DISSERTATION
ZUR ERLANGUNG DES AKADEMISCHEN GRADES
"DOCTOR RERUM NATURALIUM"
(DR. RER. NAT.)
IN DER WISSENSCHAFTSDISZIPLIN
DIGITAL ENGINEERING

# Social Media Based Personality Prediction Models


**By**

Raad Bin Tareaf

**Supervisor**

Prof. Dr. Christoph Meinel

*Faculty of Digital Engineering*
*University of Potsdam, Germany*

Hasso Plattner Institute
at University of Potsdam


March 29, 2022

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-54914
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-549142

**Eidesstattliche Erklärung**

Hiermit versichere ich, dass meine Arbeit „Social Media Based Personality Prediction Models“ („“) selbständig verfasst wurde und dass keine anderen Quellen und Hilfsmittel als die angegebenen benutzt wurden. Diese Aussage trifft auch für alle Implementierungen und Dokumentationen im Rahmen dieses Projektes zu.

Potsdam, den 29. März 2022,

Raad Bin Tareaf

**Abstract**

Individuals have an intrinsic need to express themselves to other humans within a given community by sharing their experiences, thoughts, actions, and opinions. As a means, they mostly prefer to use modern online social media platforms such as Twitter, Facebook, personal blogs, and Reddit. Users of these social networks interact by drafting their own statuses updates, publishing photos, and giving likes leaving a considerable amount of data behind them to be analyzed. Researchers recently started exploring the shared social media data to understand online users better and predict their Big five personality traits: agreeableness, conscientiousness, extraversion, neuroticism, and openness to experience.

This thesis intends to investigate the possible relationship between users' Big five personality traits and the published information on their social media profiles. Facebook public data such as linguistic status updates, meta-data of likes objects, profile pictures, emotions, or reactions records were adopted to address the proposed research questions. Several machine learning predictions models were constructed with various experiments to utilize the engineered features correlated with the Big 5 Personality traits. The final predictive performances improved the prediction accuracy compared to state-of-the-art approaches, and the models were evaluated based on established benchmarks in the domain. The research experiments were implemented while ethical and privacy points were concerned. Furthermore, the research aims to raise awareness about privacy between social media users and show what third parties can reveal about users' private traits from what they share and act on different social networking platforms.

In the second part of the thesis, the variation in personality development is studied within a cross-platform environment such as Facebook and Twitter platforms. The constructed personality profiles in these social platforms are compared to evaluate the effect of the used platforms on one user's personality development. Likewise, personality continuity and stability analysis are performed using two social media platforms samples. The implemented experiments are based on ten-year longitudinal samples aiming to understand users' long-term personality development and further unlock the potential of cooperation between psychologists and data scientists.

## Zusammenfassung

Menschen haben das Bedürfnis, sich anderen Menschen innerhalb einer bestimmten Gemeinschaft mitzuteilen, indem sie ihre Erfahrungen, Gedanken, Handlungen und Meinungen teilen. Zu diesem Zweck nutzen sie am liebsten moderne Online-Plattformen für soziale Medien wie Twitter, Facebook, persönliche Blogs und Reddit. Die Nutzer dieser sozialen Netzwerke interagieren, indem sie ihre eigenen Status-Updates verfassen, Fotos veröffentlichen und Likes vergeben und dabei eine beträchtliche Menge an Daten hinterlassen, die analysiert werden können. Forscher haben vor kurzem damit begonnen, die in den sozialen Medien geteilten Daten zu untersuchen, um die Online-Nutzer besser zu verstehen und ihre Big-Five-Persönlichkeitseigenschaften vorherzusagen: Verträglichkeit, Gewissenhaftigkeit, Extraversion, Neurotizismus und Offenheit für Erfahrungen.

In dieser Arbeit soll der mögliche Zusammenhang zwischen den Big Five Persönlichkeitsmerkmalen der Nutzer und den in ihren Social-Media-Profilen veröffentlichten Informationen untersucht werden. Öffentliche Facebook-Daten wie sprachliche Status-Updates, Metadaten von Likes, Profilbilder, Emotionen oder Reaktionsaufzeichnungen wurden zur Beantwortung der vorgeschlagenen Forschungsfragen herangezogen. Es wurden mehrere Modelle des maschinellen Lernens mit verschiedenen Experimenten erstellt, um die entwickelten Merkmale zu nutzen, die mit den Big 5 Persönlichkeitsmerkmalen korrelieren. Die endgültigen Vorhersageleistungen verbesserten die Vorhersagegenauigkeit im Vergleich zu modernsten Ansätzen, und die Modelle wurden auf der Grundlage etablierter Benchmarks in diesem Bereich bewertet. Die Forschungsexperimente wurden unter Berücksichtigung ethischer Aspekte und des Datenschutzes durchgeführt. Darüber hinaus zielt die Forschung darauf ab, das Bewusstsein für die Privatsphäre von Nutzern sozialer Medien zu schärfen und zu zeigen, was Dritte über die privaten Eigenschaften von Nutzern aus dem, was sie auf verschiedenen sozialen Netzwerkplattformen teilen und tun, herausfinden können.

Im zweiten Teil der Arbeit werden die Unterschiede in der Persönlichkeitsentwicklung in einer plattformübergreifenden Umgebung wie Facebook und Twitter untersucht. Die konstruierten Persönlichkeitsprofile in diesen sozialen Plattformen werden verglichen, um die Auswirkungen der verwendeten Plattformen auf die Persönlichkeitsentwicklung eines Nutzers zu bewerten. Ebenso werden Persönlichkeitskontinuität und -stabilität anhand von zwei Social Media Plattformen untersucht. Die durchgeführten Experimente basieren auf zehnjährigen Längsschnittstichproben mit dem Ziel, die langfristige Persönlichkeitsentwicklung der Nutzer zu verstehen und das Potenzial der Zusammenarbeit zwischen Psychologen und Datenwissenschaftlern weiter zu erschließen.

## Acknowledgement

First, I would like to express my honest gratitude and appreciation to Professor Dr. Christoph Meinel and his wonderful support and guidance to complete my Ph.D. journey. I am thankful for his valuable feedback, inspiration and guidance because the work would not have been possible without his direct support, trust and patience. I would extend sincere gratitude to my second supervisor Professor Dr. Felix Naumann for being an amazing role model and an infinite source of knowledge and inspiration which helped me tackle my research questions.

Throughout my journey at the Hasso Plattner Institute, I had the pleasure to engage and work with many gifted persons, whom I was pleased to consider more than work colleagues. Special thanks to our chair secretary Michaela Schmitz and Nicola Mertens for their administrative work.

I am thankful to my family, for all the motivation, inspiration and support they gave throughout my Ph.D studies. I have to express my gratefulness to my oldest brother, Khaled, who had aroused my obsession with computers and information technology when I was a kid at school, started by teaching me how to surf the internet using dial-up connections completed by instructing me on how to trade crypto assets in cyberspace.

I am also extremely thankful to my wife, Mais, for all the inspiration and motivation she provided me during my studies and for carrying our baby. She stood by me through all impatience. She gave me support and help, discussed abstract ideas and prevented several wrong turns in my plans. I am very thankful to my colleagues Eric Klieme, Ali Alhosseini, Muhammad Sukmana, Hanadi Traifeh and Kennedy Torkura and many others for their discussions, inputs and motivations.

# Contents

# List of Tables

# List of Figures

# 1 Introduction

## 1.1 Social Media Platforms and Online Personality

Web usage has been significantly expanded during the last decade. Users have begun to spend their online time on websites that allow anyone to write, update, and contribute to them. For this reason, many web technologies get developed by social experts to facilitate user experience and give users the ability to collaborate and contribute in an online fashion. These contributions are called user-generated content from various social media platforms. They are considered a rich source of information about individuals and the masses. Therefore, scientists collect and analyze these contents to provide new approaches for enterprises to understand their users and personalize services based on online interest and identity.

In the past few years, several types of social media platforms evolved. For example, micro-blogging platforms such as Twitter and social networking sites like Facebook (Meta) have become one of the most popular networking sites worldwide. Therefore, it is no surprise that these platforms play a huge role in individuals' social interactions. There are also other types of social media platforms such as Wiki-based knowledge sharing sites (e.g., Wikipedia), Social news sites and websites of news media (e.g., Huffington Post), Community media sites (e.g., YouTube, Flickr, Instagram), Social curation sites (e.g., Reddit, Pinterest), User reviews (e.g., Yelp, Amazon.com), Social Q and A sites (e.g., Quora, Yahoo Answers) and Location-based social networks (e.g., Foursquare). This thesis considers Facebook and Twitter platforms the primary data source to answer the proposed research questions.

These services and platforms, such as Facebook and Twitter, have different purposes and intentions in convincing and attracting online users to share their experiences, impressions, opinions, and thoughts. As Facebook was founded in 2004, the primary mission they follow is to empower individuals and bring societies closer [1]. Therefore, users utilize Facebook to stay connected with family members and friends, explore the world, and share what is valuable to them.

Based on Facebook reports of the first quarter in 2020 [2], Facebook is the third-most visited website outranked only by Google and YouTube. More than 3.21 billion people actively use Facebook, Instagram, WhatsApp, or Messenger each month compared to 1.66 billion for 2019. Twitter, on the other hand, was founded in 2006, and its mission is to provide everyone the ability and power to formulate and share ideas and information instantly without restrictions [3]. Based on SimilarWeb stats in October 2020 [4], Twitter is the 4th most-visited website in terms of traffic, with 6.1 billion visits.

In previous years, statistics reveal a massive increase in users' interest in utilizing online social platforms. It also showed that computational science attracted many researchers from different subdomains in academia to investigate and study the personality computation research areas. There is a shred of reliable empirical evidence supporting the validity of information collected from online profiles [5]. As these platforms provide users the ability to share their ideas and thoughts expressed in a natural environment, scientists found that these samples, which users themselves originally author, are considered high-quality data and thus help researchers to address many concerns that were not clear before [6]. Many studies and research which covers several topics and domains are being investigated using social media data such as behavioural interventions [7], health domains as chronic physical illnesses [8] and smart health [9], job satisfaction [10], social science and personality [11] [12], political science [13] and education [14]. These types of qualitative data offered by social media platforms allowed researchers to adequately investigate new research questions within a specific and picked group with different beliefs and backgrounds without bias.

There are studies and projects such as Mypersonality [15], and ApplyMagicSauce [16] are using textual data to infer users' personalities by utilizing the publicly available information as textual posts in social media profiles. This type of advancement in the research allows a better interpretation of human online psychology and helps create a new understanding of how users are using social networking sites.

In this research effort, we seek to extend the literature and contribute to models which are developed for inferring humans' personality traits by incorporating text as inputs, Like metadata objects as inputs, five reaction buttons, and emojis as inputs, and finally, the profile picture as inputs. Furthermore, we compare the personality development of the same user over different social platforms. In other words, we examine whether users' personalities differ from one platform to another, and in the end, a study about personality stability and change using social media data is presented. In this thesis, a comprehensive study to check whether computers could use these different data types such as (posts, likes, emojis, and profile pictures) shared by online users to predict individuals' personalities based on the big five personality model is concluded.

## 1.2 Research Questions and Contributions

This thesis investigated the relationship between users' Big five personality traits and the published information on their social media platforms. Various types of the Big 5 personality traits are examined: extroversion, agreeableness, openness, conscientiousness, and neuroticism. For this reason, an extensive literature review has been done to identify the open challenges in the domain. We engineered and created psychological features extracted from textual status updates from the Facebook social network platform by using an enormous sample of 110,000 users to address and extend the previous research advancement. Several predictive models are constructed by using different machine learning algorithms such as support vector regressor, an ensemble learning XGBoost, and neural networks. A grid search optimization for the learning parameters and cross-validation for the training samples utilizing features that were found significantly correlated with Big 5 Personality traits extracted and highlighted by Pearson correlation coefficient and boosting trees are studied. The predictive models are constructed based on the best combinations of extracted features and classifiers. The final predictive performances are evaluated based on well-established benchmarks in the domain using metrics such as root mean squared error, f-score, and mean squared error.

The second contribution is carried out by creating a new and novel methodology to estimate users' Big five traits by investigating their likes records on the Facebook platform. The previous research and past contributions use the like record directly to create statistical models to infer users' personalities. In this contribution, a new method to fit the problem by collecting extra information about the likes records from the Facebook API to create more sophisticated personality models is created. We used and employed the methodology that Facebook provided by crawling the metadata about the objects of the likes. The category and sub-category associated with likes records provided by the API to create more accurate data samples representing users' likes space is used. Therefore, by using the newly generated information about the likes data from the used dataset and applying different resampling techniques, linear regressions, boosted trees, k nearest neighbors, and neural networks to build personality profiles from users' likes are trained. The influence of the size of the used datasets is shown. We emphasized that although the data associated with the metadata of likes objects is a small proportion of the total information a user leaves on social networks, the trained models using these data can accurately estimate an individual's personality.

The third contribution is about creating a machine learning model and a novel framework that uses the new set of reaction buttons that Facebook introduced to their users (to interact with the social contents at their platform with love, care, sadness, anger, and laugh) to predict and classify posts' emotions and reactions. The proposed framework is trained on a huge dataset of 64000 public unique Facebook pages and their published 3 million labeled posts associated with the reactions collected by our scalable Facebook crawler. This contribution creates a baseline for other researchers to use the dataset we made public on the Harvard Dataverse website as a benchmark for the emotions classification task. Several filtering techniques such as minimal reaction count and reaction gap filter are used to extract further a group of features (CountVectorizer, Term Frequency–Inverse Document Frequency, Punctuation, and Google Embeddings word2vec) to train multinomial Naive Bayes classifier to handle the problem. The evaluation against standard benchmarks using the proposed new dataset and used features demonstrates promising results compared to previous research.

A contribution to predicting the personality from the profile picture, particularly from facial images, is presented in chapter four. The previous literature predicts the personality from labeled samples where users provided their face photos and answered a personality-related questionnaire. In this contribution, Twitter API is used to collect random public users' profiles to investigate tweets to further create a labeled dataset for profile pictures. The whole work is made to raise the awareness for social media users of what third parties can reveal about their private traits without their consent. A method to label profile pictures with personality scores from textual tweets is shown. To train the final models, almost 80 unique facial features from the images defined by a tool called Face++ are extracted and investigated. These features were then used to investigate the relationship between the personality scores generated from their tweets and the values of their facial features. After splitting 80% of the data to the train, the correlated facial features with the personality labels and 20% for the test to train a set of different machine learning algorithms such as support vector regressor, random forest, bagging, and adaptive boosting are investigated. The testing and the evaluation results showed promised results in automating predicting users' personalities from publicly available profile pictures.

The fifth and the sixth contributions are made to comprehend the personality of online social networks in different setups. The constructed personality models that can predict personality from textual inputs to measure the difference in the personality development for the same users' set based on their published content on the cross-platform environment as Facebook and Twitter social platforms are used. We compared the constructed personality profiles based on the shared textual inputs and showed how different platforms affect the building and the construction of the final Big 5 personality traits. In the sixth contribution, an investigation of the concept of personality stability and continuity using textual samples extracted from social media platforms is presented. The implemented experiments are based on a ten-year longitudinal study on Facebook to understand users' personality traits development over time. The concept of stability coefficient metric is applied to demonstrate users' online identity development based on the published linguistic contents over time by involving advanced machine learning and neural network algorithms.

In general, this thesis aims to investigate several of the challenges and opportunities associated with predicting and estimating users' Big five personality traits utilizing online textual and non-textual data. Specifically, different methods to engineer and create different feature sets to be further examined using machine learning algorithms for mining digital identities are shown. Furthermore, cross-domain research is presented between the data science and psychology domains to advance psychologists' understanding of how humans' personality traits develop over time using social media data. This thesis is developed to answer the following research questions;

1. How do the Big 5 personality traits get measured based on users' linguistic textual inputs from social media platforms?

2. How do the Big 5 personality attributes get estimated based on users' public likes and reaction buttons on social media platforms?

3. How does the personality get estimated based on users' uploaded images on social media platforms?

4. Does a user's personality change from one social media platform to another?

5. Does a user's online personality evolve over time?

## 1.3 Thesis Outline and Structure

This thesis is structured as follows: In chapter 2, we review the literature on the effectiveness of utilizing social media platform data to examine the relationship between users' personality traits on the Facebook social media platform to answer the first proposed research question. Using myPersonality project dataset, individuals' personality traits are considered by investigating a broad set of linguistic features that mimic personality characteristics. Eighty-two linguistic-based features are explored by implementing two feature extraction methods (Pearson correlation and gradient boosting) to evaluate the best performing features in defining the five personality traits. Therefore, three different machine learning algorithms (support vector regression, gradient boosting, and feed-forward- neural network) to predict the personality scores from textual inputs after obtaining significant features from the text by employing the closed vocabulary approach are utilized. This experiment's most significant personality prediction scores were obtained by training the XGBoost machine learning models using the defined Union features between the Common and Own sets extracted by boosting trees. The combination of the linguistic features we extracted exposes a high potential of using linguistic social network post features for automating the personality estimation tasks.

The first part of the second research question in chapter 3 introduces the concept of using likes metadata derived from Facebook API payload as features for training machine learning models for the personality computation task. We presented a method that shows how the Big Five personality scores can be quickly estimated by leveraging the information about the pages a person liked on Facebook without taking his/her post into further analysis. The results validate the significance of the correlations between users' personality and their Facebook Likes history and Likes categories.

In chapter 4, the second part of the second research question presented a novel framework for predicting emotions and reaction distributions for any given post within the Facebook social media platform. Facebook introduced a new reactions feature that allows users to express their psychological emotions regarding published content using the five Facebook reactions buttons. The potential of using Facebook reactions to identify and distinguish human emotions is examined. For this task, we gathered an enormous amount of Facebook posts associated with their reaction labels using the introduced scalable Facebook crawler. The final model can predict the five reactions distribution on Facebook posts in an automated way.

The third research question is addressed in chapter 5. An explanation of how pictures users post and share publicly on their social media platforms is motivated and explained by the psychological constructs that psychologists identify as personality traits are discussed. We examine how social media profile pictures differ based on the users' personalities posting them. Profile images from the Twitter platform whose personalities are predicted based on 1.7 million data points are handled. The outcomes suggest differences in the profile pictures selection between users from different personality traits.

In chapter 6, the answer to the fifth and the sixth research question is presented. That chapter showed how we trained a language model on Facebook data to predict personality dimensions on the Twitter platform. In the same chapter, we also explored continuity and stability in personality for social media platform users. We tracked individuals and evaluated their linguistic behaviour for ten years in a longitudinal study to reveal how personality traits develop over time. Results suggest that using advanced machine learning and stability coefficient measurements show two kinds of patterns that human personality follows in their online identity development across the life span: the inter-individual development pattern and intra-individual development patterns. Multiple cases of stability and continuity based on users' public linguistic language are showed.

Several of the ideas and findings contained in different parts of this thesis have been published previously:

**Chapter 2:**
-Raad Bin Tareaf, Philipp Berger, Patrick Hennig and Christoph Meinel.**Personality Exploration System for Online Social Networks: Facebook Brands As a Use Case**. *In Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence (WI, IEEE)* [12].

**Chapter 3**:
-Raad Bin Tareaf, Seyed Ali Alhosseini,Philipp Berger, Patrick Hennig and Christoph Meinel. **Towards Automatic Personality Prediction Using Facebook Likes Meta-**

**data**. *In Proceedings of International Conference on Intelligent Systems and Knowledge Engineering (ISKE, IEEE)* [17].

**Chapter 4**:
-Raad Bin Tareaf, Philipp Berger, Patrick Hennig and Christoph Meinel. **ASEDS: Towards Automatic Social Emotion Detection System Using Facebook Reactions**. *In Proceedings of International Conference on High Performance Computing and Communications; IEEE International Conference on SmartCity; IEEE International Conference on Data Science and Systems (HPCC/SmartCity/DSS, IEEE)* [18].

**Chapter 5**:
-Raad Bin Tareaf, Seyed Ali Alhosseini and Christoph Meinel. **Facial-Based Personality Prediction Models for Estimating Individuals Private Traits**. *In Proceedings of IEEE International Conference on Social Computing and Networking (SocialCom/IEEE)* [19].

**Chapter 6**:
-Raad Bin Tareaf, Philipp Berger, Patrick Hennig and Christoph Meinel. **Cross-platform personality exploration system for online social networks: Facebook vs. Twitter**. *In Web Intelligence Journal (WIC/ IOS Press)* [20].

-Raad Bin Tareaf, Seyed Ali Alhosseini and Christoph Meinel. **Does Personality Evolve? A Ten-Years Longitudinal Study from Social Media Platforms**. *2020 IEEE Intl Conf on Parallel Distributed Processing with Applications, Big Data Cloud Computing, Sustainable Computing Communications, Social Computing Networking (ISPA/BDCloud/SocialCom/SustainCom)* [21].

Some papers and journals that were published during the Ph.D. study but not included in this thesis because of a topic difference:

-Raad Bin Tareaf, Philipp Berger, Patrick Hennig and Christoph Meinel. **Malicious Behaviour Identification in Online Social Networks**. *In Proceedings of International Federated Conference on Distributed Computing Techniques (DAIS/Springer)* [22].

-Seyed Ali Alhosseini, Raad Bin Tareaf, Pejman Najafi and Christoph Meinel. **Detect Me If You Can: Spam Bot Detection Using Inductive Representation Learning**. *In the companion of The World Wide Web Conference (WWW/ACM)* [23].

-Raad Bin Tareaf, Philipp Berger, Patrick Hennig, Sebastian Koall, Jan Kohstall and Christoph Meinel. **Information Propagation Speed and Patterns in Social Networks: A Case Study Analysis of German Tweets**. *In Journal of Computers (J. Comput)* [24].

-Ahmed Shams, Raad Bin Tareaf, Jan Renz and Christoph Meinel. **Smart MOOC - Social Computing for Learning and Knowledge Sharing**. *In proceedings of International Conference on Computer Supported Education (CSEDU/SciTePress)* [14].

-Raad Bin Tareaf, Philipp Berger, Patrick Hennig, Jaeyoon Jung and Christoph Meinel. **Identifying Audience Attributes: Predicting Age, Gender and Personality for Enhanced Article Writing**. *In Proceedings of the International Conference on Cloud and Big Data Computing (ICCBDC, ACM)* [25].

- Hanadi Traifeh, Raad Bin Tareaf and Christoph Meinel. **E-Learning Experiences from the Arab World**. *In proceedings of International Conference on Advanced Research in Education, Sorbonne University, Paris, France [26].*

-Seyed Ali Alhosseini, Raad Bin Tareaf and Christoph Meinel. **Engaging with Tweets: The Missing Dataset on Social Media**. *In RecSys Challenge: In proceedings of the Recommender Systems Challenge (RecSys Challenge/ACM) [27].*

# 2 Personality Prediction from Textual Inputs

In this chapter, we will explain the big five personality traits model and how the model is being used to measure individuals' traits. Furthermore, we will investigate how users' personalities can be estimated based on their textual inputs on various social media platforms. As the Big Five model is the most used and universally accepted and trusted framework for personality estimation, the rest of this thesis will concentrate exclusively on the five-factor framework.

