choose these functionalities from the first page they entered the application. Therefore, we changed this page to a new page named "home" which is function-oriented with three following buttons:

- Login with one's social accounts
- Predict by text
- Explore celebrities

Restructuring and renaming prediction for users and others' pages

After changing the structure of home page based on the functionality, we renamed the "Prediction for the user" page with "predict by text" and removed the social accounts login section which is available in "home" page now. In addition, we renamed the title "Prediction for others" page with "celebrities" after removing the textbox from it.

Distinguishable elements in user interface

Different elements in user interface including buttons and labels were not distinguishable in the low-fidelity prototype. We added colours to buttons in all pages to make it clear where users touch would lead to action (navigating to other pages).

6.1.2Evaluation of Iteration 2: User evaluation

In this iteration, user evaluation is performed on the design. User-based evaluation involves getting real users to try to use the current design for observing what problems they have and where they may be confused. User-based evaluations must be conducted ethically. Ethics forms were signed by us before evaluation. Before conducting the experiment, all participants completed informed consent.

To evaluate the usability of the current low-fidelity design and considering the limitation of time for doing the project, the application was shown to two users:

a) 25-year-old female, electrical engineer, uses cell phone, computer expert, knows how to use the Internet and smartphone;

b) 51-year-old male, manager, uses cell phone, familiar with how to work with computer and mobile applications.

During the evaluation, the prototype was shown to the users separately and different tasks were given. Then the questions about the pages were asked while performing tasks. The following tasks and questions were asked:

Task 1:

Open the application, go to home page.

Related questions:

a) Is the aim of the application clear?
b) Does welcome page make you curious to work with the application?
c) What do you expect if you press the “explore celebrities” button?
d) What parts of the home page can you tap on?

Task 2:

Open the application, go to home page, log in to one of your social accounts, get the result

Related questions:

a) Is it understandable why you should sign in to your social account?
b) What do you understand from the result page?
c) What do upwards and downwards arrows mean?

Overall, users concluded that the application is easy to use: pages do not contain excessive details, the number of possible actions is limited and, therefore, it is straightforward to understand what to do.

However, some critical design issues were found by users:

Back buttons

There is no back button in the interface design, so users do not know how to come back to previous pages. This problem makes users unable to start new personality detection. Thus, in all pages except welcome page, back button was added.

Explanation of personality traits

Using the name of the personality traits such as neuroticism and agreeableness do not convey what they mean. These labels were vague and have a different meaning to each user. Therefore, we added some simple adjectives to describe each personality trait better at the bottom of each trait name.

Display the person's name (and profile picture) in the result page

A few customisations were needed on the result page. Name of the person whose result is displayed, and a probable profile picture were added to the result page.

6.1.3Evaluation of Iteration 3: Rigorous expert heuristic evaluation

On this iteration, we applied the rigorous expert heuristic evaluation. This method involves several usability experts who evaluate the design against 10 criteria called usability heuristics. These criteria are defined by Nielsen and are recognised usability principles for interaction design [22] and described briefly and simply as follows:

1) Visibility of system:

   Getting appropriate feedback

2) Match between system and the real world:

   Making information appear in a user's natural and logical order

3) User control and freedom:

   Users often choose system functions by mistake; undo and redo are needed

4) Consistency and standards:

   Choosing the same words for the same actions and situations of system

5) Error prevention:

   Not letting the users make mistakes rather than showing error messages

6) Recognition rather than recall:

   By applying visibility, the user does not have to remember information from one part to another

7) Flexibility and efficiency of use:

   By giving accelerators to experienced users, making it suitable for all users

8) Aesthetic and minimalist design:

   Just needed and relevant options should be visible to users

9) Help users recognise, diagnose, and recover from errors:

   Expressing error messages in plain language (no codes) with a solution suggested

10) Help and documentation:

    Providing short, easy-to-search and focused documents when needed

Nielson [22] believed that the best evaluation results could be achieved by using 3 to 5 usability experts. Hence, the three usability experts participated in the evaluation. They independently used the interactive prototype, assessed it according to 10 heuristics and ranked the severity of problems (1 to 4 for cosmetic, minor, moderate, and severe issues respectively). After that, usability experts compared the list of identified issues with each other and agreed on a final list. The issues that were diagnosed are described below.