## 2.1 Introduction

Users of online social media platforms produce many different kinds of content while they leverage the platforms. The importance of such data is that it can be used with the help of domain experts to address better and improve software and services based on user's preferences. Therefore, collecting and interpreting this huge content provides a new strategy for industries to personalize services and better understand their final users.

In the past few years, capturing and understanding users' traits required a lot of time and effort from both entities (users and psychologists) by asking individuals to fill out lengthy questionnaires in clinical settings. For psychologists to capture and measure individual personality traits, it takes the participants to answer 15-300 related personality questions where most of these participants will tend to be prone to social-desirability bias. Clinical setups are also costly and require higher interaction with users, which is why they did not reach a large population. However, with the advancement in digitization, researchers from various domains are competing to collect and understand the publicly available online social media data to fulfill their needs. In this chapter, we will investigate several of the previous researchers' approaches and how they utilized the data in building the Big personality profiles in online spaces and finally introduce our contributions in building and advancing the personality prediction models that can estimate and measure the Big 5 personality traits of any given users based on his/her published textual inputs.

Research has suggested that humans' behaviour can be understood by psychological constructs, which are designated as personality traits [28]. Therefore, multiple personality models have been proposed and investigated over time across multiple locations and cultures. One of the most well-known personality models is the five-factor personality model introduced and reviewed by [29]. This five-factor model is built on the relationship between terms and words users use and their reactions and experiences to specific events.

Psychologists concluded that human personality (online or offline) could be defined and classified based on the five global factors: Openness, Conscientiousness, Extraversion, agreeableness, and Neuroticism.

Paul Costa and Robert McCray [30] and the psychologist Lewis Goldberg [31] were among the first researchers who introduced the five factors models under the name of Big Five Personality traits abbreviated OCEANS with the first psychometric test. Based on the fact that the Big 5 model is built on lexical research effort, estimating individual traits is a natural language processing task that results rely heavily on the users' used language and words. Therefore, the Big 5 model's dimensions hold a unique meaning represented with a group of word categories and emotions that reflect final individual personality.

It is essential to know the general descriptions for each of the Big 5 traits [32]. Individuals assigned as Openness to experience in their personality are generally intellectually unique and willing to try new things in their lives. The personality of a user who scores high in Openness is imaginative, creative, and insightful. The second personality trait is conscientiousness. A dominant personality trait of conscientiousness reflects the trait of punctual self-controlled, and responsible people. The Personality trait extraversion is the trait that summarizes users who are passionate, talkative, and active. The neuroticism personality trait is directed to distress and dissatisfaction, which implies constant changes in mood and emotions. Highly neurotic people tend to be anxious, tense, and withdrawn. Individuals who are high in neuroticism tend to be worrying, temperamental, and vulnerable, whereas individuals with low scores in neuroticism are represented as self-satisfied, unemotional, and calm. The last personality trait is agreeableness. A dominant score of agreeableness for a user's personality means he/she is more likely to be trustworthy, generous, and kind.

It took psychologists several years in psychology research to validate the personality-related terms and the five-factor personality models. After that, the five-factor models became the most accepted and widely used model over different languages and cultures. It reflects the traits, characteristics, and terms that a user may use and adapt. However, the previous and traditional approaches require individuals to answer predefined questionnaires/surveys, which can be time-consuming and impractical in all possible scenarios. One of the most used personality-related and acceptable questionnaires is the revised NEO personality inventory questionnaire [31], which consists of 20 to 360 personality-related questions.

This chapter is organized as follows. Section 2.2 introduces the previous literature in the domain of individual personality prediction. Section 2.3 discusses data acquisition, feature selection, and implementation criteria concerns. In section 2.4, we illustrate different experiment results against various machine learning classifiers such as support vector regressor, XGboost, and feed-forward neural network and visualize insights for the following evaluation method. Finally, section 2.5 summarizes the final results with redirection and suggestions for future work in automating personality detection.

## 2.2 Background

Various features, algorithms, and datasets have been studied in machine learning for automating personality predictions from textual inputs. The research examines how writing style and content vary between different personality types, highlight the significant differences and demonstrate the extent to which such differences can be used to identify users' personality. This section will document the most important studies that leveraged textual inputs as a feature to assess online users' personalities.

In order to study and identify the personality of individuals, researchers start analyzing the textual inputs that users on various social media platforms produce. As personality is encapsulated in language [33] [34] [35], psychologists have approved that there are a clear relationship and high correlations between the personality traits and the linguistic variables (structural, lexical categories, speech type, character-level, word-level, and n-grams) and psycholinguistic characteristics such as (emotional affects, perceptions and social relationships [36] [37] [38]. Researchers generally approach this task by considering two features: linguistic textual features as described above and non-linguistic textual features as social network features such as (network size, betweenness, eigenvector centrality, and many other graph measurements). This section will focus on the previous work of predicting the Big 5 personality traits using textual features.

With recent advances in computational social science and data mining, data scientists have explored the potential of predicting personality and many other private traits based on users' online social profiles. Therefore, multiple scientists applied different methods with different features using machine learning algorithms to solve this task. For instance, some researchers applied supervised machine learning algorithms as shown in [39] [40] [41] [42], while others tried unsupervised [43] [44] [45], some used hybrid machine learning algorithms [46] [47] and lately deep learning [48] [49] [50] to estimate users personality profile.

Disregarding the final machine learning types that are used to predict the traits, all of the proposed models are trying to solve the task with minimal errors and higher accuracy. One of the first works for predicting personality using Facebook profiles was introduced by [51]. He used a 45-question version of the personality questionnaire for the used sample. After applying some preprocessing steps, he reduced the final features set, which contains different extracted features (language features and activities), from 161 to 74 per individual. Then he applied regression analysis by using the Weka tool to analyze and generalize the final produced personality models. Also, one of the efforts for predicting personality using Twitter profiles was introduced by [52]. He used the well-known MyPersonality facebook dataset to study 335 Twitter users' accounts, which was shared on the Facebook study samples. He applied the Pearson correlation between the user characteristics and the big five traits results to address the most significant predictor for the five traits; then, he generated regression analysis with cross-validation using m5 rules to solve the task of personality recognition from the Twitter platform. Oberlander et al. [36] classified conscientiousness, agreeableness, stability, and extraversion personality traits of blog authors from their posts. They carried a series of experiments with several percentiles by leveraging Naive Bayes as a classifier and n-grams as features, which concluded that automatic feature selection and binary classes yielded the most remarkable improvement over the baseline.

Naive Bayes classifier was investigated by [47] which is a simple model that describes a particular class where all of the features are class conditionally independent to predict the personality from a Chinese social networking site. He leveraged the 44-item questionnaire on his sample associated with other features such as time-related and emotion-related usage to analyze emotion usages against content. He used multiple types of classification algorithms like support vector machines and decision trees. Another study where the researcher decided to investigate a vast feature space with a ranking algorithm for features selection was introduced by [53]. To address the task, he used the Facebook MyPersonality dataset to extract more than 700 deep linguistic features, including social and demographic information. He used different types of algorithms like support vector machines and different ensemble learning-based models such as adaptive boosting and multi-boosts to create accurate personality recognition models. Also, [54] used ensemble learning-based models to enhance prediction accuracy (by joining multiple weaker learners) on the Facebook MyPersonality dataset and essay labeled samples. Using the concept of trigrams on the essays corpus, he used the ensemble methods based on meta-learning to combine predictions of various classifiers and then test the models for personality trait prediction.

Mairesse et al. [55] labeled personality with both conversation and textual inputs. The conversation was labeled using observers' judgments domain experts, and the text was labeled using self-assessment using the Big five personality questionnaire. They used different features extraction methods such as MRC, the medical research council psycholinguistic database, LIWC, and two lexical features, and predicted both personality scores and classes using SVM and M5 trees as classifiers. They published a list of correlations between Big5 personality traits and the two lexical features they adopted.

Iacobelli et al. [56] published a study that concentrated on linguistic characteristics of personality and compared different features sets extracted by open vocabulary approach 1-gram and 2-gram and closed vocabulary approach from linguistic inquiry and word count tool. Their used dataset consists of writing examples of a corpus collected from three thousand bloggers over several months. The bloggers completed a five-item personality assessment test with yes or no questions. Only high or low samples from these personality traits were finally included for further analysis after applying re-balancing techniques, which improved the overall accuracy by 2-3% compared to samples that included the whole dataset. The authors highlighted that the classification might overfit the samples because few bigram features are extracted from each personality trait's space. They trained a support vector machine and naive Bayes classifiers to classify personality traits from textual inputs, and their best-reported results are 70.51% for neuroticism and 84.36% for openness to experience traits achieved using textual inputs bigrams as features.

Another exciting work was published by Golbeck et at. [57] for predicting personality from Twitter. The authors used a sample of fifty users recruited through public posts on Twitter. The users answered a personality test to capture their big five traits scores using 45-personality related questions. Besides the personality test results from users, they collected the last 2000 tweets, including a set of statistics related such as the number of followers, number of following, the density of the social network, number of replies, number of a hashtag used, number of links and word per a tweet from their accounts. To analyze the content of the user tweets, they used linguistic inquiry and word count tool excluding the count feature to capture 80 different word categories. They also used the MRC psycholinguistic database with word by word sentiment analysis to assign words sentiment on a -1 to +1 scale. They used Pearson correlation analysis between the dependent and independent variables to predict the score of a given personality trait using a regression analysis such as Gaussian Process and ZeroR with 10-fold cross-validation in the Weka tool. In the end, they highlighted the importance of creating automatic personality prediction tools and how such tools can be integrated into personality-oriented interfaces and recommender systems.

The latest research efforts for personality prediction from the MyPersonality dataset for 250 users is proposed by Christian et al. [58]. They used multimodal deep learning architecture combined with multiple pre-trained language models such as BERT, a Bidirectional Encoder Representation from Transformers, RoBERTa, with features extracted as XLNet, sentiment analysis, TF-IGM statistical features, and NRC lexicon database for the task. They showed multiple comparisons for different features extraction methods and trained models, resulting in higher accuracy than pre-trained models. Ren et al. [59] proposed a Bidirectional Encoder Representation from Transformers also to generate sentence-level embedding with sentiment dictionary for the task, and they evaluated their model on two different public personality datasets, which are MBTI personality dataset and MyPersonality. Majumder et at. [49] developed a document modeling technique based on CNN features extractor by bypassing sentences from the samples to convolution filters to extract the sentence model in the n-gram feature vector form to be later fed into a fully connected neural network with one hidden layer. The sample they used is an essay dataset of 2467 users tagged with the big five personality scores. Their proposed document vector CNN with Mairesse using multiple engineered filters on sMLP and MP classifiers yielded the most accurate models for extraversion and agreeable personality traits.

One of the most significant studies in personality and language is presented by [60] using the collected MyPersonality dataset. The researchers investigated seventy thousand users who took standard personality tests to create and understand social media users' word usages addressed by personality traits and age. They used linguistic inquiry and word count tool, topics extracted from Facebook posts automatically by LDA topic clusters, and word phrases using n-grams of the size 1 to 3 as features to investigate several predictive models. Using word clouds, they visualized the most distinguishing female and male words, phrases, and topics. In addition, they showed words and phrases for different ages and personalities, which allowed them to know the most common words used by age and personality in their studied samples. They used a quantitative evaluation approach to understand different features extraction for the predictive evaluation. They used a linear support vector machine for classifying male and females and a ridge regression algorithm for the five personality traits and age, where both algorithms leveraged a regularization parameter reporting the performance in R score. They used principal component analysis on fifty percent of their studied samples to minimize their features space.

In this work, using writing style from LIWC tool as a feature defined by ($p < \alpha^* = 0.01$) correlation relationship reduced by ($|r| > 0.05$), we will classify all the five personality traits (Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness) with multiple machine learning algorithms such support vector regressor, decision-tree-based ensemble algorithms such as XGB and deep neural network with a sample of 110,000 users from MyPersonality dataset for the given task. Also, three different feature sets and their combinations extracted by two different approaches will be investigated to maximize the final accuracy in estimating online personality traits.

## 2.3 Materials and Methods

This section will report the complete characteristics of the used datasets and the proposed methods to learn the hidden relationships among the dependent and the independent variables. In the end, we will perform various experiments and then evaluate the final results.

### 2.3.1 Methodology

As the amount of research interested in capturing and predicting personality from social media texts is snowballing, different methodologies and frameworks that use all online fingerprints to predict personality are investigated. Our proposed personality prediction framework is shown in Figure 2.1 which starts by collecting and filtering records from the MyPersonality dataset for preprocessing and cleaning techniques. After this step, several feature engineering and extraction experiments are proposed to build different types of machine learning models. We investigated supervised and unsupervised machine learning models to solve the fitting task, then evaluated and predicted the testing samples.

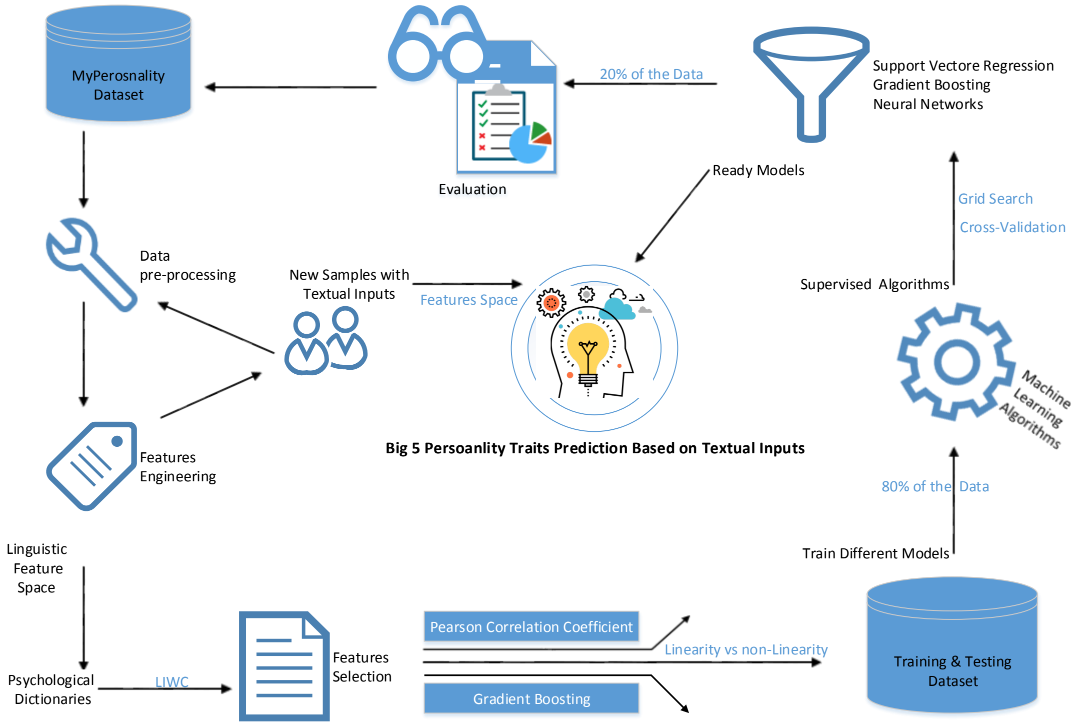



**Figure 2.1:** Personality Prediction Framework from Textual Inputs

### 2.3.2 Data Collection

For the task of personality prediction from textual inputs, very few available datasets meet the needed requirements. Therefore, after extensive research in this domain, we found a unique dataset representing a popular Facebook app in 2012. MyPersonality dataset is a well-known dataset that represents a Facebook application introduced by [61] in 2007. The authors showed how the Facebook platform could be employed in psychological studies, and they also outlined the most critical opportunities and challenges affiliated with the Facebook platform as a research tool.

The Facebook application enabled users who agreed to participate and answer different psychometric tests to share their complete profiles with their final personality scores. The personality scores are derived from a test they decided to answer using the revised NEO personality inventory [62] which they decided to share with the authors for research purposes. As mentioned by [61], more than ten million participants joined the experiment, and approximately over 2 million users decided to share and donate their data to the primary authors for further research. In general, the dataset consists of almost six million test results and almost four million unique Facebook profiles from the United States.

For this task, we used a partition of this enormous dataset that fits predefined criteria, including only the samples with at least a specific predefined amount of words and exclusively in the English language. We investigated three main tables from MyPersonality project for our research. The first table is called *Big5 Personality Scores* which represents the five-factor personality generated scores for three million users where their personality scores are shown in a range from 1 to 5. We also considered the final questionnaire size that the final personality scores are generated based on. The second table we used for this task is *Facebook Status Updates* which contains more than 25 million text status and posts updates from the participants. The final table we used from MyPersonality project for this experiment is *Demographic dataset*. The last dataset contains demographic details, age, gender, relationship status, number of friends, and timezone information for the participated participants.

All the used datasets contain anonymized user identifiers that prevent us from knowing the real Facebook account and can only be used to understand and connect other tables. Our task in this chapter is to predict individuals' five personalities from their English textual fingerprints. For this purpose, we gathered and joined the Big five personality scores data table with the status posts data to reassure the correct extraction of LIWC features, as we will explain in the feature extraction section. At the end of this step, each individual in the newly generated dataset is represented with his/her linguistic features based on the status updates annotated and labeled with the Big five personality scores. Table 2.1 gives an overview of the final used dataset for the training and testing models. We applied a random split to extract 20% of the samples for testing.

**Table 2.1:** Characteristics of the Textual Training Dataset [12] ©2018 IEEE.

| **Characteristics** | | |
|---|---|---|
| # Samples = # Users | | 108547 |
| # Male Gender | | 44844 |
| # Female Gender | | 63245 |
| # Extracted Linguistic Features | | 93 |
| # Big5 Personality Labels | | 5 |
| Avg. Age | | 27 |
| **Labels** | **Mean** | **Standard Deviation** |
| Openness | 3.8435 | 0.6759 |
| Conscientiousness | 3.4631 | 0.7358 |
| Extraversion | 3.5068 | 0.8135 |
| Agreeableness | 3.5659 | 0.7070 |
| Neuroticism | 2.7334 | 0.8003 |

*LIWC* as described in the next subsection as an application that analyses textual inputs after being manually preprocessed from stop words in terms of 82 categorical elements such as verb, pronoun, past tense, etc. Each of the 82 categories has its own list of words that a user's textual inputs can be classified. To mention some, the category auxiliary verb in LIWC dictionary includes "can", "cannot", "cant". While the category social in LIWC dictionary includes "receiv*", "ask", "say*", etc. Using such a tool, the derivatives from the same word form can be categorized as the same word, disregarding the word tense, case or plural forms.

### 2.3.3 Personality Features Extraction

In this section, we will discuss creating features from textual inputs. For this task, we used the academic version of the LIWC tool (Linguistic Inquiry and Word Count) to analyze psychological word usage for all studied samples. There are other alternatives for this tool as SPLICE and MRC, but we decided to utilize LIWC in this task as it had already shown its advantage in literature for the personality computational task. LIWC is a dictionary containing words that belong to different and multiple word categories. The tool performs a text analysis over given textual inputs to return the percentage of words that appeared in one or more of the 82 linguistic, psychological, and topical categories representing the various social, cognitive, and affective processes which are four general descriptor categories, 22 standard linguistic dimensions, 32-word categories tapping psychological constructs, seven personal concerns categories, three paralinguistic dimensions and 12 punctuation categories [63].

| ipron | article | prep | auxverb | adverb | conj | negate | verb | adj | compare | interrog | number | quant | affect | posemo | negemo | anx | anger | sad | social | family | friend | female | male |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.35 | 0.02 | 1.19 | 1.38 | 3.68 | 1.13 | 0.72 | 11.96 | 6.67 | 1.62 | 0.19 | 0.83 | 1.74 | 9.25 | 6.16 | 2.95 | 0.27 | 1.04 | 0.41 | 9.82 | 0.15 | 0.25 | 0.18 | 0.27 |
| 1.54 | 0.04 | 1.06 | 1.32 | 5.44 | 1.34 | 0.69 | 11.94 | 6.58 | 1.84 | 0.27 | 0.61 | 2.00 | 10.97 | 7.66 | 3.20 | 0.42 | 0.71 | 0.51 | 7.73 | 0.12 | 0.42 | 0.07 | 0.12 |
| 0.49 | 0.00 | 0.84 | 1.51 | 1.71 | 0.62 | 0.32 | 8.83 | 5.27 | 1.60 | 0.13 | 0.46 | 1.12 | 10.09 | 8.11 | 1.85 | 0.24 | 0.86 | 0.22 | 7.38 | 0.13 | 0.39 | 0.03 | 0.07 |
| 0.69 | 0.02 | 0.73 | 0.94 | 1.98 | 0.42 | 0.44 | 9.41 | 5.76 | 1.39 | 0.03 | 0.90 | 0.91 | 9.01 | 7.37 | 1.59 | 0.21 | 0.55 | 0.26 | 6.98 | 0.78 | 0.16 | 0.13 | 0.08 |
| 1.30 | 0.06 | 1.33 | 1.46 | 4.16 | 0.98 | 0.76 | 13.84 | 6.75 | 1.82 | 0.10 | 0.88 | 2.07 | 11.21 | 7.45 | 3.65 | 0.29 | 1.62 | 0.57 | 9.93 | 0.13 | 0.19 | 0.20 | 0.27 |
| 0.53 | 0.00 | 0.65 | 0.89 | 1.88 | 0.64 | 0.21 | 8.72 | 5.45 | 1.16 | 0.05 | 0.75 | 0.92 | 9.17 | 7.60 | 1.31 | 0.15 | 0.32 | 0.36 | 10.14 | 0.28 | 0.29 | 0.04 | 0.00 |
| 1.13 | 0.11 | 0.94 | 1.07 | 2.95 | 1.08 | 0.64 | 10.16 | 4.57 | 1.16 | 0.22 | 0.59 | 1.24 | 8.23 | 4.44 | 3.72 | 0.34 | 1.61 | 0.56 | 9.51 | 0.15 | 0.31 | 0.11 | 0.22 |
| 1.18 | 0.07 | 1.63 | 1.66 | 3.47 | 1.07 | 0.66 | 11.06 | 4.66 | 1.58 | 0.18 | 0.95 | 1.17 | 10.25 | 4.53 | 5.68 | 1.08 | 2.72 | 0.45 | 8.94 | 0.27 | 0.52 | 0.16 | 0.52 |
| 0.99 | 0.16 | 1.42 | 1.00 | 2.90 | 0.94 | 0.84 | 8.90 | 4.71 | 1.51 | 0.09 | 0.54 | 1.18 | 7.76 | 4.15 | 3.56 | 0.33 | 1.13 | 0.54 | 6.80 | 0.48 | 0.16 | 0.10 | 0.22 |
| 1.25 | 0.06 | 1.40 | 1.55 | 3.68 | 1.14 | 0.80 | 11.22 | 5.21 | 1.47 | 0.15 | 1.02 | 1.43 | 10.77 | 5.09 | 5.60 | 1.02 | 2.25 | 0.48 | 7.27 | 0.15 | 0.29 | 0.21 | 0.23 |
| 1.26 | 0.00 | 0.70 | 1.43 | 1.88 | 0.47 | 0.38 | 10.23 | 6.82 | 2.09 | 0.14 | 0.73 | 1.59 | 9.27 | 7.21 | 2.02 | 0.43 | 0.77 | 0.26 | 6.51 | 0.35 | 0.14 | 0.38 | 0.27 |
| 1.04 | 0.06 | 1.97 | 1.10 | 2.78 | 0.73 | 0.76 | 8.71 | 3.85 | 1.07 | 0.17 | 0.51 | 1.01 | 10.20 | 4.16 | 5.88 | 0.68 | 2.89 | 0.54 | 7.61 | 0.19 | 0.48 | 0.16 | 0.29 |

**Figure 2.2:** Extracting Words Categories from Textual Posts using LIWC Tool.

Each label generated by the LIWC tool is defined as a word category. The category is represented as the percentage of words in all individual textual posts as shown in Figure 2.2. However, some words in the posts are considered in multiple LIWC categories dictionaries. This leads to the sum of all word categories can be greater than 100%, which means all categories are dependent on each other. Therefore, we decided to consider three new different types of features sets per model in order to be able to evaluate options adequately.

In this setup, we enlarged the sample space in the study to include as much sample as we could for the personality prediction task from the MyPersoality dataset where the main previous research work extracted samples based on the fact that they want to predict multiple labels such as political and sexual orientations or life satisfaction and religion in which resulted in using way fewer samples from MyPersonality dataset for the task of personality prediction [16]. The first features set contains all features that are important for each trait, so for each personality trait, we have created its own features set and called it **Own_Features**. A feature set that holds a set of features that are important and shared between all personality traits, we called the set as **Common_Features**. We also investigated the last possible combination of these two features by applying union operation over **Own_Features** and **Common_Features**, and we named it as **Union_Features**. In order to know which is the best performing feature set for the task, we have to investigate and test their correlations and significance with each personality trait. The features selection section shows the evaluation of each of these features against each personality trait.

### 2.3.4 Personality Features Selection

This section will investigate several computational feature selection algorithms that we have applied for the task personality prediction. The outcome of this step is to evaluate and decide the best relevant and performing features for each trait prediction task.

In supervised machine learning algorithms, deciding which essential and relevant features for each class directly affects the final models' accuracy. Several feature selection algorithms are available to investigate, such as the CHI method, symmetrical uncertainty attribute evaluation, information gain, Pearson correlation, correlation-based feature subset, and boosting trees. We used two different approaches to select the best-performing features on the introduced feature sets in this work. The first approach is based on Pearson's Correlation Coefficient, and the other approach is based on the Decision tree features significance.

#### 2.3.4.1 Pearson Correlation

I attributed the leveraged feature sets in the following way: *X-S-M*. $X$ is the influenced trait (e. g. $E$ for Extraversion and $N$ for Neuroticism). This field is optional if the set is independent from the trait. $S$ is one of the introduced feature sets. $M$ defines the used method. $P$ for Pearson and $B$ for gradient boosting. As an example *E-own-P* describes the feature set *own* defined with the Pearson approach for the trait Extraversion.

We applied pairwise Pearson correlation between all the introduced features and the Big five personality scores using *Pearson product-moment correlation* as discussed in detail in [12]. This leads to $m = 5$ correlations with the same feature set. The more inferences are made, the more likely Type-I errors happen. To minimize this multiple comparison problem, we have employed *Bonferroni correction* to the global significance level of $\alpha = 0.05$ to determine the local significance levels: $\alpha^* = \frac{p}{m} = \frac{0.05}{5} = 0.01$.

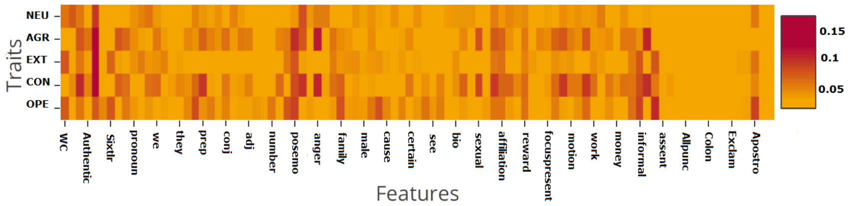



**Figure 2.3:** Pearson Correlation Coefficient Heatmap. Red = High Coefficient [12] ©2018 IEEE.

Correlation results and significance levels between features and all five traits can be found in heatmap 2.3 and heatmap 2.4. Unsurprisingly, not all features are correlated with personality scores. Examples are the word categories *Dash*, *QMark* or *Period*. These punctuations are consistently barely used in status updates and are poor discriminators. Features with overall high relative correlation coefficients are e. g. *tone*, *negemo*, *netspeak* and *Apostro*. Applying the Pearson correlation between the extracted LIWC categories from Facebook posts and personality labels in the training dataset resulted in interesting insights. LIWC category *insight* which represents the (think and know) keywords was found in the language of openness to experienced users. Also, openness to experienced users tends to use perceptual processes and reward categories *percept* and *reward* which contain language keywords look, heard, feeling and take, prize, benefit. The consciousness personality trait tends to use more achievement *achieve*, biological processes *body*, personal concerns *work* and *relig* related keywords such as win, success, body, religion and work.

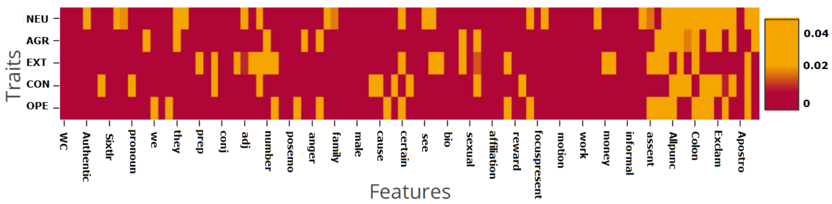



**Figure 2.4:** Pearson Correlation Significance Heatmap. Red = Low Significance [12] ©2018 IEEE.

To get important features to a trait (*X-own-P* feature sets), we only considered features with a significant correlation ($p < \alpha^* = 0.01$) to the selected trait $X$. We reduced the number of features even further by selecting features with high correlations ($|r| > 0.05$). The *common-P* feature set was obtained by selecting features that significantly correlated with all trait labels with $p < 0.01$. Table .1 in appendix .1 shows all resulting feature sets *O-own-P*, *C-own-P*, *E-own-P*, *A-own-P*, *N-own-P* and *common-P*.

#### 2.3.4.2 Gradient Boosting

A different method to determine a subset of features is to utilize the relative importance of features within a gradient boosted regression tree. The concept of this method is to produce a model on all available features. This means we will have to train five models corresponding with each trait to define each feature's importance at the final prediction. The resulting importance features extracted by this method for openness personality traits is shown in figure 2.5.

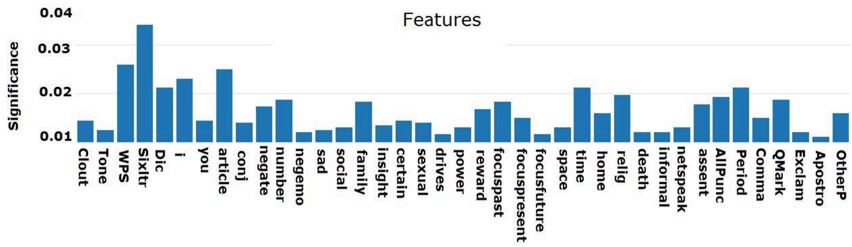



**Figure 2.5:** Significance and Relative Importance for Openness Trait: The diagram contains all features with a relative importance higher than 0:011 value [12] ©2018 IEEE.

We trained one model for each personality trait with all features as input, to retrieve the *own* feature sets. In order to arrange the set *common* we applied the intersection between all *own* sets. Table .1 in appendix .1 shows all resulting feature sets *O-own-B*, *C-own-B*, *E-own-B*, *A-own-B*, *N-own-B* and *common-B*.

## 2.4 Learning Predictive Models

Following the correlation analysis between the dependent variables and independent variables in the previous section, we carried out several experiments to evaluate different machine learning algorithms in this section: Support Vector Regression, Boosted Regression Trees, and Neural Networks will be described below.

As the five-factor model grasps the five distinct personality characteristics, we trained five models for each trait. Each classifier is trained on the three different proposed feature sets (Own, Common, and Union), which we already apprehended using two different approaches (Pearson correlation and Gradient Boosting) for all five traits. This originated 30 models per classifier (three features set, two extraction methods, five personality traits, three machine learning algorithms). As a result, we trained 90 personality models as shown in Figure 2.6 and then evaluated against the 20% test samples, which were split from myPersonaltiy dataset randomly.

The final prediction models should not underfit or overfit the data, and instead, the models should be low in both variance and bias. As personality trait scores are being extended between 1 and 5, estimating individual personality from text is a regression problem. Therefore, we decided to train regression models to map the function from the feature vector to a continuous output variable. After applying a grid search combined with folds for cross-validation, the best performing hyper-parameters are selected to train the models. We compared the 90 generated personality models by calculating the performance on the testing dataset using quality measures. To evaluate our models, we used the root mean squared error ($RMSE$) to estimate the error in predictions for the regression task. $RMSE$ calculate the difference between the actual and the predicted values. $RMSE$ value ranges from 0 to $\infty$, and lower values refer to a lower error rate and, therefore, a better model. With a sample size of $n$, we identified instances by their number $t = 1, \ldots, n$. $y_{t,act}$ stands for the actual observed values, and $y_{t,pred}$ stands for the values predicted by the model defined by the following formula:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (y_{t,act} - y_{t,pred})^2}{n}} \tag{2.1}$$

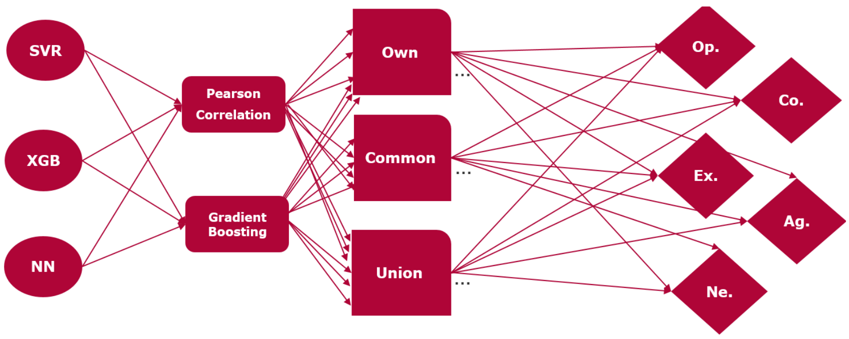



**Figure 2.6:** Actual Learning Phase for the 90 Personality Prediction Models.

### 2.4.1 Support Vector Regression

Support Vector Regressor converts data into a higher dimensional feature space for non-linear base data. As trials with a linear kernel gave unsatisfactory results for the task, I resolved the task by using a Gaussian Radial Basis Function ($rbf$)-kernel which transforms the input vectors in an infinite-dimensional vector space and is defined as [64]:

$$K_{RBF}(\boldsymbol{x}, \boldsymbol{x'}) = \langle \phi(\boldsymbol{x}), \phi(\boldsymbol{x'}) \rangle \tag{2.2}$$

$$= exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x'}\|^2}{2\sigma^2}\right) \tag{2.3}$$

It measures the similarity of two feature vectors $\boldsymbol{x}$ and $\boldsymbol{x'}$ in the input space. The Kernel $K_{RBF}(\boldsymbol{x}, \boldsymbol{x'})$ is large if the euclidean distance between the two feature vectors $\|\boldsymbol{x} - \boldsymbol{x'}\|$ is small. The $rbf$-kernel has one free parameter $\sigma$. The regularization parameter $C$ of SVR and the two hyperparameters can thus be adjusted in the training phase. Figure 2.7 shows the results of the SVR models trained on the feature sets selected by Pearson correlation and Table 2.2 compare the best features set for both approaches against the baseline. Features selected by Pearson correlation are the most acceptable performing models which use the *common-P* feature set. This feature set consists of more features than the trait-specific ones, which after comparison, it yielded better results as the common features set to have a more prominent representation than trait-specific features.

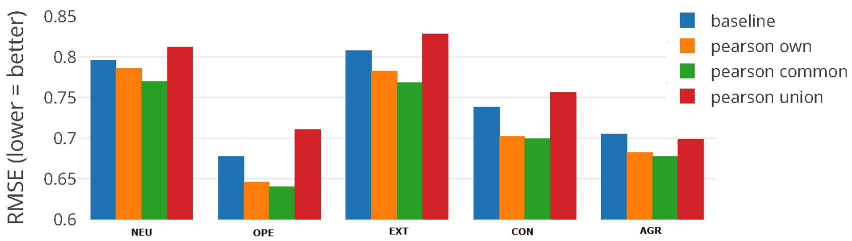



**Figure 2.7:** Evaluation Using RMSE Metric for Different Features Sets Extracted by Pearson Correlation and Trained on Support Vector Regressor [12] ©2018 IEEE.

Figure 2.8 shows the ("RMSE") of the models trained on the feature sets selected by gradient boosting. Models trained on *union* features have a considerably higher ("RMSE") than the other models besides the model for trait Agreeableness.

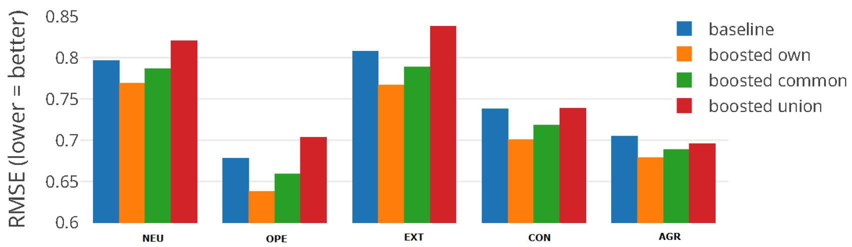



**Figure 2.8:** Evaluation Using RMSE Metric for Different Features Sets Extracted by Boosting Methods and Trained on Support Vector Regressor [12] ©2018 IEEE.

### 2.4.2 XGBoost Models

For gradient boosting training and testing, we used the same method to identify the features' importance. Gradient boosting algorithms consider each leaf to have a scalar as output. As the tree has rigid decision boundaries, the scalar has a fixed number of values. Figure 2.9 shows the results of the ("XGB") models trained on the feature sets selected by Pearson correlation. The models trained on *X-own-P* have lower errors compared to all other models for the traits Openness, Agreeableness, and Conscientiousness. Traits Neuroticism and Extraversion could not outperform the models trained on *common-P* and *X-own-P*. On average, the models trained on *X-own-P* have the highest accuracy.

**Table 2.2:** ("RMSE")s Comparison of the best Feature Sets on Pearson and Boosted with Support Vector Regressor. Bold Values Indicate Lowest Error for the Personality Trait [12] ©2018 IEEE.

| Trait ($X$) | Baseline | *common-P* | *X-own-B* |
|---|---|---|---|
| **Neuroticism** | 0.7963 | 0.7696 | **0.7689** |
| **Openness** | 0.6779 | 0.6402 | **0.6382** |
| **Extraversion** | 0.8079 | 0.7686 | **0.7672** |
| **Conscientiousness** | 0.7381 | **0.6991** | 0.7003 |
| **Agreeableness** | 0.7051 | **0.6774** | 0.6786 |

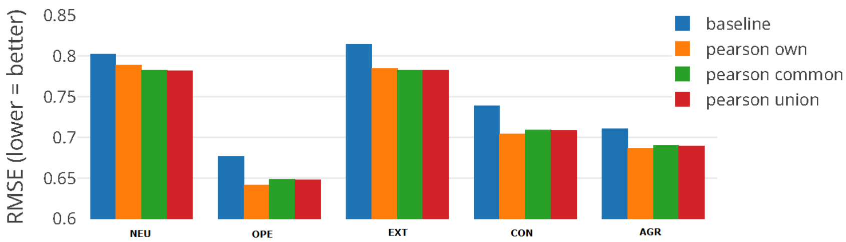



**Figure 2.9:** Evaluation Using RMSE Metric for Different Features Sets Extracted by Pearson Correlation and Trained on XGBoost [12] ©2018 IEEE.

For the feature sets selected by gradient boosting and trained on XGB, there is a higher variance in the resulting ("RMSE"). Figure 2.10 illustrate the results of all possible combinations. Table 2.3 compared the two features selection methods used to train the model. The feature set *common-B* performs the worst compared with other features set for all traits. Models trained on *X-union-B* outperform all other models for all personality traits.

### 2.4.3 Feed-Forward Neural Network

For the neural network part, we trained a feed-forward Neural Network where every perceptron is connected as an input to every perceptron of the next layer. We used four hidden layers with 1024, 512, 256, and 128 perceptrons with some additional layers. Directly following the input layer, we deployed a normalization layer for the input variables, which we also learned during the training of the Feed-Forward Neural Network. We also deployed two more layers as dropout layers to cut a specific *rate* of connections among the perceptrons while training.

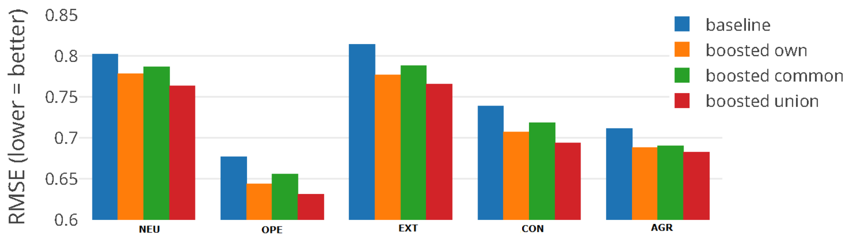



**Figure 2.10:** Evaluation Using RMSE Metric for Different Features Sets Extracted by Boosting Methods and Trained on XGBoost [12] ©2018 IEEE.

**Table 2.3:** ("RMSE")s Comparison of the best Feature Sets on Pearson and Boosted with XGBoost. Bold Values Indicate Lowest Error for the Personality Trait [12] ©2018 IEEE.

| **Trait ($X$)** | **Baseline** | ***X-own-P*** | ***X-union-B*** |
|---|---|---|---|
| **Neuroticism** | 0.7963 | 0.7890 | **0.7635** |
| **Openness** | 0.6779 | 0.6416 | **0.6309** |
| **Extraversion** | 0.8079 | 0.7845 | **0.7655** |
| **Conscientiousness** | 0.7381 | 0.7045 | **0.6934** |
| **Agreeableness** | 0.7051 | 0.6864 | **0.6823** |

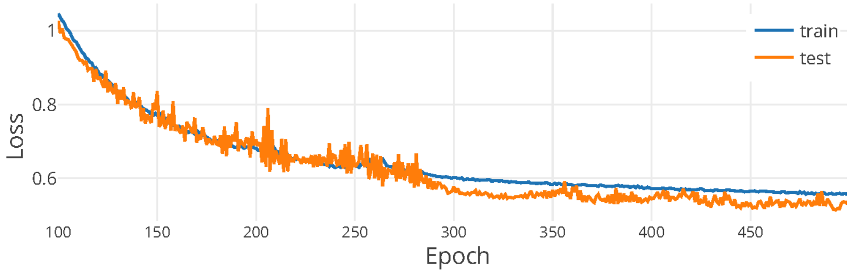



**Figure 2.11:** Loss function for Neural Net based on *A-own-B* starting with epoch 100 [12] ©2018 IEEE.

Fig. 2.11 present a very low noise in loss function for the Agreeableness personality trait model based on the feature *A-own-B*. In other words, the chosen batch size of 50.000 samples during the training is big enough. Comparing the test loss to the training loss, we can conclude that the model has an appropriate state of generalization on the data and does not overfit the problem. Additionally, Table 2.4 shows the RMSE for the best performing algorithm. Lastly, the XGB models, which trained on *X-union-B* outperform other models where SVR better predicted only the Agreeableness trait.

**Table 2.4:** ("RMSE")s Comparison of the Support Vector Regressor and Neural Nets models Trained on *X-own-B* features where XGBoost models Trained on *X-union-B* features. Bold values indicate lowest error for the associated trait [12] ©2018 IEEE.

| Trait | Baseline | SVR | Boosting | NN |
|---|---|---|---|---|
| **Neuroticism** | 0.7963 | 0.7689 | **0.7635** | 0.7611 |
| **Openness** | 0.6779 | 0.6382 | **0.6309** | 0.6525 |
| **Extraversion** | 0.8079 | 0.7672 | **0.7655** | 0.7877 |
| **Conscientiousness** | 0.7381 | 0.7003 | **0.6934** | 0.7279 |
| **Agreeableness** | 0.7051 | **0.6786** | 0.6823 | 0.7069 |

## 2.5 Chapter Summary

In this chapter, we provided an outline and insights for research on social networks and personality from online textual records. The chapter studied the literature on social media platform data as a feature by investigating the relationship between users' personalities and the published public posts on the Facebook social media platform. To estimate an individual's personality, we adopted the MyPersonality dataset to examine a large set of linguistic features that play a crucial role in defining different personality traits. The results reveal that numerous insights can be obtained from studying the linguistic features associated with personality.

We found that using the LIWC dictionary can help advance the cross-domain research domain by opening the road for more data scientists and psychologists to work together. We realized that the 82 linguistic features are vibrant for personality estimation tasks, compared to other features extracted from social media platforms, as shown in the literature review. Consequently, we decided to examine two feature extraction methods: Pearson product-moment correlation and gradient boosting between the available personality traits scores and the linguistic features extracted by linguistic inquiry and word count tool. We used three machine learning algorithms (support vector regression, gradient boosting, and feed-forward- neural network) to predict the personality scores from texts after we extracted all possible features using a closed vocabulary approach. The most outstanding personality prediction scores were achieved by training the XGBoost machine learning models using the defined Union features between the common and own sets of features extracted by boosting trees. Overall, the combination of all linguistic features we extracted reveals a high potential of using linguistic social network posts features for personality estimation tasks. The final models can adequately distinguish between personality dimensions by investigating a broad set of combinations between the extracted linguistic features with state-of-the-art machine learning classifiers.

# 3 Personality Prediction using Likes as Inputs

## 3.1 Introduction

Social media users play a dominant role on social media platforms by generating various types of content as they interact with each other. These interactions vary from sharing messages, liking articles and pages, or reading posts, leaving them digital footprints that can be recorded, extracted, and analysed. Accessing other people's personalities is an essential component of everyday life, steering many interpersonal decisions [65]. The information disclosed through social media platforms' data can align advertisements, political campaigns, and other information streams to the users. Predicting a user's personality using social media data is not straightforward since various aspects contribute. Examples for exciting data points are the like ids of the user, the metadata about these likes, status updates, the content of posted pictures (colors used, are people smiling, situation), metadata about posts (when was it posted, how many likes did it receive) and much other information.