Ambiguity in homepage's goal

Level: moderate

Criteria: help and documentation

In the homepage, there is no description, metaphor or sign that why they should log in or what celebrities' section is about. It can be solved by adding a descriptive text at the top of the page (which can cover all sections) such as:

"You can predict your or others' personality by choosing one the following ways:"

Different button colours

Level: cosmetic

Criteria: consistency and standards

Buttons on "home" and "predict by text" have different colours, although they have the same meaning to the user. They should use the dark blue colour of our colour scheme.

6.2 Evaluation of algorithm and model

In machine learning algorithms, we separate the data into training and testing set or use another dataset for the test. Although this separation amount could be of any ratio, 80 percent training and 20 percent for testing is a good starting point. However, different ratios and number of records could cause overfitting or underfitting issue.

The cross-validation and testing are used to reduce the risk of overfitting and under-fitting when one trains a model in data science applications. The overfitting and under-fitting are two common issues in statistical models. Therefore, if we do not have a good balance of training and testing portion, the numbers dataset records would not be enough, or the dataset quality associated to different classes would not be good enough, therefore, we will suffer from overfitting or under-fitting.

If one fits their model too much to their train set it is probable that your model overfits and do not predict new samples correctly. On another side, if the data amount is not enough, it is possible to fell into under-fitting and one's model could not separate different classes correctly.

K-fold cross-validation is one the cross-validation techniques by which we randomly partition the dataset into k equal amount of data. One of the k samples is used for validating the data and testing purpose and k-1 remaining partitions fed into the model's training data. The process continues until each fold is used once for validation, which means we test the data k times in total. When we have all the k models, we take an average of the error terms which means the error divided by k. Buhagiar and Juan [23] demonstrated the k-fold validation procedure for k=10 in Figure 11:

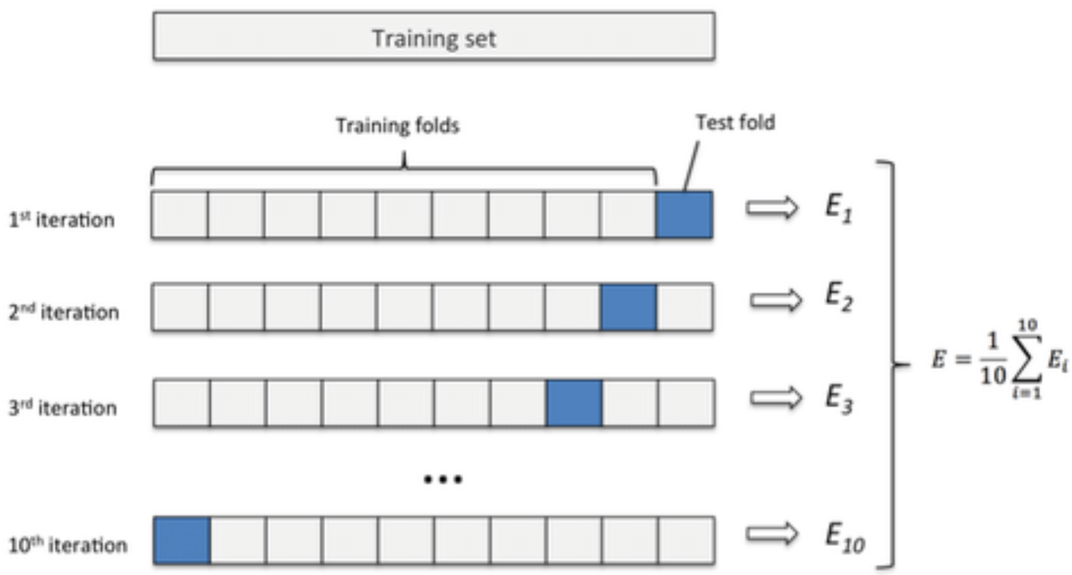


Figure 11 – A demonstration of 10-fold cross-validation process [23]

In this project, we used a famous credible dataset for personality traits [14] which enables us to compare it with state-of-the-art papers. With using 10-fold cross-validation as a measurement metric for our models, the following results are obtained:

	EXT	NEU	AGR	CON	OPN
SVM	**58.33**	56.79	54.20	55.86	62.95
Naive Bayes	57.24	56.71	54.08	55.13	**63.11**
State-of-the-art model 1 (using the convolutional neural network) [24]	55.07	**59.38**	55.08	55.14	60.51
State-of-the-art model 2 [25]	55.45	58.33	**56.03**	**56.73**	60.68

Table 1 - Model accuracy with different classifiers and comparison with other works

In the state-of-the-art model 1, as they used different settings, the MLP classifier with 2,3,4 convolutional filter has been selected from their work.