In this section, we created a mapping from each like object to the corresponding category and assessed different machine learning algorithms in predicting a user's characteristics using the relative counts of each like category. We uncovered a relationship between the metadata of the like objects and created different statistical models to predict users' personalities using their likes history.

This chapter is organized as follows. The subsequent section 3.2 presents the related works of predicting users' personalities from their Likes records within social networking sites. Section 3.3 describes the datasets we used and provides a stable view of the actual implementation of the features selection process, followed by an overview of the used and applied research methodology in the research. In section 3.4, we applied a multi-class classification evaluation and decided the most appropriate algorithms for the task. Finally, in Section 3.5, we summarize the final results with redirection of future work for automatic personality prediction using Likes.

## 3.2 Related Work

Only a little previous work has been done to investigate personality from likes. Michal Kosinski et al. [61] demonstrate how Facebook likes can be employed to categorize individuals across a large variety of personal characteristics, e.g., sexual orientation, religion, cultural ethnicity, and more. Their results indicate that more accurate information about an individual personality can be determined computationally by extending and enriching the data, e.g., employing other metadata or browsing history.

Bachrach et al. [66] summarize that personality traits are correlated with patterns of social network use. He showed that an individual's personality could be more accurately predicted using the proposed combination of extracted features. He utilized the number of Facebook "likes" rather than examining which objects were liked for the personality prediction task. He also showed that the Openness personality trait is positively correlated with the number of users' likes, where Conscientiousness is negatively associated with users' number of likes.

Youyou et al. [67] showed that computer-based personality judgments and assessments of individuals are more accurate compared to those made by their Facebook friends. Their experiment also utilized Facebook likes among other data sources to predict Big Five personality scores, along with other statistics, such as life outcomes. They identified key-like objects and handled the likes overlap between users to predict the similarity of users and, therefore, their personality traits score. They attribute these results since more information leads to better informed and, therefore, more accurate personality judgments as computers surpass humans' capabilities in storing and processing these vast amounts of knowledge.

Thilakaratne et al. [68] experimented and applied natural language processing on Facebook textual posts to predict user's personality traits. They considered semantic features rather than syntactical ones to judge human traits. Their research reveals that the number of posts and other statistical information can be used to assess personality traits. In contrast, how users phrase their posts entails much information about their personality. They improved existing computer-based personality assessments by supplementing the syntactical features with their semantical ones.

## 3.3 Methodology

The proposed Likes-based personality prediction models employ users' likes records to formulate a complete personality profile. The framework design is shown in Figure 3.1. We created a new dataset by using the old records of myPersonality dataset to query the Facebook API for the full metadata associated with each like user made to be further mapped and analyzed. The generated dataset is then divided into train and test sets to be used for the training and the evaluation steps. Multiple experiments for features engineering, extraction, and normalization are discussed in the following sections. Neural networks, linear regressions, boosted trees, and k-nearest neighbors algorithms are investigated to fulfill the final personality prediction task.

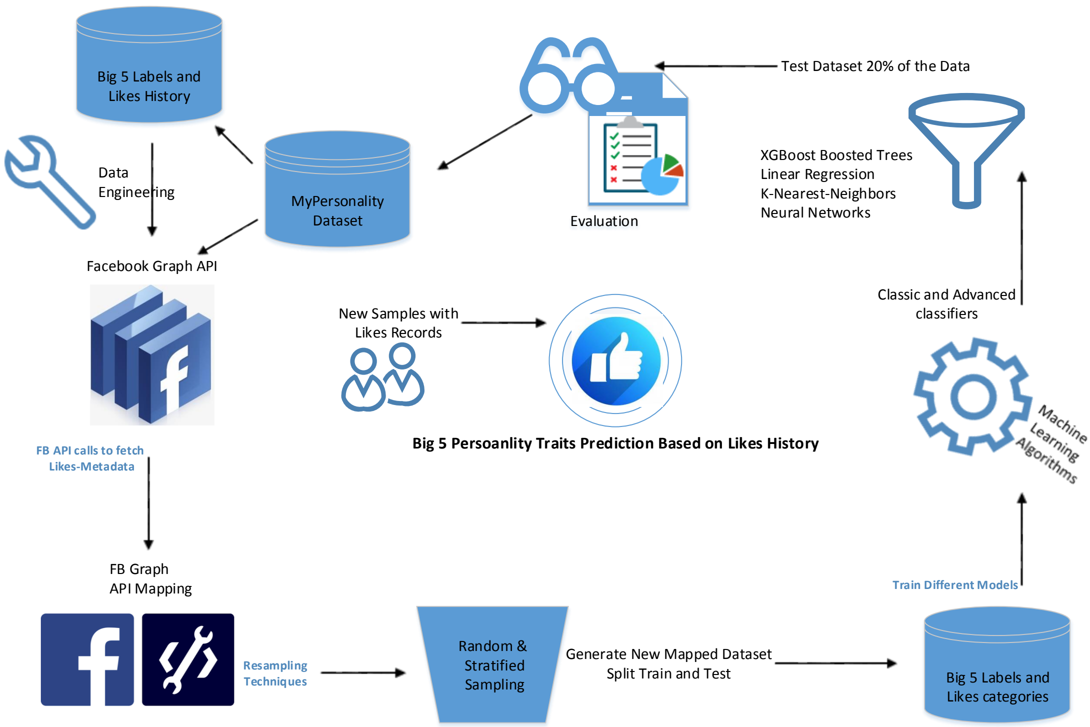



**Figure 3.1:** Personality Prediction Framework from Likes Input

### 3.3.1 Dataset

As a basis of this experiment in this chapter, we used the same datasets introduced in the previous chapter, but instead of using the records of the textual post, we used the likes history datasets from the myPersonality project [61]. This dataset includes data about the Big Five personality scores and Facebook likes of more than half a million Facebook users who donated their data for research purposes. The personality scores are described on a scale utilizing the Big Five personality model scheme. The dataset also provides information about the questionnaire size that each individual answered mapped with a unique user id. This dataset can be linked with the information afforded in the user likes dataset by mapping the user id to a Facebook-like id. However, the Facebook-like id solely does not provide any knowledge about the object, e.g., the category of the liked page. Therefore, considering the ids are the primary Facebook page ids, the missing data can be queried from the Facebook API[1]. For this reason, we developed a query bot that utilized the Facebook Graph API to get all available information about each like itself, as a category, subcategories, engagements, verification status, price ranges, and many more associated metadata. For this chapter, we utilized the category and subcategories of the like objects. Finally, we used the queried information to generate a new dataset that maps a user id to the corresponding Big Five scores and the number of likes that the user has in each category and subcategory.

### 3.3.2 Features Selection

After comprehensive research analyzing the previous work predicting individual personality traits using Facebook likes, we found that using single likes as features has already been established. However, in these experiments, none of the authors have used the metadata about the like objects as a feature to train a machine learning algorithm. Therefore, we decided to create a mapping technique from each like id to its corresponding category and subcategories using the Facebook Graph API. This new dataset can be used now after being normalized as features considering the total number of likes as the feature for my final personality traits prediction.

As classification and regression algorithms make predictions based on a quantitative scale, many features have to be treated differently than a smaller number of features. When a user has 50 likes in sports and 300 likes in education, he or she might be less interested in sports than another user who has 40 likes in sports but no other likes in any other categories. Therefore, we divided each like count by the number of total likes for this user to turn the absolute scale into a relative scale. A value of 0.1 in sports means that 10% of his likes are from the category sports. The final results section presented in the final results section, normalizing the features improved our RMSE values by about 5%.

### 3.3.3 Classification

As the goal is to predict numerous outputs on a linear scale, we have to convert it to a categorical scale first. Therefore, we used "buckets" for each Big Five category, holding related user objects. we started by learning a classifier that gets a search query as input and predicts for every page if it is irrelevant in relation to the query. The most simplistic classifier would estimate the most frequent value and always return the result for every page. The algorithm will have a recall of 1, and since most pages are indeed irrelevant, the algorithm will also have a precision of near 1. Despite the most straightforward algorithm being inefficient in terms of finding out whether a page is relevant or not, the result of other learning algorithms should therefore be examined against the simplest classifier

[1] https://developers.facebook.com/docs/graph-api

and evaluated if the sampling method or the cost function resulted in an underfitting of minority categories. Since the simplest algorithm constantly returns the same value for each element, it cannot be used to distinguish between users or make meaningful decisions based on the algorithm's output. Therefore, the output of every learning algorithm should be validated using precision/recall, f1 measure, mean-squared-error, root-mean-squared-error, or similar metrics and how resembling it gets to the simplest classifier that always just returns the most frequent value.

We investigated the Random Forest algorithm as a classification algorithm based on many decision trees. In order to create a decision tree, the data should be mapped into an n-dimensional space where each dimension stands for one like category. The algorithm then finds a decision boundary and divides the dataset into two non-overlapping partitions. This is continued until a remaining group can be perfectly separated, resulting in tiny buckets and overfitting having a very low bias but a high variance. To moderate overfitting, stopping criteria have to be introduced, e.g., a maximum tree size assigning the majority value as a result to a bucket (tree pruning). Another mechanism to mitigate overfitting is to generate many trees and average over these while forcing each split to only consider a subset of the predictors, resulting in de-correlated trees as random forest.

### 3.3.4 Regression

Linear regression is a learning technique that represents the relationship between a numeric variable y and one or more feature variables. In our case, the produced features are the number of likes a user has in a particular category. Let's suppose we want to predict the personality, specifically the openness trait of a given user, based on the feature number of likes in Boxing Studio (numLikesBoxingStudio). In this instance, we attempt to discover the optimal values for $\theta_{0,1}$ so that the following function is optimal:

$$\theta_0 + \theta_1 * \text{numLikesBoxingStudio} = \text{OpennessScore}$$

These links are found by linear regression and modeled by the predicting function. An illustrative example is visualized in Figure 3.2 inspired by examples mentioned in [61].

We investigated the Boosted Trees algorithm for this task. Boosted trees can be used in both regression and classification tasks. Boosted trees are pretty similar to the Random forest classifier we explained before. To utilize random forests for regression tasks, each leaf must be labeled with a scalar output instead of a category output. The output, in that case, is not a linear function as in linear regression but can be compared to a grid or stairs because there are rigid decision boundaries. Nevertheless, boosted trees are an additive technique establishing what is already learned and adding one new tree at a time.

We also applied the k-nearest neighbors algorithm, which draws each sample into an n-dimensional space where n is the number of various like categories and labels the data points with their personality score in each category. The data point that is used

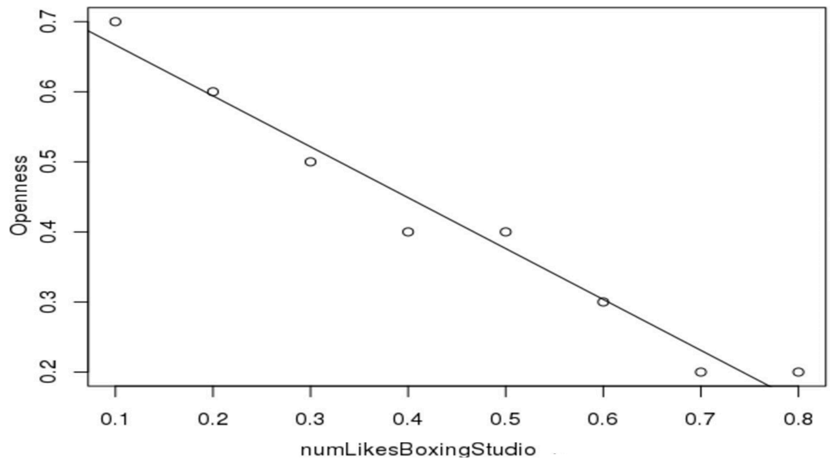



**Figure 3.2:** Relationship between a Feature and Openness to Experience Personality Trait Represented by Linear Regression: An Example [17] ©2019 IEEE.

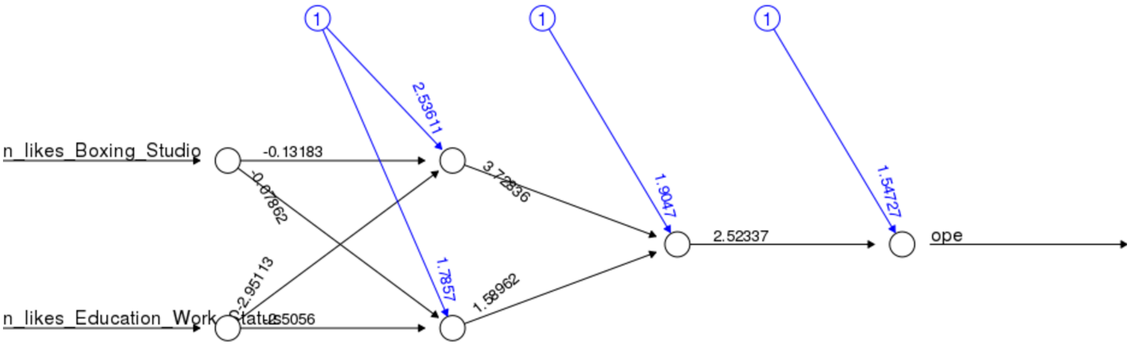
for prediction is then also outlined into the n-dimensional space and gets the average personality score of his k-nearest neighbors. For k = 1 the algorithm would always return the same score as the most equal person from the training set. For k = N (where N is the number of training users), the algorithm would always return the average value for each category because all people from the training set are neighbors.

Neural Nets Neural nets are built on the concept of a perceptron. A perceptron is a particular node that receives input variables, employs a function, and generates an output value. In general, Linear regression can be modeled using a single perceptron. Therefore, a neural net can be modeled as a network of perceptrons in which the first layer takes the features as input, and the following inner layers get the output of the previous perceptrons as input. An example is shown in Figure 3.3 where we show a neural network to predict openness trait based on two input features. In each node, the logistic function is employed to each input value with a weight to assemble the output value for the next node.

To make the result of our prediction from Facebook-likes accessible, we established a prototype on which users can predict their own personality using their Facebook Likes data. The prototype is divided into two parts, a Python backend that performs prediction tasks on given like information and a NodeJS frontend that serves the webpage and queries the like information of the user. A random user personality prediction from the dataset using the likes is shown in Figure 3.4. The figure represents the personality prediction using the first 100 likes a user has compared to 250 likes.

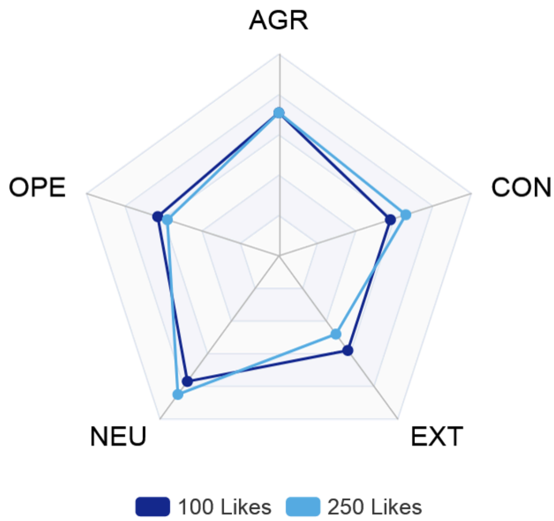
Initially, when opening the webpage, the user is faced with an explanation text about the product and a Facebook login button. The login button triggers a permission request for the assigned Facebook app and returns an access token if the user accepts to login

**Figure 3.3:** Openness to Experience Personality Trait Prediction using Neural Network with Two Input Features:
$(\mathrm{num}_{L}ikesBoxingStudios)featureand(num_{L}ikesEducationWorkStatus)feature.$




and provides the user likes permissions. The access token is then sent to the NodeJS backend, where it is used to query all likes of the user and convert the like ids into the matching categories. It then aggregates the data and sends it to the Python backend, which calculates the Big Five scores. Once the scores are returned, they are sent back to the webpage, which displays a radar chart using ChartJS3. It also generates a text describing the user personality using different text blocks4 based on a threshold for each dimension.

**Figure 3.4:** Radar Chart Representing the Big Five Personality Traits for a User using his/her Likes History (100 - 250 Likes).



## 3.4 Experiments Results

### 3.4.1 Choosing the Most Appropriate Class of Algorithms

Different groups of algorithms shine for different tasks. The first step of deciding which algorithm to use is to define what exactly we want to predict. In this case, there are two different ways we could consider how to predict the five personality traits of a user: we could define categories to which we assign users, or we could calculate a score for each user. A categorization could be in the form of defining ranges, e.g., low", medium", and high", for each score. A more sophisticated categorization technique could be in the form of personality profiles, e.g., " The Athlete", " Intellectual", or " The Reliable". However, the data set did not include any of such information, and we could not find a data set of sufficient size to pursue that idea.

The second way of predicting personality traits from likes is to strictly follow the Big Five personality model and predict a score for each personality trait. To decide which group of algorithms fits our problem better, we trained two models: Random Forest", an algorithm typically used for classification, and Linear Regression", which is used to predict continuous values. Early on, regression models outperformed classification models. We achieved precision values of about 40% with an optimized set of features for Random Forests", but the task at hand is not a classification task. A classifier that just picked a class randomly would achieve a precision of 33% because three groups were equal in size, making it only slightly worse at predicting labels than our model. Additionally, the class labels do not provide much information about the person itself, a score would provide much more detail. Therefore, we decided to focus solely on algorithms that predict continuous values.

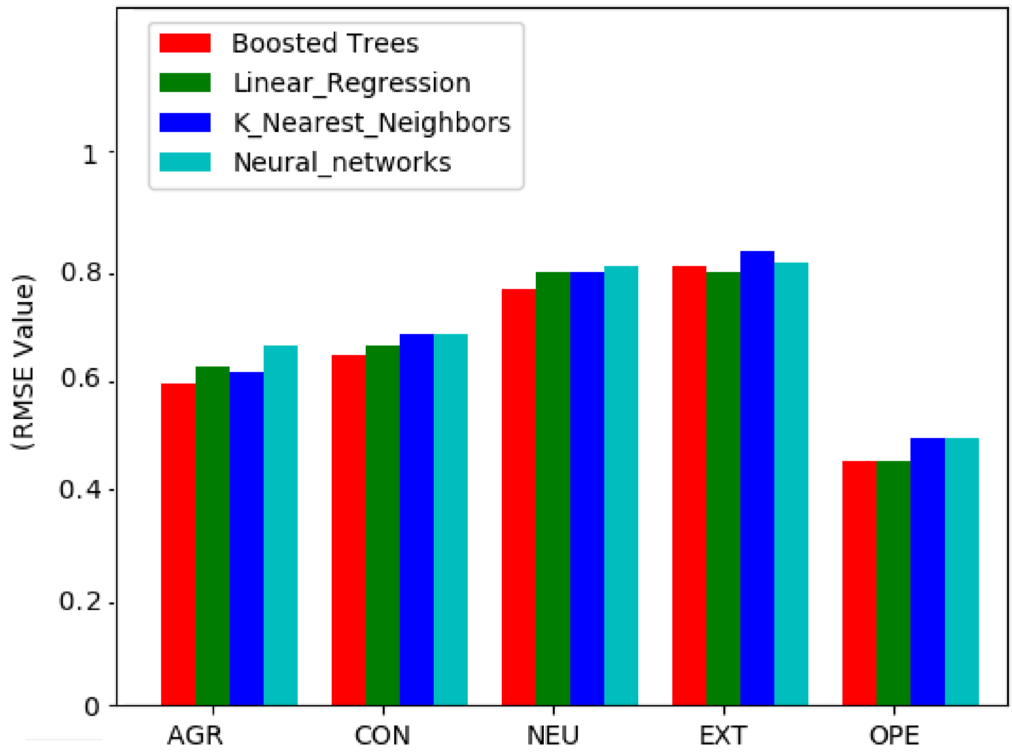
### 3.4.2 Advantages of Simple Algorithms

There are a variety of models that can be used to predict continuous values. This research evaluated four different algorithms: linear regression, k-nearest neighbors, regression trees, and neural networks. A summary of the RMSE values of the best models for each algorithm can be seen in Figure 3.5.

All metrics indicate that the best-performing algorithm of the four is boosted trees, followed by linear regression. Linear regression is a fairly simple, and very well-understood mathematical model. The advantage is that it can investigate where the model has its weaknesses. In turn, these weaknesses can be addressed, e.g., by sampling differently or spending more time on feature engineering tailored to the deficits of the predictions. This is a considerable difference to algorithms much harder to comprehend, e.g., neural networks. While these algorithms offer much flexibility in their tasks, they sometimes introduce unnecessary complexity. In this case, we deal with a rather traditional regression task. The linear regression models are capable of high-quality predictions and perform even better with the feature set we optimized for. Since the output is linear, it is not limited to the logical boundaries of the setting where output is between one to five. Therefore, we clipped the output to have all values above five to be rounded down to five and all values below one to be rounded up to one.

The algorithm we used for boosted trees is called “xgboost”[2]. In recent years, it has increased popularity by being the winning algorithm in several machine learning competitions. Boosted trees strive to address the model’s weaknesses during the training stage automatically. In a given step, it calculates which training data it has difficulty predicting and then produces a tree exceptionally trained to predict those remarks adequately. It continuously calculates a reasonably sophisticated model that consists of multiple smaller trees, each optimized to predict specific properties of the training set. It is therefore eminent to have a representative training set.

**Figure 3.5:** Evaluation using RMSE Metric for ( Boosted Trees, Linera Regression, K Nearest Neighbors and Neural Network) for Predicting the Big Five Personality Traits Using Likes Records as Inputs ©2019 IEEE.




[2]https://github.com/dmlc/xgboost

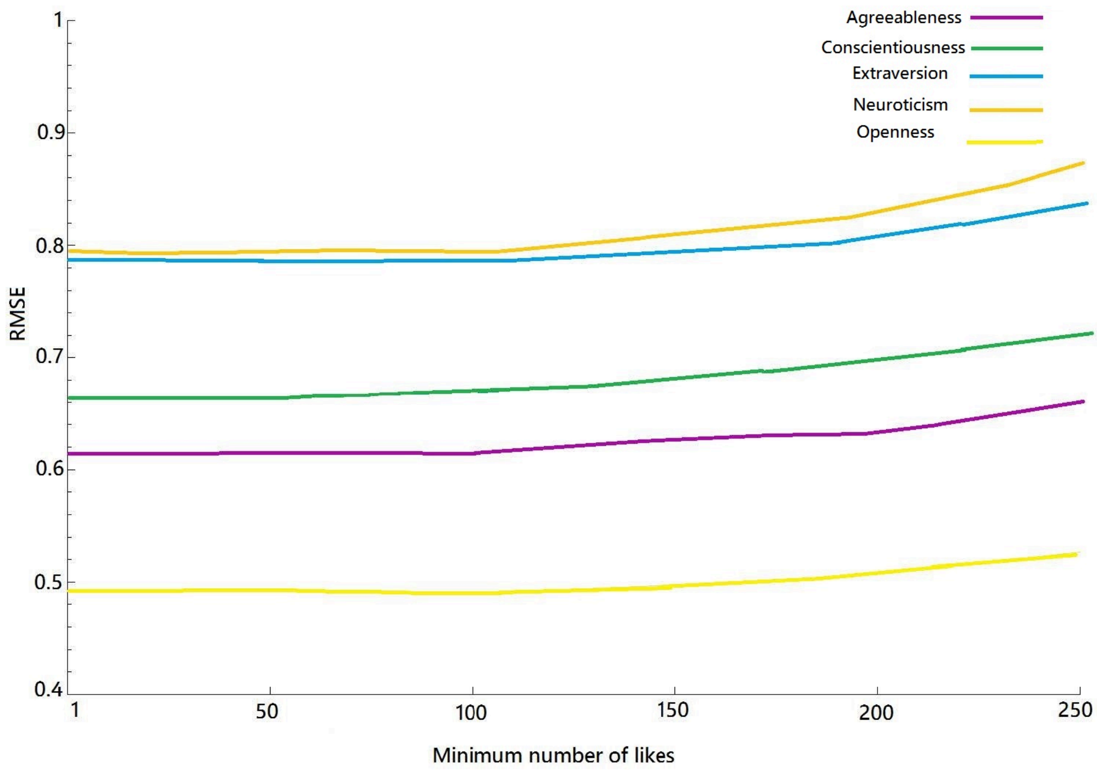
The k-nearest neighbor model we developed could compete with linear regression but could not accomplish similar metrics. We examined different numbers for k, and found 10 <= k <= 15 to yield the best results. The k-nearest neighbor algorithm is distance-based, and we also used different functions to penalize different sorts of distance, e.g., within and across categories. Based on several trials with different forms of penalties, we decided to count the number of non-overlapping categories the users have likes and use this as the basis for the penalty. Consequently, the more categories in which the observation has likes, but the neighbor does not, the higher the penalty. One way to further improve the model is to analyze the importance of each feature and significantly reduce the number of features to the ones that are most impactful on prediction performance. Also, users might be penalized heavily for having similar likes, which Facebook assigns to different subcategories. Generalizing these might help improve prediction performance because distance-based metrics would put them closer together. The fourth algorithm we trained belongs to the group of neural networks. While we are able to get decent results for the predictions, we could not statistically improve the model. Today's research still struggles to understand the inner workings of neural networks, which makes feature engineering a lot harder. In summary, the results show that we could predict a user's Big Five openness score within about 8% on average with the boosted trees algorithm. The openness" personality trait can be predicted with a minor error, indicating that it correlates most with the pages that a user liked on Facebook. This allows for a quite accurate assessment of a person's personality traits and can be used in a wide variety of fields, e.g., in political or marketing campaigns.

### 3.4.3 The More Likes, The Better

Typically, the more data points available for the algorithm to be learned, the better the algorithms are at predicting target variables. Since we use each like of a user and create features out of them, the more likes a user has, the better we expect to be the final personality predictions. On the other hand, filtering the data for a minimum number of likes significantly reduces the number of observations that we can work with. In fact, the data set is reduced by about 75% when specifying a minimum number of 250 likes for all users. Consequently, a much smaller part of the data is available to train the models, which in turn negatively influences prediction performance, as shown in Figure 3.6.

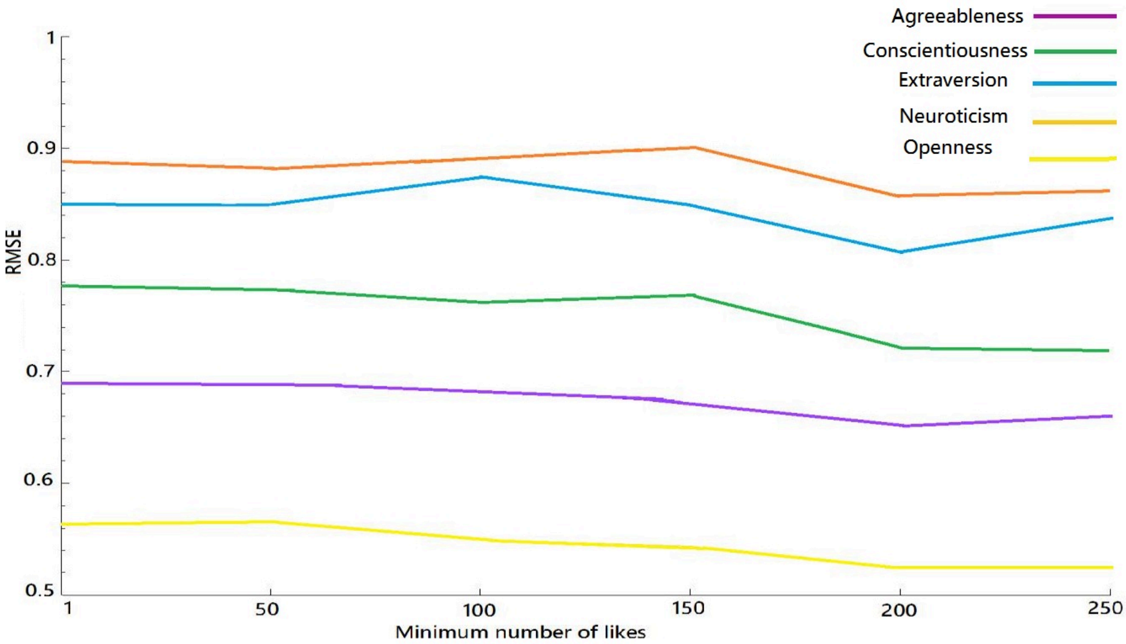
Another round of experiments confirms this hypothesis. In this case, we used a fixed-size training set in every round. Therefore, in contrast to the previous experiments, we trained the model on the same number of observations, no matter how many observations exist in the whole data set with the respective number of likes. In each round, the data set consists of close to 14,000 observations - the number of observations with at least 250 likes. If we had more than 14,000 observations available -as is the case for rounds in which the minimum number of likes is smaller than 250 - we randomly choose 14,000 observations

**Figure 3.6:** Evaluation using RMSE Metric for Different Minimum Number of Likes with the Maximum Size of Training Set [17] ©2019 IEEE.



and disregard the others. We can see that prediction performance, in that case, does improve when filtering only for users with a higher number of likes, as shown in Figure 3.7.

**Figure 3.7:** Evaluation using RMSE Metric for Different Minimum Number of Likes with the Fixed Size of Training Set [17] ©2019 IEEE.




### 3.4.4 The Influence of Stratified Sampling

Stratified sampling defines a method of producing the training set for an algorithm to learn from the available data points. In stratified sampling, the first step is defining some groups that can typically be considered buckets for the target variable. Every observation is assigned to exactly one group. There are different methods to sample from those groups, of which we experimented with two different ones.

The First is called Equal Allocation and means that the number of samples drawn from each group is the same, and it does not depend on the number of observations in the group. Using Equal Allocation might help predict users with very low or high scores in a personality trait more accurately. These scores are less common in our case, and therefore there exist fewer users with those scores in the training set. If random sampling was used, the algorithm had fewer data to train the model to predict scores at either end of the scale. By using Equal Allocation instead, we tried to even out the distribution and help the algorithm learn to predict more extreme scores. Unfortunately, our training data was not well-equipped to support this kind of sampling. The number of observations with very low or high scores was not sufficiently high to train a model that could come close to the prediction performance we reached when using simple random sampling.

The second method is called Proportional Allocation and represents the process of using a sample size for each group individually, depending on the size of the group. The number of observations drawn from each group depends on how large the group's share in the whole population of all observations is. The observations are simply chosen randomly from each group, e.g., every observation in the group has the same chance of being drawn. Additionally, all sets of observations with the same number of samples are equally likely. If a group has only very few observations, the training set might not include it if only simple random sampling is used. The advantage of Proportional Allocation compared to simple random sampling is that we can ensure that every group is represented in the training set proportionally to how it is represented in the whole population. As a result, Proportional Allocation provides equal or better precision than random sampling. However, in these experiments, we could not measure a significant difference in prediction performance when utilizing proportional allocation.

## 3.5 Chapter Summary

Although the data associated with the metadata of the like objects is a small proportion to the total information, a user gives on social networks, the learning models that we developed confirmed that they can correctly estimate an individual's personality. Ultimately, the Like-based trained models can now be joined and integrated into a knowledge ensemble method that considers other relevant information such as posts to build personality prediction systems better. The results validate the significance of the correlations between users' personality and their Facebook Likes history and categories. Therefore, predicting users' characteristics and preferences can be used to enhance numerous product services and reveal new opportunities for personalizing interfaces. We used the hierarchy that Facebook uses to categorize pages as features to train the models and estimate users' Big Five scores for each personality trait. While the performance varies between the traits, the results show that we can predict the personality score within only 8-15% of the actual value. It is also shown that the more data we have, the more accurate the final models will be. The Likes features with its metadata information can now be combined with the models of other features types to create a learning ensemble that can predict the personality of a social media user to an acceptable approximation.

# 4 Emotions Prediction Using Facebook Reactions and Emojis Buttons

## 4.1 Introduction

The Tremendous use of social media platforms has afforded unique approaches for people to show their opinion and sentiment online. Facebook launched the features of the new reaction buttons in 2016 to let users express and show their psychological responses to online content. Using these five reaction buttons, we built a framework for predicting the distribution of Facebook post reactions by analyzing the textual inputs that are publicly available from public Facebook pages. For this task, we created a unique data crawler to collect and gather an enormous amount of Facebook posts associated with their reaction tags using the proposed scalable Facebook crawler as described in detail in the data collection section. We collected from 64,000 public unique Facebook pages from various categories for the machine learning phase, resulting in more than 3 million labeled Facebook posts associated with their reaction records. The evaluation at different standard benchmarks using the proposed extracted features manifests encouraging results compared to previously published research.

Using the Facebook platform, users share their stories and ideas and show their support or endorsement by "liking" posts of other network members. The users apply the newly introduced feature to express specific emotions in response to content. Previously, users had only one type of choice to like a post or not. Nowadays, Facebook introduced a new set of dynamic features called Facebook reactions such as Like, Love, Care, Haha, Wow, Sad and Angry, allowing users to choose between seven different types of interactions. Thus, reactions are more meaningful compared to direct simple like as presented in Figure 4.1. The Care reaction is not considered in this research because it is still relatively new, and not enough labeled training data can be retrieved from the API or the proposed crawler.

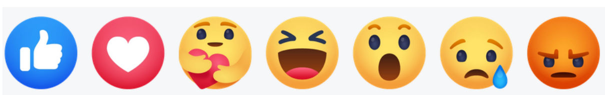


**Figure 4.1:** Extension of Facebook Like Button: Seven Different Facebook Reactions Button ©2018 IEEE.

Extracted social media streams draw many researchers' attention from different domains. For this reason, we collected public pages containing a set of posts, including how people reacted and sensed while exploring their feeds. This presents us the opportunity to develop machine learning models that can predict the reactions of a post, given merely its plain text. Consequently, we decided to create an automatic emotion detection system by utilizing Facebook reactions as representatives for emotion labels trained and evaluated on well-known benchmarks in the field of emotion recognition.

This chapter is constructed as follows. Section 4.2 introduced the previous work in this domain. Section 4.3 clarified how to pick Facebook pages from various Facebook page categories to assemble and train an unbiased universal prediction model. The exact section delivered insights into comprehensively retrieving a massive dataset from public APIs and outlines the difficulties that emerge when experimenting with an enormous pool of heterogeneous data. Section 4.3 also explained how features were extracted, examined, and selected based on their final prediction performance. We also presented the methods we applied to evaluate the generated models against different datasets, applied different filtering criteria, and evaluated the selection of final models' features. Finally, section 4.4 compared the results of our models against other models from previously published research and concluded with redirection for future work.

## 4.2 Previous Work

As Facebook Reactions were first released in 2016, experimentation on them and their influence is still limited and further investigated. Pool et al. [69] proposed a system that can identify emotions by analyzing the retrieved data from Facebook posts. It begins by describing how they selected Facebook pages from different categories, mainly based on their intuition. Almost 1,000 posts were collected from each Facebook page with their associated reactions. They excluded Normal Likes and only collected the reactions of "love", "haha", "surprise", "sadness" or "anger". The authors used standard textual features like the tf-idf, bag of words, character and word n-grams, as well as the presence of negation words and the punctuation in the published posts. Models evaluation was done against three different datasets annotated with emotions manually throughout domain experts.

Chaffar et al. [70] utilized a mixed dataset obtained from blogs, fairy tales and news headlines to train machine learning classifiers for automatic emotion recognition from textual posts. They examined the model's performance when it is being trained independently over homogenous datasets. They explained the significant statistical improvement when they used the SMO classification algorithm. Gray et al. [71] used the Facebook reaction data for examining bias in Facebook Pages of United States 2020 Senate candidates using Facebook Reactions. Felbo et al. [72] trained a model based on 1,246 million Twitter posts that include one of 64 predefined Emojis. Their dataset included 56.6 billion tweets and was filtered for language and URLs. They pointed out that enlarging the set of noisy labels increases the performance of emotion detection and sentiment analysis.

Another framework for predicting the distribution of Facebook post reactions was presented by Krebset al. [73]. The framework was trained on a customer service dataset from several supermarkets' Facebook posts. The study confirmed that a baseline sentiment miner could be used to identify a post sentiment/emotion. Subsequently, the results can be merged with the output of the neural network models to predict the final Facebook reactions. Basile al. [74] created a simple regression model to estimate the entropy of a post's reactions based on the Facebook reaction feature. The proposed measure is a proxy to predict the news controversy, where the higher the entropy (indicated by highly mixed reactions), the bigger the controversy. They run trials both within and across communities, explained by the Facebook pages of particular newspapers. In almost all cases, the trained model beats the baseline in cross-validation of the same source data. The strategy they have emerged is based on discrete linguistically motivated features. This has shown a clear influence on the built model as it cannot generalize enough when dealing with low-frequency features and unseen new data in the testing dataset.

Pool and Nissim (P+N) did the pioneering work on using Facebook posts and the users' reactions for emotion recognition task Pool et al. [69]. They presented three different models, one for each evaluation dataset: The Best-Model (B-M) is based on Facebook posts from *Time*, *Disney* and *The Guardian* and was built to perform best on the Affective Text dataset. For the Fairy Tails dataset, they selected the pages *ESPN*, *CNN* and *HuffPostWeirdNews* and created the FT-M. A combination of *Time*, *CookingLight* and *The Guardian* builds the foundation for the ISE-M, a model tailored to the ISEAR dataset.

Based on the Affective Text dataset introduced 4.3.2, Strapparava and Mihalcea published [75]. In addition, they presented five systems for emotion detection task in [76]. Their model (LSE all emotion words) yields the best results for the recall and f-score and is included in our evaluation as (S+M). They experimented with automatizing the analysis of emotions in textual inputs. They explained the development of a large data set labeled for six basic emotions: anger, disgust, fear, joy, sadness, and surprise. As mentioned earlier in the text, they introduced and assessed various knowledge-based and corpus-based approaches for automatically classifying emotions.

Four different models are proposed by Kim et al. [77] to predict a particular emotional label based on textual inputs implemented in Matlab. A Matlab toolkit was adopted to produce the term-by-sentence matrix from the text. They used the kappa score as an unbiased reliability metric for comparing the four methods. The results can be generalized of which of these methods is the best performer because results fluctuate among different datasets. The authors further examine this association to recognize more powerful strategies suitable to a generic universal dataset. They are still examining the advancements in the used methodology proposing a combination of the method employing emotional modeling and empirical methods. We included the results of their (CATEGORIAL NMF MODEL), as it turned out to perform the best in their evaluation. They also evaluated on the Affective Text, ISEAR and Fairy Tails datasets but limited their set of labels to "anger", "joy", "fear" and "sadness".

## 4.3 Methodology

The proposed emotions and reactions prediction framework is shown in Figure 4.2. The flow of the work starts by introducing the scalable crawler. As an initial seed, we assign any random public Facebook page, and then the proposed crawler will be able to assign multiple jobs to collect all other publicly available pages found on the first page. All data is stored and filtered based on several proposed filtering criteria, as mentioned in the data filtering section. Later, different features are extracted and examined, such as (tf-idf, google embeddings, count and punctuation) to train multinomial Naive Bayes Classifier to predict the emotions for any given textual post. The evaluation is done against top published datasets considered to be the main benchmarks in emotion prediction tasks from textual inputs.

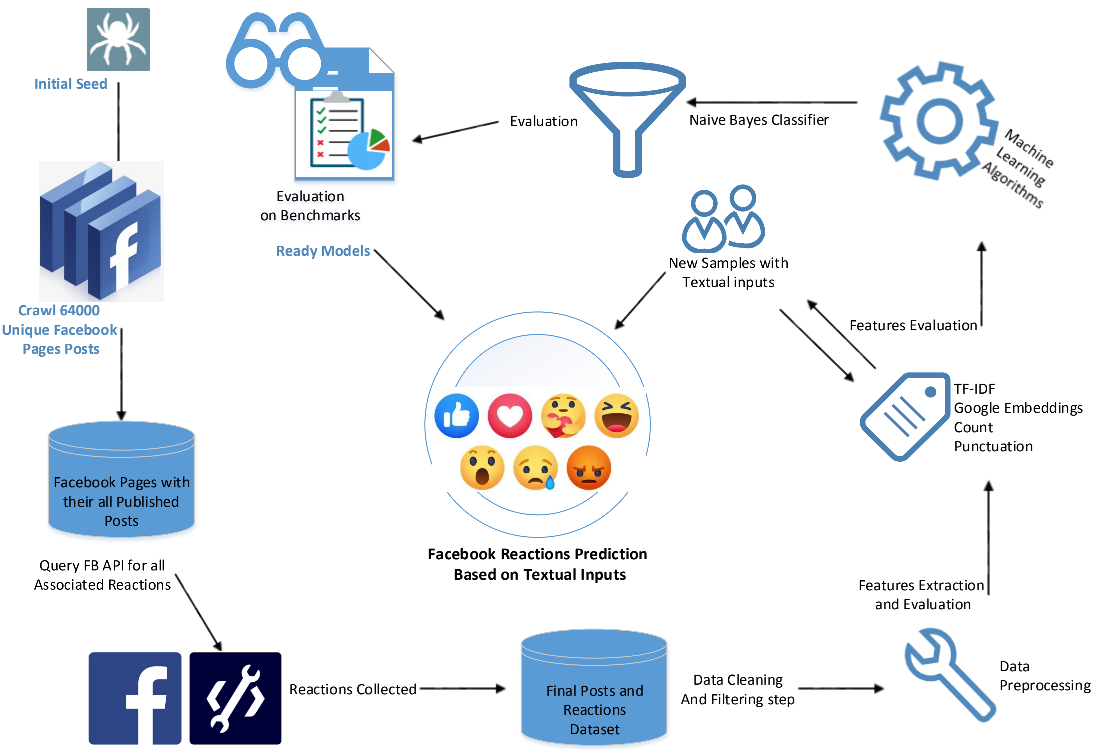



**Figure 4.2:** Emotions and Reactions prediction Framework for Facebook posts.

### 4.3.1 Data Planning

I started collecting data by crawling 1,000 posts from three different pages that share their posts publicly on the platform. This gave insights into the initial dataset and offered a starting point to formulate scripts that filter and normalize it. Simultaneously, the first set of features was chosen as described in the features selection section.

#### 4.3.1.1 Data Crawling

The first shot of training for the model was done. From this primary setup, the dataset was built continuously. After manually choosing three Facebook pages and collecting about 3,000 posts in total, including their reactions, we realized that the dataset had to be expanded because of the imbalance with labels.. The first step is started by implementing our scalable crawler that can effectively retrieve the latest 1,000 posts for each page by feeding a list of public Facebook page ids. The first crawl process showed that joyful reactions are utilized much more compared to sad, angry, or surprised ones. Considering this, a new list of categories was determined that mainly have sad, surprised or angry reactions. Figure 4.3 presents the list of categories, each attached with the most common reaction in the category. Based on the list, a subset of top Facebook pages was picked. After collecting these pages and training models based on the data, the evaluation results improved, as we showed in Table 4.1.


**Figure 4.3:** Facebook Pages Categories List Associated with the Most Common Retrieved Reactions (Sad, Surprised and Angry) [18] ©2018 IEEE.

The stated crawler automatically defines a list of ids for Facebook pages in order to collect their posts and reactions. To produce a Facebook page ids list, the crawler identifies a starting position using a set of pages from the user's history. Then a set of liked pages is requested for each page (pages having at least 10,000 users), and the algorithm begins recursively determining new pages on the liked pages set to be included as shown in Figure 4.4.

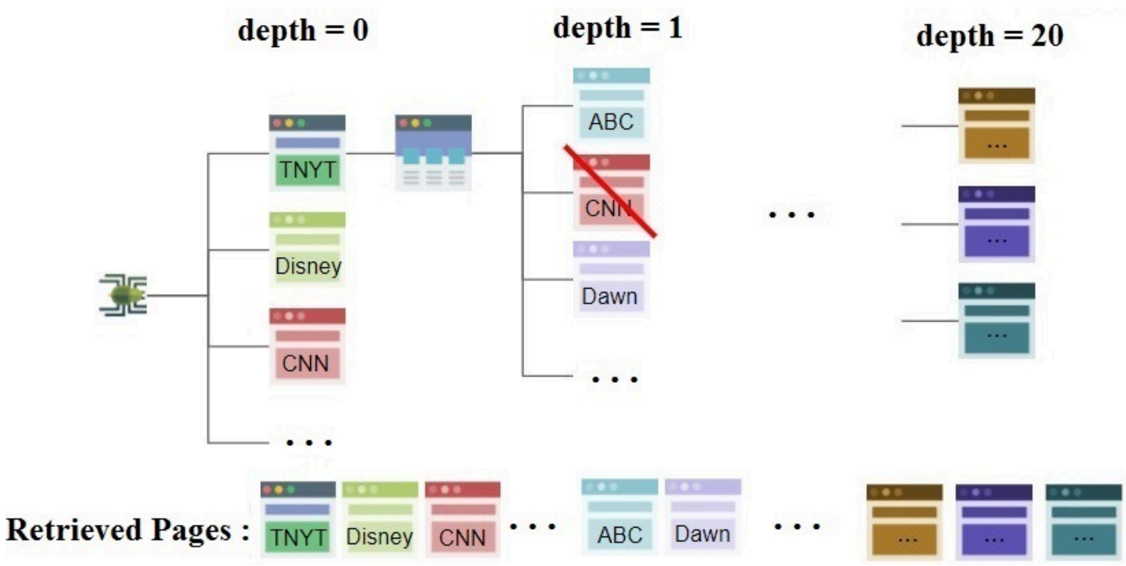



**Figure 4.4:** The Proposed Recursive Crawler for Facebook Pages (Assigned max_depth = 20) [18] ©2018 IEEE.

Executing the script with a maximum depth of n=20, a set of 64,000 unique Facebook pages was discovered. More than 3 million posts, including reactions, were retrieved handling this set of pages. A model was trained and tested using the produced dataset against state-of-the-art models, as shown in the results section. Table 4.1 we show the evaluation results for different datasets consisting of various sets of pages. Datasets were automatically retrieved by the page finder and crawled the first 20,000 posts. The Facebook Crawled dataset shows the best results in the average score and on three out of the four labels.