As can be seen from the Table 1, our SVM model gets the best result for extravert (EXT) trait among other classifiers by 58.33 percent accuracy in its prediction comparing with state-of-the-art model 1 and 2 with 55.07 and 55.45 percent accuracy measure respectively. For neuroticism (NEU) personality, state-of-the-art model 1 performs better than our classifier with almost 3 percent more accurate predictions. State-of-the-art model 2 precited slightly better than other three classifiers in agreeableness (AGR) and conscientiousness (CON) categories by approximately 1 percent. Both out classifiers including Naive Bayes and SVM outperform in Openness (OPN) personality trait by 63.11 and 62.95 percent respectively.

6.2.1Test on specific cases

On the better evaluation of our algorithm, we tested our model against two specific cases: Donald Trump and Hillary Clinton. We first explained what Five Big Factors means to some people and then asked them to tell us about the personality trait of Trump and Clinton, in addition to some blog posts investigation [26] [27] [28] [29]. It is noticeable that each candidate's real personality may be different from the personality that they reflect in their social accounts, especially because they are politicians and sometimes a team of writers and psychologists polish their posts before publishing them in their social accounts. The following result has been achieved:

Trump personality traits:

1) Extroversion (EXT): Positive

He is outgoing, talks dominantly without fear and tries to broaden his social connections.

2) Neuroticism (NEU): Positive

Although he seems fearless in his speeches, he is emotionally unstable and reacts quickly.

3) Agreeableness (AGR): Negative

He shows less empathy and kindness towards others. He also withdrew some international agreements such as Paris Climate Agreement in June 2017 which implies he may not consider others or next generations even in environmental world-range issues.

4) Conscientiousness (CON): Negative

He does not seem to be aware of the needs of others. He believes that the United States can perform better without getting help from other nationalities and countries. One of his famous quotes against H1-b visa (a United States visa type which allows employers to employ foreign nationals as skilled workers) is "hire American".

5) Openness (OPN): Negative

He presents his attitudes and viewpoints directly in a way that he is unwilling to express an apology. Although this is a sign of openness, these speeches from him convince us to label his personality trait for openness as negative:

“build the wall”

“total and complete shutdown of Muslims entering the United States”

Another sign of low openness in him is that sometimes he shows a strong commitment to traditional ideas, without considering changes happened to the world such as in his famous slogan:

“make America great again”

Hillary Clinton personality traits:

1) Extroversion (EXT): Negative

She prefers to give fewer public speeches. In addition, she strictly tends to make decisions rationally rather than emotionally which reduces her close social interactions and connections.

2) Neuroticism (NEU): Negative

She seems to be adjusted and calm under pressure. In addition to her firmness when she faces critics, she hesitates to act when she is nervous which is a sign of low neuroticism in her personality.

3) Agreeableness (AGR): Negative

In her political competition, she occasionally considered her fellow Democrats and supported them which could imply her agreeableness. However, there are cases she defended her opinions against all opponents in both political factions which persuades us to label this personality trait as negative (though it is not that much high or low)

4) Conscientiousness (CON): Positive

Her support for unprivileged people during her life is a sign that she is aware of others and cares about them. Another feature of her personality is that she is highly diligent and perfectionistic. Therefore, she considers the details and is self-organised.

5) Openness (OPN): Positive

She is open to new experiences as she has taken charge of positions during her career path which was unusual for women. Another reason for this selection is related to her resistance against traditions and conventions as a sign of openness.

In the next stage, our model had been tested on around 6444 tweets of these American presidential candidates in 2016 (3218 tweets for Trump and 3226 tweets for Clinton) to figure out the predicted personality trait of them. The result achieved by models for Trump and Clinton is shown in Table 2 and 3 respectively.

	Our Assumption	SVM model	NB model	SVM model score	NB model score
EXT	positive	negative	**positive**	39.18	63.83
NEU	positive	negative	negative	0.71	34.93
AGR	negative	**negative**	**negative**	21.78	38.32
CON	negative	positive	**negative**	97.07	42.70
OPN	negative	**negative**	**negative**	0.83	29.21

Table 2 - The model's results for Donald Trump

	Our Assumption	SVM model	NB model	SVM model score	NB model score
EXT	negative	**negative**	positive	43.83	73.28
NEU	negative	**negative**	**negative**	0.96	27.81
AGR	negative	**negative**	**negative**	21.95	41.82
CON	positive	**positive**	negative	97.12	33.35
OPN	positive	negative	negative	0.34	25.26

Table 3 - The model's results for Hillary Clinton

As can be seen from Table 2, Naive Bayes (NB) algorithm predicts closely to our assumptions for Trump case with the same prediction for 4 personality traits out of 5. However, the SVM classifier predicts 40% of cases equivalent to the assumptions. For Hillary Clinton case, the predictions performed by SVM is more accurate than NB classifier according to our assumptions, as shown in Table 3. We obtained the same result as expected for all Clinton personalities with SVM classifier except the openness (OPN).

7 Conclusions

In this work, we create a mobile application for predicting the personality traits of people by applying two different classifiers as our machine learning engine. These models have suitable performance in terms of predicting accuracy comparing with two other current state-of-the-art models.

The architecture of the mobile application is analysed in terms of its data input and machine learning model components. We discussed the different possibilities of deployment of our models inside the mobile application, considering each design's advantages and disadvantages.

The interface design is another critical aspect of the mobile application design similar to any other software design cycle. Therefore, we created our interface design according to interactive interface design principles in three iterations. In each iteration, we applied different evaluation methods such as quick and dirty expert evaluation, user evaluation, and expert heuristic evaluation. Thus, the final interface seems to have the minimum error.

8 Further work

In the future, we plan to improve the project in different aspects including the model, the mobile application, and the interface.

In current work, the model is optimised for the English language which can be expanded to other languages by recognising the grammar, structure and word delimiter of different languages. In addition, deep learning paradigms in machine learning could be applied to see if they can achieve better accuracy considering all the conditions for mobile applications such as accepted response time and efficient battery consumption on in-device models. Moreover, predicting emotions as well as personality in our models could give better insight to users about themselves. Apart from this, emotion detection may help us to have better personality results as it seems there are some connections between them.

We intend to implement some of our ML models inside the mobile applications and check the feasibility and efficiency compared to the current server-side prediction deployment.

Because we need the same machine learning models and almost the same user interface for Android devices, developing the same application for the Android platform could expand its users which is one of the future plans.

From the user interface perspective, we want to evaluate our design with more users by which we can reduce more user interface and even user experience flaws. In addition, we can design a new interface based on tablet devices to make the application available to them as well.

9 References

1. Digman JM. Personality structure: Emergence of the five-factor model. Annu Rev Psychol. 1990;41(1):417–40.
2. McCrae RR, Costa Jr PT. Personality trait structure as a human universal. Am Psychol. 1997;52(5):509.
3. Goldberg LR. The structure of phenotypic personality traits. Am Psychol. 1993;48(1):26.
4. Tsytsarau M, Palpanas T. Survey on mining subjective data on the web. Data Min Knowl Discov. 2012;24(3):478–514.
5. Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on human language technology and empirical methods in natural language processing. 2005. p. 347–54.
6. Liu B. Sentiment analysis and opinion mining. Synth Lect Hum Lang Technol. 2012;5(1):1–167.
7. Hagenau M, Liebmann M, Neumann D. Automated news reading: Stock price prediction based on financial news using context-capturing features. Decis Support Syst. 2013;55(3):685–97.
8. Balahur A, Steinberger R, Kabadjov M, Zavarella V, Van Der Goot E, Halkia M, et al. Sentiment analysis in the news. arXiv Prepr arXiv13096202. 2013;
9. Tandera T, Suhartono D, Wongso R, Prasetio YL, others. Personality prediction system from facebook users. Procedia Comput Sci. 2017;116:604–11.
10. Colneriĉ N, Demsar J. Emotion Recognition on Twitter: Comparative Study

and Training a Unison Model. IEEE Trans Affect Comput. 2018;

11. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey. Ain Shams Eng J. 2014;5(4):1093–113.

12. Wikipedia. Big Five personality traits [Internet]. 2018 [cited 2018 Jul 25]. Available from: https://en.wikipedia.org/wiki/Big_Five_personality_traits

13. Cooper A, others. The inmates are running the asylum:[Why high-tech products drive us crazy and how to restore the sanity](Vol. 261). Sams Indianap. 1999;

14. Pennebaker JW, King LA. Linguistic styles: Language use as an individual difference. J Pers Soc Psychol. 1999;77(6):1296.