#### 4.3.1.2 Data Filtering and Normalizing

After collecting multiple Facebook page posts with their reactions, we applied different filtering criteria to check inconsistent data from influencing the training process. We applied some filters to the dataset before we proceeded further by removing all posts containing URLs. Removing the URLs from our samples is justified because users might have posted a reaction to the URL's content instead of the post itself. Furthermore, only posts with minimum characters and reactions are considered. Additionally, the main reaction of a post (excluding Likes) has to be unambiguous. This means that the number of reactions of the most frequently used reaction must hold more reactions than the second

**Table 4.1:** F-Score Results for the Task Emotions Prediction using Different Datasets [18] ©2018 IEEE.

| Dataset | Joy | Surprise | Sadness | Anger | Avg. f-Score |
|---|---|---|---|---|---|
| CNN | 0.43 | 0.22 | 0.32 | 0.30 | 0.343 |
| CNN + TNYT | 0.52 | 0.23 | 0.29 | 0.28 | 0.380 |
| Disney + CNN + TNYT + GP | 0.55 | **0.26** | 0.29 | 0.26 | 0.407 |
| Facebook Crawl | **0.63** | 0.20 | **0.45** | **0.44** | **0.509** |

most common reaction. Table 4.2 shows the evaluation results for the datasets filtered by different top reaction gaps. Highlighted values exceed the values of other models in that category, indicating that the first model, filtered by a top reaction gap of 10%, is the best. The next filtering criteria can be optimized using this optimal value for the top reaction gap filter.

**Table 4.2:** F-Score Results for Different Top Reaction Gap Filters for the Emotions Prediction Task [18] ©2018 IEEE.

| Min. Reactions | Top Gap | Joy | Surprise | Sadness | Anger | Avg. f-Score | # of Postss |
|---|---|---|---|---|---|---|---|
| 25 | 10% | **0.60** | **0.16** | **0.35** | **0.41** | **0.458** | 532732 |
| 25 | 25% | **0.60** | 0.12 | 0.31 | **0.41** | 0.449 | 511708 |
| 25 | 50% | 0.57 | 0.12 | 0.31 | **0.41** | 0.435 | 461308 |

Table 4.3 shows the evaluation results for the dataset filtered by different minimal reaction counts. The table also examines the results of the differently filtered training sets against the results of Pool et al. [69]. For instance, the best model is the one filtered with the lowest minimal reaction count and the smallest top reaction gap. A likely reason could be that filtering the dataset with higher thresholds decreases the number of total posts which are left for training.

The final filtered generated Facebook dataset and the evaluation datasets (the benchmarks) had different structures. For this challenge, the labels had to be normalized first before proceeding with the training phase. We normalized the main reaction of a post assigned to it. Furthermore, the main reactions were also mapped to their corresponding emotions using the same recommended mapping technique proposed by [69]. Table 4.4 shows the adjusted mapping technique from reactions to emotions.

**Table 4.3:** F-Score Results for Different Minimal Reaction Count Filters [18] ©2018 IEEE.

| Min. Reactions | Top Gap | Joy | Surprise | Sadness | Anger | Avg. f-core | # of Posts |
|---|---|---|---|---|---|---|---|
| 10 | 10% | **0.63** | 0.20 | **0.45** | **0.44** | **0.509** | 741281 |
| 25 | 10% | 0.60 | 0.16 | 0.35 | 0.41 | 0.458 | 532732 |
| 50 | 10% | 0.60 | 0.18 | 0.30 | 0.41 | 0.444 | 400879 |
| 100 | 10% | 0.58 | 0.17 | 0.33 | 0.41 | 0.440 | 291029 |
| Pool et al. TFIDF | | 0.47 | **0.30** | 0.31 | 0.32 | 0.368 | ∼ 1000 |
| All + Google-Emb. | | 0.57 | 0.20 | 0.39 | 0.44 | 0.469 | ∼ 1000 |

#### 4.3.1.3 Features Extraction and Selection

For predicting emotion distribution for any given post on the Facebook social platform, the following features were examined:

**Count features**: The count features are extracted by the CountVectorizer provided via scikit-learn. For each word present in the text a feature is created, containing the number of occurrences for the respective word for each document.

**TF-IDF features**: Term Frequency Inverse Document Frequency feature weight the occurring words based on their inverse frequency in the documents. The often-used words will become less important, compared to rarely used ones.

**Punctuation features**: The punctuation feature takes punctuation characters used within posts into consideration as mad content creators might use exclamation marks more excessively, than a sad person, which in turn might use more dots.

**Google Embeddings features**: Based on an already pre-trained word2vec model which includes 300-dimensional vectors for 3 million words and phrases and was trained on a part of the Google News dataset. Therefore, we implemented a feature for assigning a value based on these vectors to each post.

**Multinomial Naive Bayes classifiers** were trained which are commonly used for text classification tasks. We used a multinomial naive Bayes classifier instead of the Bernoulli model, as the superiority already presented by McCallum et al [78]. Considering a multinomial naive Bayes classifier, the model is employed for document classification [79] with cases describing the existence of a word in a single document as showed below (where b =$\log p(Ck)$ $and\ wki = \log pki$) :

**Table 4.4:** Mapping of Facebook Reactions to Emotions in Different Datasets [69] ©2018 IEEE.

| Affective Text | Fairy Tales | ISEAR | Facebook | Mapped |
|---|---|---|---|---|
| Anger | Angry/Disgusted | Anger | Angry | Anger |
| Disgust | Angry/Disgusted | Disgust | - | Anger |
| Fear | Fearful | Fear | - | - |
| Joy | Happy | Joy | Haha/Love | Joy |
| Sadness | Sad | Sadness | Sad | Sadness |
| Surprise | Surprise | - | Wow | Surprise |
| - | - | - | Care | - |
| - | - | Shame | - | - |
| - | - | Guilt | - | - |

$$\log p(C_k \mid \mathbf{x}) \log \left( p(C_k) \prod_{i=1}^{n} {p_{ki}}^{x_i} \right) \tag{4.1}$$

$$= \log p(C_k) + \sum_{i=1}^{n} x_i \cdot \log p_{ki} \tag{4.2}$$

$$= b + \mathbf{w}_k^\top \mathbf{x} \tag{4.3}$$

We used the CNN-TNYT-GP-Disney as a training dataset as it yields better results while keeping the time frame for training and evaluation shorter than the large Facebook crawl set. All evaluation is done on the development part of the Affective Text dataset. Results are compared against the results of [69].

The results for the proposed combinations of features are presented in Table 4.5. All reported results are in f-scores and represented in micro-averaged values. As the results revealed, the combination of the Count, TF-IDF and the Google Embed-dings feature exceeds all other presented approaches. However, the TF-IDF feature on its own already performs great, especially for the “Joy” emotion, which is over-represented in the training dataset. Adding the Count feature to the TF-IDF feature could decrease the poor performance on “Surprise”, but could not fix the overall score, as “Joy” and “Anger” both slightly decreased. Joining them with the punctuation feature has neither improved nor weakened the performance.

**Table 4.5:** Performance Comparison for the used Features and their Proposed Combinations [18] ©2018 IEEE.

| **Feature** | **Joy** | **Surprise** | **Sadness** | **Anger** | avg. **f-score** |
|---|---|---|---|---|---|
| Count | 0.55 | 0.23 | 0.28 | 0.23 | 0.398 |
| TF-IDF | 0.56 | 0.13 | 0.26 | 0.27 | 0.407 |
| Count + TF-IDF | 0.55 | 0.26 | 0.29 | 0.26 | 0.407 |
| C. + T. + Punct. | 0.56 | 0.23 | 0.29 | **0.29** | 0.407 |
| C. + T. + G.-E. | **0.57** | **0.31** | **0.32** | 0.22 | **0.426** |

### 4.3.2 Datasets for Evaluation

In this section, the proposed models are evaluated and compared against previous models and former approaches, by using three different benchmarks. For instance, a short description for each dataset used for evaluation is given, followed by brief introductions about the comparative approaches. To compare our resulting model with other approaches, specifically to [69], we used the same datasets for evaluation. Before presenting and analyzing the results, the three datasets that are often used for emotion recognition tasks are briefly introduced.

In the "International Survey On Emotion Antecedents And Reactions", short ISEAR, students were asked to describe situations and their reactions, in which they felt one of the following emotions: "joy", "fear", "anger", "sadness", "disgust", "shame" and "guilt". In this survey, directed by Scherer and Wallbott in the 1990s, about 3,000 people from a total of 37 countries participated[3], which allows a transcultural view on emotions and reactions.

The proposed approach does not cover all seven emotions used in this dataset, we mapped "disgust" to "anger" and discarded entries with the labels "fear", "shame" and "guilt", considering that Facebook do not provide any reaction buttons at this time that represent such emotions. Mapping technique is shown as Table 4.4.

The Affective Text dataset was provided by Strapparava and Mihalcea [75] in the course of task 14 of the SemEval-2007 [4], an international workshop on semantic evaluations.

Two corpora are available, both contain news headlines taken from news websites, like CNN or Google News, and were annotated by six emotions "anger", "disgust", "fear", "joy", "sadness" and "surprise". They chose news headlines, as they are very likely to be

[3] http://www.affective-sciences.org/home/research/materials-and-online-research/research-material/
[4] http://web.eecs.umich.edu/ mihalcea/affectivetext/

very emotional, but are also rather short and concise. A single annotation for one headline includes a score from 0 to 100 for each emotion, to express how present an emotion is in the given sentence. The development corpus consists of 250 headlines, whereas the test set comprises 1,000 annotated headlines. Using the dataset for evaluation purposes, "disgust" was mapped to "anger" and headlines annotated with "fear" were disregarded.

Alm [80] provides a dataset containing 176 fairy tails [5], broken down to sentences and annotated with the labels "angry", "disgusted", "fearful", "happy", "neutral", "sad" and "surprised". The same mapping technique as for the Affective Text set was applied, by mapping "disgust" to "anger" and ignoring "fearful".

All of the three datasets need to be processed in order to adapt the usage of labels to our approach. The distribution of emotions in each dataset is illustrated in Figure 4.5. The ISEAR dataset now includes 4,282 entries, with about one half "anger", "joy" and "sadness" representing a quarter each. "Surprise" is not included in the dataset, as this emotion was not part of the survey. All four labels are represented in the Affective Text dataset, containing 1,056 headlines after mapping, mainly "joy". In the resulting Fairy Tails dataset, the emotions are distributed for more 5,087 sentences.

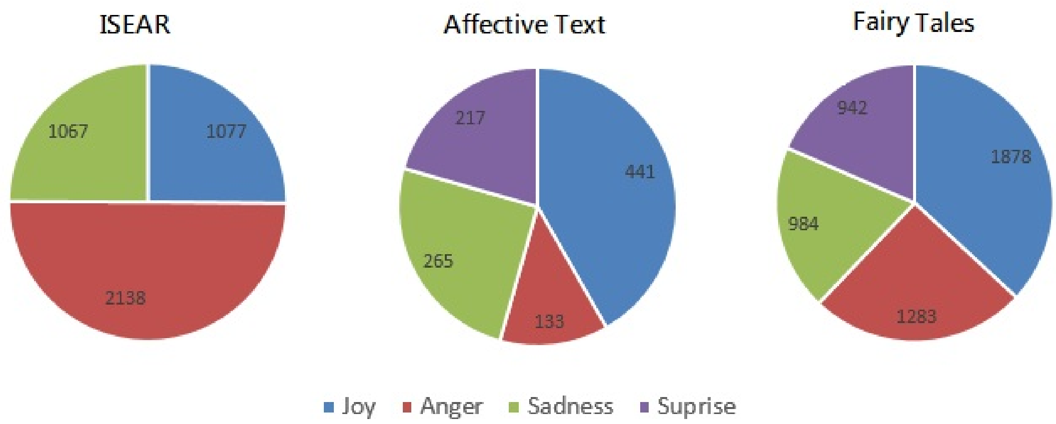



**Figure 4.5:** Emotion Distribution in the Evaluation Datasets [18] ©2018 IEEE.

### 4.3.3 Experimental Results and Performance Evaluation

In the following sections, we evaluated the performance of our model for each of the evaluation datasets introduced. Table 4.6 shows the average micro f-score for our approach, as well as for the Pool et al. [69] models, which are evaluated in the subsequent detail-sections for each dataset and should supplement the detailed values specified there. The

[5]http://people.rc.rit.edu/ coagla/affectdata/index.html

resulting numbers are from the multinomial Naive Bayes classifier, using a Count and TF-IDF feature and trained on the *Facebook Crawl* dataset.

**Table 4.6:** Evaluation of Average Micro F-Score Compared to Pool et al. [69] ©2018 IEEE.

| | **ISEAR** | **Affective Text** | **Fairy Tails** |
|---|---|---|---|
| Approach | Avg. Micro f-score | Avg. Micro f-score | Avg. Micro f-score |
| Our approach | 0.385 | **0.473** | 0.391 |
| (P+N) B-M | 0.411 | 0.409 | 0.459 |
| (P+N) FT-M | 0.336 | 0.412 | 0.408 |
| (P+N) ISE-M | **0.422** | 0.405 | **0.460** |

#### 4.3.3.1 ISEAR Evaluation

**Table 4.7:** Evaluation on the ISEAR Dataset [18] ©2018 IEEE.

| | **Joy** | **Sadness** | **Anger** |
|---|---|---|---|
| Approach | p,r,f | p,r,f | p,r,f |
| Our approach | 0.32, **0.87**, 0.47 | 0.51, **0.49**, **0.50** | **0.80**, 0.08, 0.15 |
| B-M | 0.41, 0.79, 0.53 | 0.50, 0.39, 0.44 | 0.72, 0.06, 0.11 |
| FT-M | **0.42**, 0.63, 0.50 | **0.79**, 0.28, 0.41 | 0.57, 0.10, 0.17 |
| ISE-M | 0.40, 0.83, **0.54** | 0.51, 0.38, 0.44 | 0.74, 0.06, 0.11 |
| (K et. al) | 0.39, 0.01, 0.01 | 0.37, 0.01, 0.25 | 0.41, **0.99**, **0.58** |

Table 4.7 represent the performance of our approach compared to Pool et al. [69] and Kim et al. [77] on the ISEAR dataset. For “sadness” emotion, our model outperforms the other approaches concerning recall and f-score, we reported an exceptional results in recall for “joy” sentiment, as well as a noteworthy score for the precision on “anger” emotion. As a result, all reported micro average f-scores are quite similar, even though, the model especially created for ISEAR by Pool et al. [69], ISE-M, performs best (Table 4.6) for averaged micro f-scores.

**Table 4.8:** Evaluation on the Affective Text dataset [18] ©2018 IEEE.

| | **Joy** | **Surprise** | **Sadness** | **Anger** |
|---|---|---|---|---|
| Approach | p,r,f | p,r,f | p,r,f | p,r,f |
| Our approach | 0.56, 0.73, 0.63 | 0.25, 0.16, **0.20** | **0.53**, 0.32, 0.40 | 0.33, 0.46, 0.38 |
| B-M | 0.39, 0.85, 0.54 | 0.20, 0.05, 0.08 | 0.51, 0.21, 0.30 | 0.50, 0.35, **0.41** |
| FT-M | 0.41, 0.77, 0.54 | 0.25, 0.17, **0.20** | **0.53**, 0.28, 0.37 | **0.51**, 0.30, 0.38 |
| ISE-M | 0.39, 0.82, 0.53 | **0.27**, 0.08, 0.12 | 0.49, 0.21, 0.29 | 0.48, 0.35, 0.40 |
| (S+M) | 0.19, **0.90**, 0.31 | 0.08, **0.95**, 0.14 | 0.12, **0.87**, 0.22 | 0.06, **0.88**, 0.12 |
| (K et. al) | **0.77**, 0.58, **0.65** | | 0.50, 0.45, **0.48** | 0.29, 0.26, 0.28 |

#### 4.3.3.2 Affective Text Evaluation

On the Affective Text dataset, comparing the reported average micro f-score, we outperformed the models presented by [69]. Disregarding approaches that are not based on Facebook data (Strapparava et al. and Kim et al.), our model reported the highest f-score for the labels "joy", "surprise" and "sadness", as shown in Table 4.8. Overall, Strapparava et al. reported so far the highest recall values across all labels. However, they noted a wretched performance in their reported precision. Kim et. al discloses the best f-scores for "joy" and "sadness" emotions, and produced the highest precision on "joy" tendency.

#### 4.3.3.3 Fairy Tales Evaluation

**Table 4.9:** Evaluation on the Fairy Tails Dataset [18] ©2018 IEEE.

| | **Joy** | **Surprise** | **Sadness** | **Anger** |
|---|---|---|---|---|
| Approach | p,r,f | p,r,f | p,r,f | p,r,f |
| Our approach | 0.40, **0.90**, 0.55 | **0.36**, 0.05, 0.08 | 0.34, 0.23, 0.28 | 0.62, 0.02, 0.05 |
| B-M | 0.49, 0.77, 0.60 | 0.12, 0.04, 0.06 | 0.43, 0.39, 0.41 | 0.33, 0.04, 0.07 |
| FT-M | 0.49, 0.69, 0.58 | 0.14, **0.33**, **0.19** | 0.50, 0.24, 0.33 | 0.27, 0.02, 0.04 |
| ISE-M | 0.48, 0.81, 0.60 | 0.17, 0.04, 0.07 | 0.43, 0.34, 0.38 | 0.36, 0.05, 0.08 |
| (K et. al) | **0.80**, 0.76, **0.78** | | **0.71**, **0.82**, **0.77** | **0.77**, **0.56**, **0.65** |

Based on the results reported against the Fairy Tales dataset, all Facebook based approaches perform well on “joy” and “sadness” emotion, as presented in Table 4.9. However, all approaches scores low with the other two labels, “surprise” and “anger”. Overall, the ISE-M model by Pool et al. [69] reaches the highest average micro f-score on the Fairy Tails dataset. We exceeded all previous reported models in terms of the recall results of 0.90 for “joy“ label, and presented the highest precision score for “surprise” emotion among others.

### 4.3.4 Discussion

The best performance which is reported by our model was on the Affective Text dataset. This can be abstracted due to the diverse variety of the crawled Facebook pages as well as for the feature-selection method. Besides, the Affective Text dataset holds many “joy”-annotated sentences, which serves our model and the underlying training set.

Also, all prior described models are based on Facebook posts which are often formulated in humans everyday language. Even though evaluations were executed on the same data basis, datasets used for estimating the final recall, precision and f-scores might vary from project to another. For instance, Pool et al. [69] published that distribution of emotions in the ISEAR set to be about one third “joy”, one third “anger” and one third “sadness”. But, Kim et. al [77] and our strategy, found a distribution of about one half “anger”, one quarter “joy” and one quarter “sadness”, even though we used the same label-mapping as presented by pool et al. [69]. Such inconsistencies in the final datasets will certainly influence the stated measurements.

## 4.4 Chapter Summary

We presented a novel framework for predicting post emotions and reactions distribution within the Facebook social media platform. We had investigated the potential of using Facebook reactions to identify and distinguish emotions. The proposed framework is trained on a dataset collected using the presented scalable Facebook posts crawler. While there has been plenty of research on sentiment interpretation in general, emotion classification is still primarily undiscovered, and this work contributes to this direction and aims to create a universal model using a single training set. The crawled dataset is available for other researchers and can also be used as a baseline for performing further experiments[6]. We published the dataset on Harvard DataVerse which can be found using the URL below or by searching: (Bin tareaf, Raad, 2018, "ASEDS: Towards Automatic Social Emotion Detection System Using Facebook Reactions", Harvard Dataverse, V1).

[6]https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/XJN5L5

Future work involves investigating several mapping techniques for emotion labels. As described in the data normalization section, we joined the labels “haha” and “love” to “joy” labels. Intuitively, these labels are considerably comparable, but combining them under one label might moderately twist the results. Furthermore, a balanced dataset could be crawled to examine whether the performance of more under-represented labels like “surprise” and “anger” can be fixed. Nevertheless, as the collected and crawled datasets declared, leaving “Like”, “haha”, and “love” are the most commonly and frequently used reactions. Overall, hand-picking Facebook pages for emotion/sentiment-related identification tasks based on their posts and reactions do not outperform larger, ”non-curated” datasets.

# 5 Personality Prediction Using Social Media Profile Image

## 5.1 Introduction

The increased attention in identifying online users' personalities raised the questions on how to obtain individuals' personalities automatically from their public online fingerprints. The legacy and standard approaches are to complete a self-report judgment: answering a questionnaire that is employed to estimate the user's personality. For instance, a well-known NEO Personality Inventory contains 20 to 360 personality-related questions [62]. The models originated from applying factor analysis of word usage, which produces what so-called the human personality. Five main global factors abbreviated as OCEAN or FFM for the Five-Factor model describe the main traits: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Yet, questionnaires and surveys are considered time-consuming and intrusive tasks.

Many researchers have strived to predict users' personalities and interests across diverse contexts and environments in rapid and cost-effective approaches. Lately, they could accurately apprehend users' personalities from digital footprints across different social media platforms. In some cases, the prediction was more precise and accurate than the evaluation made by users' friends or family members as mentioned in [16]. The reality of users' profile images on their social media platform is motivated and explained by their psychological constructs characterized as personality traits [81]. In this research, we investigated how social media profile pictures vary based on the users' personality posting them on their social networking sites. Many studies have managed to successfully build models to predict a wide range of user private traits, these studies use different types of features to build the final prediction models, ranging from social network attributes features [82] [83], linguistic-based features from posts/tweets text [52], likes history [16] [18] [17] to profile picture preference [84] [85].

As images get widely spread, especially between younger people [86] and recent social networks are focusing on visual content such as Snapchat or Instagram, personality dimensions, in this case, can be calculated by running content analysis which is based only on images as presented in previous researches [87]. Images, in general, contain complex information in multiple variables such as scenes, compositions, colors, emotions, facial presentations, and facial expressions where these attributes can be fetched by various computer vision algorithms such as [88]. These various descriptions can be then leveraged to investigate the difference among users' personality and image uploading behavior over various types of social networking sites. In this chapter, we examined all above-mentioned image properties and show their relationships with the big five personality traits.

The following chapter is divided into multiple sub-sections. Section 5.2 listed the previous work to solve the problem of predicting personality from user images. Section 5.3 discussed the collection of data used for the training phase as well as for the testing. The same chapter also represened the process of picking up the most significant features using (LIWC) [7] (Linguistic Inquiry and Word Count tool) and Face++ tool. Section 5.4 showed the process of training various regressions and ensemble algorithms and the final quality measurements. Section 5.5 concluded and discussed limitations and future work.

## 5.2 Related Work

The predominance of social platforms triggered multiple research efforts in psychology for a further understanding of human nature. One of the richest types of data that are available on social media platforms is the profile picture. This chapter will examine the potential of using the user profile pictures as a feature for machine learning models for estimating and predicting users' personality traits. Also, we will address to which extent such a prediction is possible from solely human appearance. This part summarizes recent research efforts in predicting individuals' personalities from face images. In contrast to traditional ways to calculate users' personalities, leveraging social footprints such as images for estimating personality assure simple and fast intuition.

Liu et al. [84] performed a large-scale analysis for profile images and personality association at the Twitter microblogging platform. They used a broader range of interpretable aesthetic and facial features to statistically understand the correlations with personality in order to build accurate personality models. They pointed out that each personality dimension has a different profile picture posting behavior. For example, people who are high in conscientiousness or extraversion uses pictures with at least have one face in them, and they prefer to present energetic and positive emotions within their facial expressions. In the evaluation part, they tested the predictive performance of their used features and published relatively robust precision at the testing samples.

Skowron et al. [89] suggested a unique technique that combines multiple inputs such as texts, images, meta-features and integrates it out from two different social networking sites, which are Instagram and Twitter to better address the task of personality prediction. Using subsampling techniques, they also addressed the effects of dimensionality reduction and noise reduction in data preparation. Furthermore, they employed random forest regression models for creating a low bias and low variance model by averaging regression tree decisions.

[7] https://liwc.wpengine.com/interpreting-liwc-output/

Their outcomes presume that the collection of features and social networking sites produce fluctuations in final regression results. The superior outcomes for all personality traits are delivered by understanding engineered features that are derived from jointly social networking sites.

Nie et al. [90] also addressed the problem of how to estimate social user personality by only using users' uploaded profile pictures. They presented personality as reverberating individual behaviors on a specific social platform. They divided the used sample into different data sets and then labeled them with various personality dimensions by a clustering approach. They introduced low-level features to train and estimate personality scores out of personal photos. The final tests confirm the validation of using such a method. They also highlighted the importance of features refinement and features design as well as the necessity of enlarging the used samples for training.

Cristani et al. [91] examined the growing size of multimedia information that users generate and engage in an online manner and consider it as a probable contributing factor to what so-called online appearance. Their work summarized tests on the interrelation among users' personalities and Flickr images. The investigation draws attention to new challenges in this domain as detecting visual patterns that get together with personality traits are harder to harvest compared to regular linguistic features extracted from textual inputs. The paper also confers that visual patterns correlate with personality score and can be used to predict personality where also they observed that the favorite images users assign in his/her profile can be used eventually to build prediction models to estimate their preserved online personality.

The recent advancement in computational personality research is about capturing online traits from audio and video. A multi-modal personality prediction system by Suman et al. [92] was proposed in 2022 and published in the Knowledge-Based Systems journal. The authors proposed automatic personality prediction visual and audio content by extracting images features from the videos and feeding it to a deep neural network setting to create the final prediction. They released the used dataset in Chalearn-16 and managed to combine both the extracted audio and visual features to create personality models and they showed that better performance could be achieved by using a subset of images from video compared to using all images from video.

Some researchers analyzed the relationship between personality traits and filters applied to images on Instagram, the popular social networking site. Ferwerda et al. [87] studied the effect that users' personalities have on using the available filters proposed on the Instagram platform. On Instagram, users can apply a filter to their captured photos to create a specific representation that they want to openly and comfortably reflect for their audience. They examined the suggested relationship by conducting an online survey where they asked their participants to fill in a personality questionnaire as well as granting them access to their Instagram account through the Instagram API. They collected data from 113 participants, resulting in more than 22k extracted Instagram pictures. They concluded that they found that distinct picture features such as hue, brightness, and saturation are highly correlated to personality traits. It can be used to expand research in this domain to ultimately create a more sophisticated recommender system based on users' personalities.

## 5.3 Implementation

### 5.3.1 Methodology

The proposed framework for estimating individual personalities from their face profile picture is shown in Figure 5.1. The workflow starts by collecting tweets from public Twitter profiles associated with their profile picture URL. The URL is then used to collect their face image to label the dataset with personality scores using our previous models that can predict personality from textual inputs. The image URL then feeds to the well-known algorithm for facial features extraction called Face++[8]. The tool can extract more than 50 unique facial features from face images that every human has. Later, these facial features are correlated with their generated personality scores from the tweets they shared. The new dataset containing users' personality scores and facial features is further studied to train a group of ensemble machine learning algorithms to predict the five personality scores. Multiple facial features with several machine learning algorithms are investigated to evaluate the best optimal feature set and algorithm for the personality prediction task from images.

### 5.3.2 Sample and Procedure

We handled two different data sets in the experiments from Facebook and Twitter. For the personality estimation task, very few trusted ground-truth datasets are available for researchers to be used. MyPersonality's app dataset [61] is considered one of the most advanced and widely used benchmarked datasets for the task personality prediction. In this research work, we used MyPersonality's dataset as the training/testing source of data to build the final personality prediction models from textual inputs, as we have previously shown in chapter two. But, for the task of estimating individuals' personalities through portrait images, we did not find any available datasets that we can use to train/test my models. Therefore, we decided to collect publicly available samples from Twitter microblogging social platform for this task.

As shown in Figure 5.2, we took into consideration the recommended ethical agreement that no information or images that could lead to the identification of study participants have been published. Public random users' full tweets history got collected associated with the URL of their profile picture on Twitter. The collected dataset was not labeled and

[8]https://www.faceplusplus.com/face-detection/

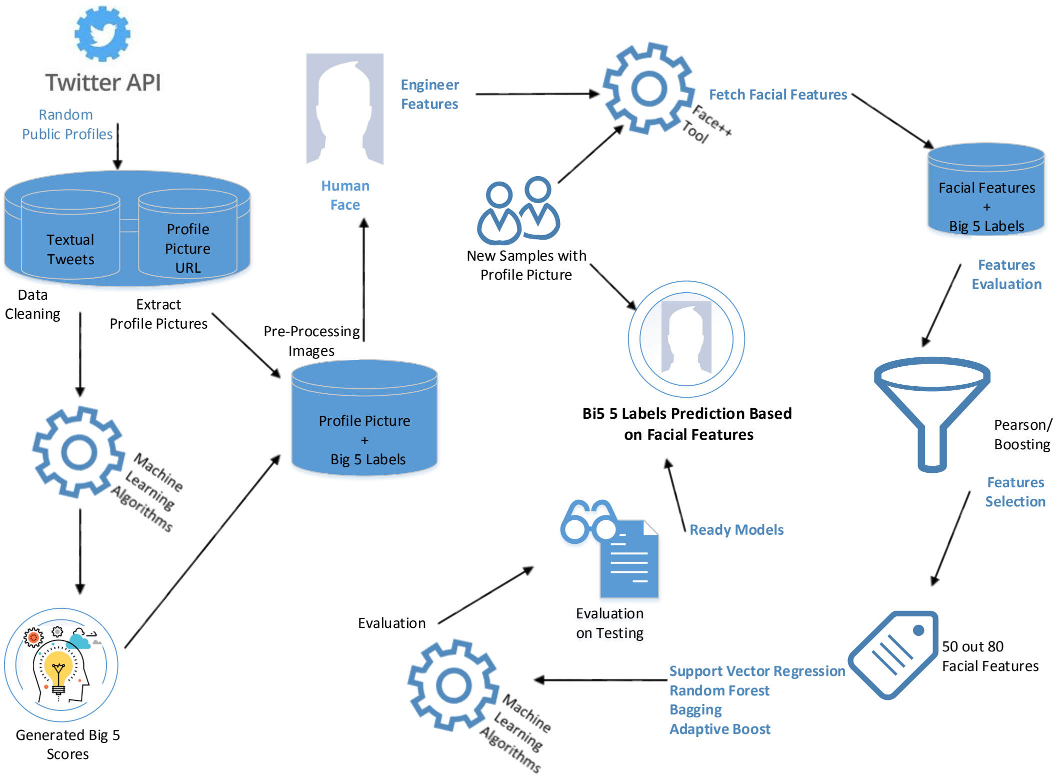



**Figure 5.1:** Personality Prediction Framework from Profile Pictures as Input.

**Figure 5.2:** Example of Profile Pictures Available in the Training Dataset. Stack Blur Filter with Radius 7.

can not be leveraged to let different algorithms learn to predict personality from pictures. So we used the Facebook self-reported samples to build personality prediction models as shown in chapter two to predict personality from texts for the collected Twitter samples. After that, we concatenated the predicted Twitter personality scores with their previously collected profile picture URL. We explained the validity of the following learning approach in-depth and demonstrate how we generated personality scores from text to image samples.

#### 5.3.2.1 Facebook Dataset

Mypersonality dataset is a famous Facebook application created by [61] back in 2007. At that time, the application let participants take part in diverse psychometric evaluations such as the Five-Factor questionnaire as NEO Personality Inventory from [62]. Almost 30 percent of the participants who participated in the questionnaire were allowed to share their data with the application. MyPersonality application dataset contains almost six million psychometric questionnaire results for four million unique participants. In this chapter, we again used the three main datasets from the application dataset, which are:

*Demographic Details Table:* It outlines the demographic information of four million Facebook participants with their Facebook ID, gender, age, relationship status, events attended, Interested In information, language, number of friends and more additional relevant information.

*BIG5 Personality Scores:* Includes the personality scores ranging from $[1-5]$ with the information about the survey length each associate took.

*Facebook Status Updates:* It is a table that holds the aggregated status updates for the participants who participated in the application. It presents a twenty-five million status update for all participants.

The considered dataset also has the (Linguistic Inquiry and Word Count) labels for almost 153.617 Facebook participant as shown in Table 2.1. The features scores were calculated by performing the (Linguistic Inquiry and Word Count) tool [63] at the participants' level. The (Linguistic Inquiry and Word Count) analysis tool determines the varying emotions and thinking styles or social concerns in any given text and returns output in many desired formats. Each single annotation which is also referred to as word categories is expressed as a percentage of words for all individuals status posts.

Comparative analysis of feature selection algorithms for computational personality prediction Facebook is done. Multiple feature extraction methods (pairwise Pearson product-moment and gradient boosting trees) have been evaluated to obtain the best correlated and significant features for each personality trait and then leverage them to build sophisticated personality models that can precisely predict individuals' personalities from their textual inputs. Moreover, multiple machine learning algorithms as (Support Vector Regressor, XGB Boost, and Feed Forward Neural Network) are built and used to assess the final models as reported in Table 2.4 and Table 2.2. For more details about all followed methodologies and concepts that generate the final models, please refer to Bin Tareaf et al. [12] in chapter two.

#### 5.3.2.2 Twitter Dataset

I randomly collected 1.7 million data points from Twitter API by only choosing public profiles that meet our filter criteria (only one face in the associated picture) checked by Face++ API [88]. The method yielded more than 600 unique Twitter users (383 males and 227 females), as it shown in Table 5.1. For each user, we obtained almost 2800 tweets on average associated with a hyperlink for the user's current profile picture.

We used the Facebook prediction models we described before (trained on features extracted by Pearson correlation) to label the current Twitter samples with the five-factor personality scores. Therefore, we ran the final prediction algorithms on the collected Twitter samples, and eventually, the prediction methods returned all the Big Five personality measurements for all samples. Furthermore, we used the link of the profile picture for each user to feed it directly to a deep learning-based system named Face++ API [88] to extract all facial features from each user image. The final facial features are presented in the appendix in Table .2.2. Face++ algorithm has approved its superiority in providing accurate face recognition, face demographics, and facial presentation [88].

For our task, we created a new data set that contains users' Twitter personality scores associated with his/her facial features scores. The generated dataset that we used to train the prediction models has 1.700.000 data points for almost 600 Twitter users. This amount of data allows me to use a fair amount of samples to train the final models. In order to evaluate the final output model, we used the random split sampling technique to pull out almost 20 % of the available samples to anticipate them as final testing samples for our final prediction models. The next subsection introduces the features extraction and features evaluation methods.

### 5.3.3 Features Extraction

In this subchapter, we handled a study to investigate a possible association between facial characteristics and actual personality scores assessed by the big-five personality models, which are Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness(A), and Neuroticism (N). Digital pictures of male and female faces were obtained, labeling the sample based on their published tweets using personality models to predict personality based on linguistics-based models that follow the NEO-FFI personality inventory standards. Facial photographs were analyzed to extract a set of facial measurements as features utilizing a digital image processing tool. Overall, the data on actual personality confirmed previous reports on predicting personality from human face images. For this reason, we decided to investigate the association between the extracted facial features and personality scores.

**Table 5.1:** Characteristics of Twitter Collected Data [19] ©2019 IEEE.

| **Characteristics** | | |
|---|---|---|
| # Samples = # Users | | 610 |
| # Tweets | | 1.7 m |
| # Male | | 383 |
| # Female | | 227 |
| # Facial_Features by Face++ | | 50 |
| # Generated Labels by Facebook Models | | 5 |
| **Predicted Labels** | **Mean** | **Standard Deviation** |
| Openness | 3.6277 | 0.7553 |
| Conscientiousness | 3.3581 | 0.7819 |
| Extraversion | 3.1162 | 0.7537 |
| Agreeableness | 3.3157 | 0.7932 |
| Neuroticism | 2.5691 | 0.7108 |

The primary analysis resulted in better results as we built the models following the concept of gender-based personality prediction models. Accordingly, we have investigated various male and female facial features associated with their personality scores. For each personality trait, we trained a dedicated model to predict it. We examined several feature sets for each predictor and then decided which facial feature affects the presence of specific personality traits and how gender as a fundamental determinant plays a principal role when it comes to predicting personality solely from users' profile images.

### 5.3.4 Face++ API

I used the Face++ API [88] which is an AI open Platform developed in 2012. It enables developers to use computer vision tools to better engineer facial features for their samples. Face++ API allows using the tool for different types of functionalities such as face detection, face comparing, and face searching. Using Face++, developers can identify and analyze human faces within the image they provide. For instance, Face++ Detect API can identify all the faces with any given image. Each identified face gets what so-called face_token in order to allow developers to keep track of each operation applied on given faces. With the standard free-of-cost developer API Key, users can define a rectangle region within the picture to apply face detection operation. This will return the face landmarks and attributes data from the studied sample such as (gender, age, smile value, blur value, head pose value, eye status value, skin status value, eye gaze value, mouth status value, beauty value, ethnicity, face quality value, and the associated emotion) and the payload of the face++ API request is returned in JSON format as shown in Figure 5.3. The full extracted facial features are presented in the appendix in Table .2.2.

After yielding all of the possible facial features from the examined samples, we used two different approaches to examine and select the best appropriate facial features for both genders. The first approach considers Pearson's Correlation Coefficient, and the second approach utilizes the Boosted Decision Trees' feature significance. The two mentioned methodologies for extracting feature sets for both genders are discussed in the following two sections. Figure .1 represents the inter-correlation between the independent facial features for both females and males Twitter samples.

### 5.3.5 Features Selection

#### 5.3.5.1 Pearson Correlation

The Pearson correlation coefficient is the most widely used feature evaluation method. It estimates the strength of the linear relationship between the studied variables. Therefore, we applied the correlation test as pairwise correlation analysis among the extracted facial features and personality records by utilizing the *Pearson product-moment correlation*. Following this methodology, we were able to derive $m = 5$ correlations with the same feature set (because the task is to predict five different personality traits). In order to overcome the various comparing issues, we applied what so called the *Bonferroni correction* to our global significance level of $\alpha = 0.05$ to decide the local significance levels: $\alpha^* = \frac{p}{m} = \frac{0.05}{5} = 0.01$.

In order to comprehend the correlation values and significance levels between the extracted facial features and the five personality traits, we visualised the results in a heatmaps as presented in Figure 5.4 for male samples and in heatmap Figure 5.5 for females sample. Typically, psychological variables have a correlational upper bound between 0.3 - 0.4 as presented in [93]. Surely, not all facial features are correlated with personality scores. To mention some, there are a facial features returned by Face++ algorithm but they have no correlation and significance with the personality traits for both genders, such as *blurness_threshold*, *gaussianblur_threshold*, *glass_value* or *motionblur_thresold*.

Facial features with overall high relative correlation coefficients for females are e. g. *smile_value*, *skinstatus_health*, *mouthstatus_open* and facial *emotions* in general as happiness emotion, where for males are e. g. *right_eye_status*, *mouthstatus_close* or *skinstatus_darkcircle*. Correlations for the Neuroticism trait specially among females samples are

```
"image_id": "9jrDd2xUw92jr0oDHHS",
"request_id": "1418427047286868,dacf2ff1-ea45-4842-9c07-6e8418cea78b",
"time_used": 136,
"faces": [{
    "landmark": {
        "mouth_upper_lip_left_contour2": {
            "y": 185,
            "x": 146
        },
        "contour_chin": {
            "y": 231,
            "x": 137
        },
        "right_eye_pupil": {
            "y": 146,
            "x": 205
        },
        "mouth_upper_lip_bottom": {
            "y": 195,
            "x": 159
        }
    },
    "attributes": {
        "gender": {
            "value": "Female"
        },
        "age": {
            "value": 21
        },
        "glass": {
            "value": "None"
        },
        "headpose": {
            "yaw_angle": -26.625063,
            "pitch_angle": 12.921974,
            "roll_angle": 22.814377
        },
        "smile": {
            "threshold": 30.1,
            "value": 2.566890001296997
        }
    },
    "face_rectangle": {
        "width": 140,
        "top": 89,
        "left": 104,
        "height": 141
    },
    "face_token": "ed319e807e039ae669a4d1af0922a0c8"
}]
```


**Figure 5.3:** Face++ API Sample Response When Request is Succeeded (Faces landmark, Attributes, Age, Gender and Face Rectangle Measurements).

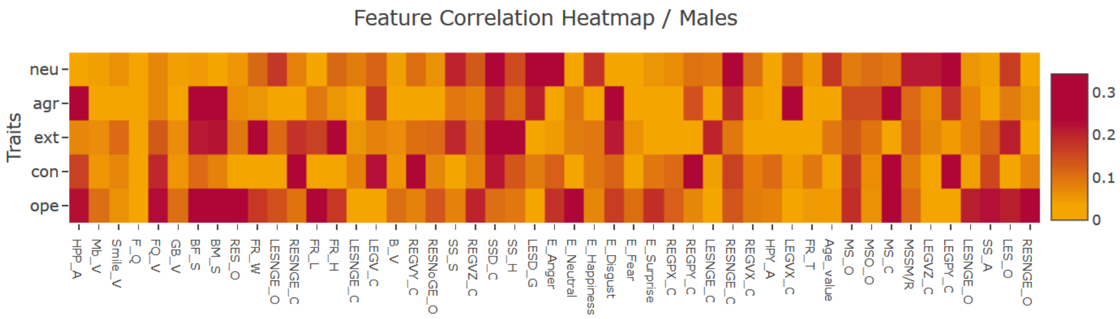



**Figure 5.4:** Pearson Correlation Coefficient Heatmap - Male Samples. Red = high coefficient, Orange = low coefficient [19] ©2019 IEEE.

harder to detect compared for other traits. Interestingly, this conclusion come in line with previous research we conducted which also reported difficulties in predicting Neuroticism trait from normal textual inputs.

To define all influential features for a specific personality trait, we took into consideration only features that show a significant correlation ($p < \alpha^* = 0.01$) for the studied trait $X$. As a result, we minimized the number of features by choosing features that presented strong correlations ($|r| > 0.05$). For example, the final feature set for the personality trait *Openness* is selected by choosing the features that significantly correlated with $p < 0.01$.

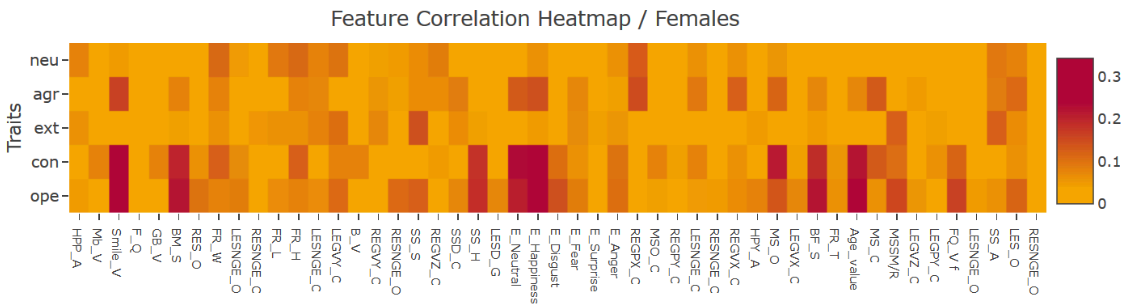



**Figure 5.5:** Heatmap for Pearson Correlation Coefficient - Female Samples. Red = high coefficient, Orange = low coefficient [19] ©2019 IEEE.

#### 5.3.5.2 Gradient Boosting

The second features extraction method we investigated to extract an adequate set of features for each personality dimension is computing the relative importance score by leveraging the gradient boosted regression tree [94]. Gradient boosted regression supports binary and multi-class classification and is used in multiple areas and domains such as web search ranking. Meanwhile, the approach which we consider to extract features further is the gradient boosting method, it manages to define multiple essential features for each personality factor. Therefore, we tested the hypothesis of having non-linear relationships among facial features and personality traits from this information.

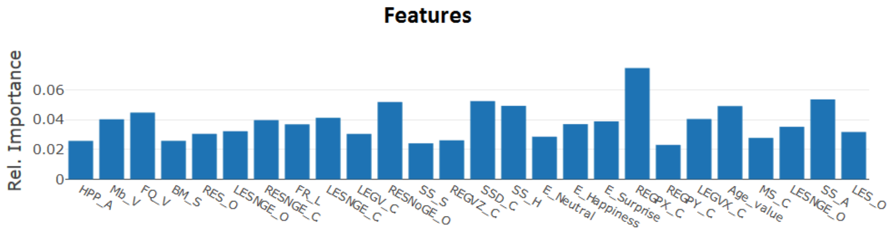



**Figure 5.6:** Significance and Relative Importance for Male Agreeableness : The Diagram Contains all Features with a Relative Importance Higher than 0.011 value [19] ©2019 IEEE.

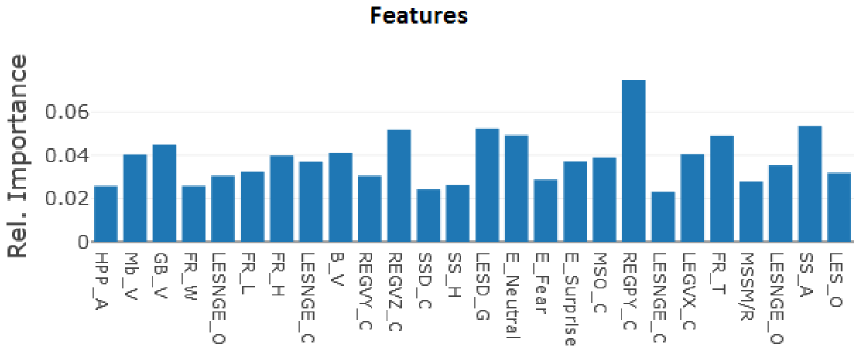



**Figure 5.7:** Significance and Relative Importance for Female Agreeableness : The Diagram Contains all Features with a Relative Importance Higher than 0.011 value [19] ©2019 IEEE.