15. Wikipedia. tf–idf [Internet]. 2018 [cited 2018 Aug 5]. Available from: https://en.wikipedia.org/wiki/Tf–idf

16. Wikipedia. Support vector machine [Internet]. 2018 [cited 2018 Aug 7]. Available from: https://en.wikipedia.org/wiki/Support_vector_machine

17. Houde S, Hill C. What do prototypes prototype? In: Handbook of Human-Computer Interaction (Second Edition). Elsevier; 1997. p. 367–81.

18. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res. 2011;12(Oct):2825–30.

19. Apple. App Store - Support [Internet]. 2018 [cited 2018 Aug 5]. Available from: https://developer.apple.com/support/app-store/

20. Hardt D. The OAuth 2.0 authorization framework. 2012.

21. Apple. Core ML [Internet]. 2017 [cited 2018 Jul 25]. Available from: https://developer.apple.com/documentation/coreml

22. Nielsen J. 10 usability heuristics for user interface design. Nielsen Norman Gr.

1995;1(1).

23. Buhagiar J. Automatic Segmentation of Indoor and Outdoor scenes from Visual Lifelogging. University of Malta; 2017.
24. Majumder N, Poria S, Gelbukh A, Cambria E. Deep learning-based document modeling for personality detection from text. IEEE Intell Syst. 2017;32(2):74–9.
25. Mohammad SM, Kiritchenko S. Using hashtags to capture fine emotion categories from tweets. Comput Intell. 2015;31(2):301–26.
26. Owens M. The One Personality Trait That Won Donald Trump the Presidency [Internet]. 2016 [cited 2018 Jul 26]. Available from: https://www.truity.com/blog/one-personality-trait-won-donald-trump-presidency
27. Ramey A, Klingler J, Hollibaugh G. What Trump and Clinton's personality traits tell us about how they might govern as president [Internet]. 2016 [cited 2018 Jul 26]. Available from: http://blogs.lse.ac.uk/usappblog/2016/11/02/what-trump-and-clintons-personality-traits-tell-us-about-how-they-might-govern-as-president/
28. Calico J. The ocean of trump's & clinton's personalities [Internet]. 2016 [cited 2018 Jul 26]. Available from: https://thepsyoflifeblog.com/2016/09/05/the-ocean-of-trumps-clintons-personalities/
29. Sherman R. The Personality of Hillary Clinton [Internet]. 2016 [cited 2018 Aug 12]. Available from: https://www.psychologytoday.com/us/blog/the-situation-lab/201610/the-personality-hillary-clinton

10 Appendix

10.1 Ethic form for user evaluation

Ethics Sign-off Form
Department of Computer Science

This form applies only studies that involve the use of other people ('participants') to collect information. For example, to carry out simple evaluations of a system or a system design, or for getting information about how a system might be used.
If no other people have been involved in the collection of information, then you do not need to complete this form.

By signing this form you confirm that your study satisfies each of the following ethical principles:

1. Participants are not exposed to any risks greater than those encountered in their normal working life.
2. Experimental materials are paper based or comprise software running on standard hardware.
3. All participants give explicit consent to take part.
4. No incentives are offered to participants.
5. No information about the evaluation or materials is intentionally withheld from the participants.
6. No participant is under the age of 16
7. No participant has an impairment that might limit their understanding or communication.
8. I am not in a position of authority or influence over any of the participants.
9. All participants are informed that they can withdraw from the study at any time.
10. All participants are informed of my contact details.
11. Participants are debriefed about the aims of the study and given the opportunity to ask questions.
12. All the data collected from participants is stored in an anonymous form.

COURSE / PROJECT TITLE: Profiler Application using Sentiment Analysis

NAME: Soheil Novinfard

STUDENT NUMBER: 170801896

SIGNATURE: Novinfard

DATE: 5/8/18

1

10.2 User-based user interface evaluation forms

User-based Evaluation

Iteration Number: 2

Tester: 25-year-old female / electrical engineer

Expert involved during test: Soheil Norinford

Task 1:

Open the application, go to home page.

a) Is the aim of the application clear?

Yes, to some extent

b) Is welcome page make you curious to work with application?

Yes, I want to know what is my Personality

c) What do you expect if you press the "explore celebrities" button?

To see some information about them

d) What parts of the home page you can tap on?

Social buttons, the two other buttons (predict by text and celebrities)

Task 2:

Open the application, go to home page, login to one of your social accounts, get the result

a) Is it understandable why you should sign in to your social account?

yes

b) What do you understand from result page?

There are some word that maybe related to my personality but I'm not sure.

c) What does upwards and downwards arrow mean?

upwards mean more and downwards is the opposite

other comments: If the result page could display my name, it would make me sure that I logged-in successfully and the result is related to me.

7/8/18

User-based Evaluation

Iteration Number: 2

Tester: 51-year-old male/manager

Expert involved during test: Soheil Novinfard

Task 1:

Open the application, go to home page.

a) Is the aim of the application clear?

Yes

b) Is welcome page make you curious to work with application?

Yes, but I should be sure it's safe

c) What do you expect if you press the "explore celebrities" button?

Some personality analysis of famous people

d) What parts of the home page you can tap on?

All buttons and icons

Task 2:

Open the application, go to home page, login to one of your social accounts, get the result

a) Is it understandable why you should sign in to your social account?

Yes

b) What do you understand from result page?

I don't understand anything from this page.

c) What does upwards and downwards arrow mean?

more or less

other comments: how can I come back to home page to see the celebrities personality analysis?

7/8/18

10.3 Rigorous expert user interface evaluation forms

Rigorous Expert Heuristic Evaluation

Iteration Number: 3

Expert: Soheil Novinfard

1) Visibility of system

Rating (1 to 10): 8

Founded issues with severity:

[minor]

Celebrities button seems a bit vague in home page

2) Match between system and the real world

Rating (1 to 10): 9

Founded issues with severity:

3) User control and freedom

Rating (1 to 10): 10

Founded issues with severity:

4) Consistency and standards

Rating (1 to 10): 7

Founded issues with severity:

5) Error prevention

Rating (1 to 10): 10

Founded issues with severity:

6) Recognition rather than recall

Rating (1 to 10): 9

Founded issues with severity:

7) Flexibility and efficiency of use

Rating (1 to 10): 7

Founded issues with severity:

8) Aesthetic and minimalist design

Rating (1 to 10): 8

Founded issues with severity:

9) Help users recognise, diagnose, and recover from errors

Rating (1 to 10): 8

Founded issues with severity:

10) Help and documentation

Rating (1 to 10): 9

Founded issues with severity:

Signature and date:

10/8/18

Novinfard

Figure 12 - Expert evaluation 1 for user interface

Rigorous Expert Heuristic Evaluation

Iteration Number: 3

Expert: Manuel Campos

1) Visibility of system

Rating (1 to 10): 9

Founded issues with severity:

2) Match between system and the real world

Rating (1 to 10): 8

Founded issues with severity:

3) User control and freedom

Rating (1 to 10): 9

Founded issues with severity:

4) Consistency and standards

Rating (1 to 10): 8

Founded issues with severity:

5) Error prevention

Rating (1 to 10): 9

Founded issues with severity:

6) Recognition rather than recall

Rating (1 to 10): 9

Founded issues with severity:

7) Flexibility and efficiency of use

Rating (1 to 10): 8

Founded issues with severity:

8) Aesthetic and minimalist design

Rating (1 to 10): 8

Founded issues with severity:

9) Help users recognise, diagnose, and recover from errors

Rating (1 to 10): 9

Founded issues with severity:

10) Help and documentation

Rating (1 to 10): 6

Founded issues with severity: minor
the reason for signing to social accounts is not clear and for 2 other buttons there should be some description to acknowledge the user of the purpose of them

Signature and date: 10/8/2018
Campos

Figure 13 - Expert evaluation 2 for user interface

Rigorous Expert Heuristic Evaluation

Iteration Number: 3

Expert: Deni Setiawan.

1) Visibility of system

Rating (1 to 10): 8

Founded issues with severity:

2) Match between system and the real world

Rating (1 to 10): 9

Founded issues with severity:

3) User control and freedom

Rating (1 to 10): 7

Founded issues with severity:

4) Consistency and standards

Rating (1 to 10): ~~8~~ 5

Founded issues with severity:

Cosmetic level

buttons color are different in different pages in apps, but, they convey the same functionality.

5) Error prevention

Rating (1 to 10): 8

Founded issues with severity:

6) Recognition rather than recall

Rating (1 to 10): 7

Founded issues with severity:

7) Flexibility and efficiency of use

Rating (1 to 10): 8

Founded issues with severity:

8) Aesthetic and minimalist design

Rating (1 to 10): 8

Founded issues with severity:

9) Help users recognise, diagnose, and recover from errors

Rating (1 to 10): 9

Founded issues with severity:

10) Help and documentation

Rating (1 to 10): 7

Founded issues with severity:

Signature and date: 10/08/2018

Figure 14 - Expert evaluation 3 for user interface