The central concept of gradient boosted regression tree as feature extraction selection method is to learn a prediction model using all available features on that dataset. The final estimator will implicitly have the importance of each feature informing the final prediction. The features, in that case, are always randomly permuted at each split. Therefore, the best-found separation may fluctuate, even with the same training data and max_features=n_features [95]. To achieve a deterministic behavior when fitting, random_state has to be experimented. Gradient Boosting for regression has several parameters, to mention the most important ones:: loss function, learning rate, number of boosting stages to perform, subsample parameter used for fitting the individual base learner, function to measure the quality of the split, random_state to control the random seed for each tree at every boosting iteration, max_feature parameter for the number of features to consider when searching for best split, and ccp_alpha complexity parameter leveraged for minimal cost complexity pruning.

Indeed, we found that Twitter picture features are associated with the user's personality. As shown in the significance and correlations results, the most important correlations are found within the conscientiousness personality trait as shown in Figure 5.8 and Figure 5.9

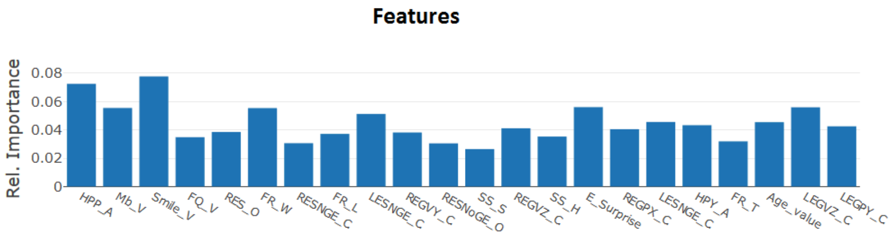



**Figure 5.8:** Significance and Relative Importance for Male Conscientiousness: The Diagram Contains all Features with a Relative Importance Higher than 0.011 value [19] ©2019 IEEE.

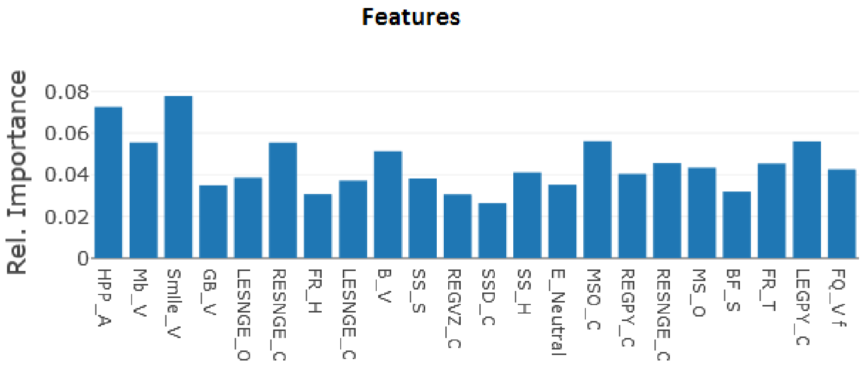



**Figure 5.9:** Significance and Relative Importance for Female Conscientiousness: The Diagram Contains all Features with a Relative Importance Higher than 0.011 value [19] ©2019 IEEE.

followed the agreeableness trait as shown in Figure 5.6 and Figure 5.7. Meanwhile, the weakest significance is found within the openness to experience personality trait as shown in Figure .4 followed by neuroticism personality trait as shown in Figure .3. The significance of most important examples of the extracted features is shown in the appendix as Figure .2 for male-Extraversion, Figure .3 for male-Neurotosim, Figure .4 for male-Openness.

## 5.4 Training Ensemble Machine Learning Algorithms

Based on the fact that the big five personality results are continuous scores ranging between 1 to 5, we decided to examine the regression algorithms to estimate individuals' five personality dimensions. In general, regression algorithms estimate a mapping function from the feature vector to a continuous output variable. Considering the collected learning samples, we trained different machine learning algorithms that support vector regression, Random Forest, Bagging, and Adaptive Boosting. All of these algorithms are characterized in the next part.

The five-factor personality model characterizes an individual's personality through five personality dimensions. Therefore, we decided to train the algorithms and built and trained a separate model for each personality dimension. Every classifier is learned on two different trait-specific features, defined by two different methodologies for all five personality dimensions (see section 5.3.3).

All trained estimators need some tuning for the hyperparameters in order to catch the best possible combination between them. We used the Mean Squared Error metric for evaluating the final accumulated error. The used metric for evaluation is detailed in section 5.4.1.

Support Vector regressor represent the input vectors in an infinite dimensional vector space SVR calculate the likeness of two feature vectors $\boldsymbol{x}$ and $\boldsymbol{x'}$ in the input space. The Kernel $K_{RBF}(\boldsymbol{x}, \boldsymbol{x'})$ is large if the euclidean distance between the two feature vectors $\|\boldsymbol{x}-\boldsymbol{x'}\|$ is small. The *rbf*-kernel has one free parameter $\sigma$. Together with the regularization parameter $C$ of SVR two hyper-parameters can thus be optimised in the learning phase.

Random Forest as its known, is an ensemble learning method capable of performing both regression and classification tasks based on multiple decision trees. The data is mapped into an n-dimensional space where each dimension stands for one feature to create a decision tree. The algorithm then defines a decision boundary and divides the dataset into two non-overlapping partitions. This is continued until a remaining group can be perfectly separated, resulting in very small buckets and overfitting having a very low bias but a high variance. To mitigate overfitting, stopping criteria could be introduced like a maximum tree size assigning the majority value as a result to a bucket (tree pruning).

Bagging is known as the Bootstrap Aggregation method, where we utilized it when I intended to minimize the variance of a decision tree. The central concept is to make several subsets of data from the learning sets decided randomly with replacement (where random forest takes the random selection of features instead of utilizing the whole features to enlarge trees). Every group of the data is utilized to learn their own decision trees. Consequently, it generates an ensemble of various models. Therefore, instead of using a single decision tree, an average of the estimators from several trees is utilized to have better final results.

Known as Adaptive Boosting and used mainly to let the model learn from mistakes by increasing the weight of misclassified data points. The learning process starts by initializing the data points' weights and then training a decision tree for each trait. Later, we calculate the weighted error rate of the decision trees known as $e$. We calculate the decision tree's weight in the ensemble and then update the weights of wrongly classified points. We repeat the mentioned steps until we make a final prediction. Obviously, the tree with a higher weight will have more influence on the final decision.

In other words, the first step is to train a weak learner on the real data. For each iteration, the sample weights are independently adjusted and the learning algorithm is rerun to the reweighted data. At that point, training samples that were wrongly predicted by the boosted model produced by the former step have their weights increased, whereas the weights are decreased for those that were predicted correctly. As repetitions continue, samples that are hard to estimate, experience a growing influence. Therefore each following weak learner is required to focus on the samples that are missed by the prior ones in the sequence as described in [96] [97].

### 5.4.1 Quality measures

Multiple evaluation metrics are available for data scientists to evaluate the correctness of their final prediction models. As a result, we decided to assess the final prediction models using the mean squad error metric. It calculates the variation between personality predicted scores against the real actual self-reported scores. The metric is represented with the following formula:

$$MSE = \frac{\sum_{t=1}^{n} (y_{t,act} - y_{t,pred})^2}{n} \tag{5.1}$$

### 5.4.2 Results

In this section, we visualize the effectiveness of the four trained algorithms' among multiple feature sets. Finally, we evaluate the performance of every algorithm we trained. Both feature extraction approaches we introduced before produce very similar outputs based on the ("MSE"), which also perform better comparing it to the baseline for all personality dimensions.

As Table 5.2 points out, the Random Forest algorithm produced the most accurate results when trained on trait-specific features resulting in lower errors for four personality traits concerning the males' samples. SVR performed very well (confirms several previous types of research in this area) for the prediction task. Interestingly, SVR outperformed Random Forest prediction for only the Neuroticism personality trait. Extraversion was the easiest among all personality traits to predict on this dataset, followed by conscientiousness, agreeableness, and neuroticism, where Openness was the hardest for males samples. Extraversion was the easiest of all personality traits to predict for the female samples, followed by conscientiousness, Openness, and agreeableness. Neuroticism was the hardest to predict for female samples. Despite the small sample size, Bagging and AdaBoost were the worst predictors for the task as presented in Figure 5.12 and in Table 5.2 and in Table 5.3.

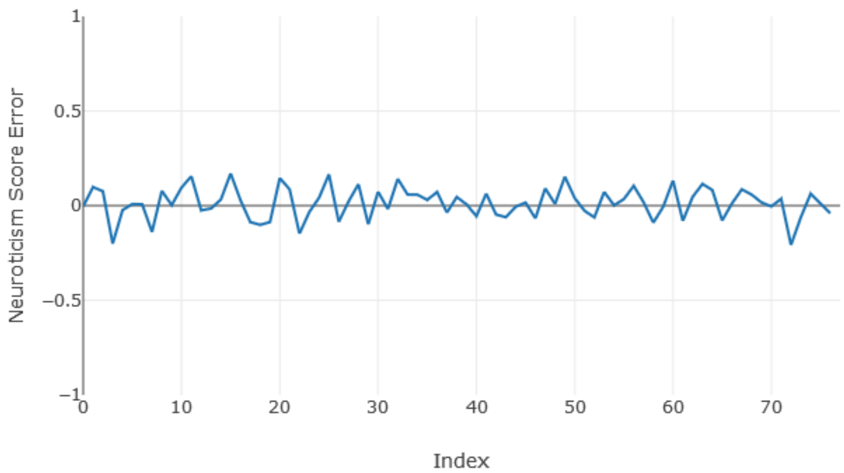



**Figure 5.10:** Mean Squared Error Results for Neuroticism Personality Trait With Support Vector Regressor among RBF Kernel at the Males Samples [19] ©2019 IEEE.

We report the behaviour of the testing male data set in Figure 5.10. All traits behaviour with all possible algorithm combinations experiments. Below are a few illustrated for space's sake.

Females' predictive models performance is reported in Table 5.3. Random Forest algorithm had the best results when trained on females' trait-specific features resulting in higher accuracy in predicting Openness, Extraversion, Agreeableness, and Neuroticism from their facial features. Support Vector Regressor was the best performing algorithm only in predicting Conscientiousness personality traits with 0.07 accumulated mean squared error as presented in Figure 5.11. In general, SVR and Random Forest were the best performing algorithms to predict personality traits for both genders. On the female dataset, Bagging was the worst prediction algorithm.

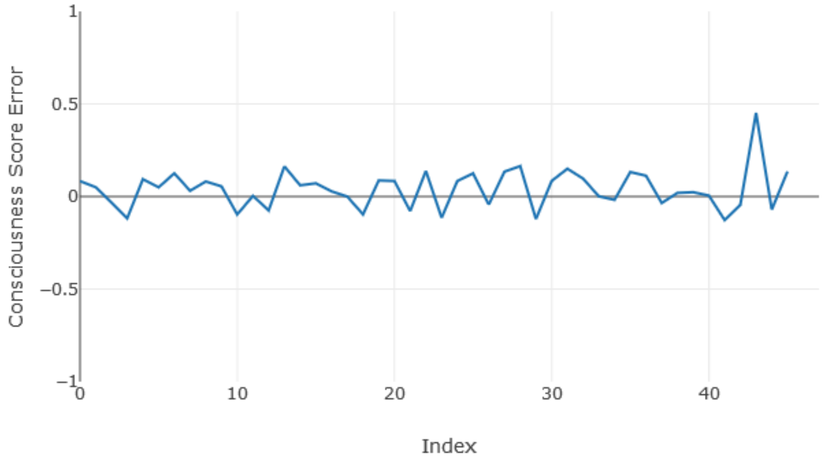



**Figure 5.11:** Mean Squared Error Results for Conscientiousness Personality Trait With Support Vector Regressor among RBF Kernel at the Females Samples [19] ©2019 IEEE.

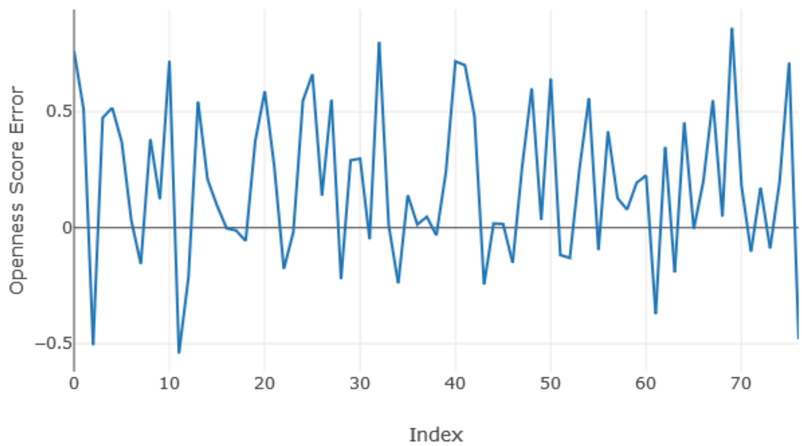



**Figure 5.12:** Mean Squared Error Results for Openness to Experience Personality Trait With Adaptive Boosting Algorithm at Males Samples [19] ©2019 IEEE.

The results of the mean squared error for the models trained using support vector regressor at males samples for openness to experience trait is reported in the appendix in Figure .5, for females neuroticism trait, the error for the support vector regression models is reported in Figure .6. The results of the mean squared error for the models trained using adaptive boosting approach at males samples for conscientiousness personality trait is reported in the appendix in Figure .7 and Figure .8 for agreeableness considering females samples with adaptive boosting approach models. The results of the mean squared error for the example models trained using random forest at males samples for conscientiousness personality trait is reported in Figure .9.

**Table 5.2:** Observed Predictive Performance using Mean Squared Error using Support Vector Regressor, Random Forest, Bagging and AdaBoost for Males Twitter Sample [19] ©2019 IEEE.

| Trait | SVR | R_Forest | Bagging | AdaBoost |
|---|---|---|---|---|
| **Openness** | 0.10619 | **0.09882** | 0.3265 | 0.2983 |
| **Conscientiousness** | 0.07426 | **0.06649** | 0.2621 | 0.2564 |
| **Extraversion** | 0.06683 | **0.05485** | 0.2530 | 0.2482 |
| **Agreeableness** | 0.08650 | **0.08444** | 0.2721 | 0.2459 |
| **Neuroticism** | **0.07694** | 0.09254 | 0.2029 | 0.1974 |

**Table 5.3:** Predictive Performance observed using Mean Squared Error using Support Vector Regressor, Random Forest, Bagging and AdaBoost for Females Twitter Sample [19] ©2019 IEEE.

| Trait | SVR | R_Forest | Bagging | AdaBoost |
|---|---|---|---|---|
| **Openness** | 0.1170 | **0.08266** | 0.5307 | 0.4773 |
| **Conscientiousness** | **0.0751** | 0.07991 | 0.4312 | 0.4067 |
| **Extraversion** | 0.0706 | **0.06714** | 0.4611 | 0.4172 |
| **Agreeableness** | 0.1000 | **0.08966** | 0.4267 | 0.4054 |
| **Neuroticism** | 0.1029 | **0.09384** | 0.3230 | 0.3117 |

## 5.5 Conclusion

In this research, we only used one user profile image to make the final prediction. Future work efforts would be to conduct and investigate a study with experience sampling where uploaded images from users are collected and analyzed over time. Further, more enhanced and engineered generated features from users' profile pictures will be investigated once we manage to increase our initial training samples. Although the information revealed within the profile picture is only a small part of the total information a user leaves on social platforms, the generated models from this study can be integrated into a learning ensemble that considers other relevant information such as Likes and Posts into further and more sophisticated prediction frameworks.

Because a significant amount of cross-domain research in data science and psychology has been done to grasp human online characteristics, estimating individuals' personality scores from their face and appearance is still largely unexplored, and the goal of this research is to investigate the best possible features to build automatic personality estimation models. The research tests are carried out with ethical and privacy concerns in mind. The purpose is to improve social media users' knowledge of what third parties may learn about them from what they publish and how they act on various social networking sites.

The final model utilizes two distinct approaches to select feature sets and evaluates four different types of machine learning algorithms. The final models can adequately estimate users' personality scores by analyzing a huge set of combinations among facial features with state-of-the-art machine learning models. We concluded that the evaluation of the best Random Forest models results reveals a considerable relationship between users' personalities and the photo they choose as profile pictures.

# 6 Personality Evolvement and Cross-Domain Assessment

## 6.1 Introduction

The personality of individuals, in general, is represented as the characteristic group of behaviors, cognitions, and emotional thinking patterns [98] [99]. Psychiatrists and domain experts interpret the personality of individuals to uncover what makes them unique and distinct compared to other individuals. This type of investigation serves a variety of purposes in the different research areas. For example, it estimates the outcomes virtually in many important life domains such as [100] physical health [101], job satisfaction [102], income [103] and relationship success [104].

The importance of personality traits assessment applications raised the awareness of identifying users' personalities automatically from their online fingerprints. The former and the legacy approach in estimating users' personalities is to complete a lengthy questionnaire containing 20 to 360 personality-related questions [62]. The NEO personality inventory model is based on linking words to determine five global factors which are *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness* and *Neuroticism*. However, the legacy approach of utilizing questionnaires to capture personality is observed as a time-consuming and complicated task. Therefore, the modern advancement in cross-domain research between data science and psychology can be joined and leveraged to overcome the cumbersomeness of using questionnaires and better capture personality traits in accelerated and cost-effective methods with no need to fill up any surveys or questionnaires.

Recent trends and advancements in the personality prediction field guided the exploration of various types of possible data available in social media platforms that can be employed to build automatic prediction models. Multiple experiments utilized several types of data ranging from social network contents and activities to ordinary textual inputs from posts/tweets [52] [25], likes history [18] [17], profile picture choice [84] [19], nonverbal behaviors [105], touchscreen-based interaction [106]. As manifested by Kosinski et al. in [16], the final trained personality prediction models from Like history were more accurate than assessments made by users' friends or family members.

One of the most persistent questions in psychological research history is whether human personality stays stable or changes over time [30] [107] [108] [109]. Some authors have developed theories and considered this topic as a debate, they stand with the belief which maintains personality is stable over human life-span [110] [111] [112]. while other researchers stand with the point of personality variation over time [113] [114] [115]. As it has been confirmed that personality is encapsulated in natural human language [60] [116], in this chapter, we decided to investigate the big five personality traits development for a large sample collected from publicly available social media users that we tracked for over ten years within Twitter microblogging platform. Also, we will test the hypothesis that assumes users' language on the Facebook platform offers more self-disclosure than the Twitter platform and visualize how it affects the final personality prediction score. However, to our knowledge, no study has applied and used natural language processing or linguistic analysis to address personality development over time, and this chapter contributes to this area.

For the personality assessment task, popular brands and trademarks gradually appeared on social media platforms for advertisement, customer support, and public relation purposes; by now, it has become a necessity throughout all social network sites. This online appearance and interaction can be identified and represented as a brand personality that reflects how a brand is perceived by its customers. We exploited recent text analysis and personality detection research to build an automatic brand personality prediction model on top of the (Five-Factor Model) and (Linguistic Inquiry and Word Count) feature extracted from publicly available benchmarks.

The rest of this chapter is organized in the following style: the subsequent section 6.2 gave a brief literature review of personality development and evolvement over time as well as for the cross-domain personality assessment studies. Section 6.3 discussed the data collection, cleaning, features selection, and implementation methods. The experimental results are presented and discussed respectively in section 6.4, where we visualized in the same section how users' personalities have evolved and matured for over 10-years. Finally, section 6.5 showed year-by-year personality development for randomly selected users from the sample space. Also, we visualized some use cases where personality is comparable between two different social platforms. In section6.6, further research is identified, and some conclusions are drawn in the final section.

## 6.2 Related Work

### 6.2.1 Personality Evolvement

The personality evolvement field is an emerging sub-field of psychology referred to as the intersection of personality development and social psychology. Psychologists from the personality psychology domain confirm that personality change and stability research suffer from paradoxical issues where researchers might have large samples to analyze. However, very few personality assessments are continued and carried over the years because most of the previous studies intended to believe that personality is stable assuming and does not change [117]. The optimal datasets have to be big enough with multiple and frequent personality assessments over time to be able to understand the time scale when such a change unfolds. Also, psychologists confirm that the optimal personality assessment would not only have to have self-reports from individuals. Instead, reports that do not require any human involvement that relies on behavioral activity such as mobile sensing and digital social media footprints or even biomarkers are better and more accurate than traditional questionnaires.

The history of personality computational research is mainly centered on personality traits score estimation [118] [57] [119] [120] [20], rather than analyzing and understanding its stability/change over time [121] [122] [123]. Therefore, in this study, we decided to investigate personality stability using our machine learning algorithms that have previously shown promising accuracy in estimating users' personalities by analyzing the textual inputs they share in social media platforms [25] [12].

Costa et al. in [122] research paper established and presented proof by a set of experiments the stability of personality, and he placed one of the first theories which state that humans' personality is unchangeable as he termed it as "set like plaster". He attested empirical evidence on personality stability by analyzing data from longitudinal studies of personality, including the Institute of Human Development studies [124], the Normative Aging Study [125] and the Duke Longitudinal Study [126]. He argued that these studies are all different in the sample composition, the age of the participants, and the tools used to measure personality, but they were nearly all comparable in their agreements and conclusions on the stability of personality in adulthood.

Neyer et al. [123] evaluated personality and human social relationships of a general population sample for 489 German young adults across four years. He claims the stability and change of personality and relationships can be investigated from different perspectives. He also addresses the different types of stability and assesses two kinds of personality-relationship stability: mean-level stability and rank-order stability, the main type of stability of personality in young adulthood. The limitations of such a study rely on self-report data. The longitudinal study only measured the personality of users on two occasions. The study itself contains a relatively large time interval and did not consider short-term variations in personality and social relationship development.

Helson et al. [30] studied the influence of gender and birth cohort on personality stability by creating a cross-sectional study combining personality inventory data from two different longitudinal cohorts to examine the change of personality with age over adult life. They investigated two longitudinal studies with hierarchical linear modeling for 212 original members of the Oakland growth study born in the 1920s. The samples were interviewed four times in their adulthood and answered the California Psychological Inventory at approximately four different ages, which are 33,49,62, and 75. The author found that most personality changes happen before the age of 30. The longitudinal evidence concluded that the global cohort effect could be considered the main contributor to personality change with age at particular spaces comparable to culture, cohort, and gender.

Damian et al. [127] addressed the first study in psychological science to test the stability and change of personality for over 50 years by investigating two independent samples N=1795, one is cross-sectional, and the other is short-term longitudinal to validate the personality scales and measurement error. He analyzed the data using four different methods: rank-order stability, mean-level stability, individual-level change, and complete profile stability. He also showed that almost 60% of the sample had maturation and reliable change with all personality traits where gender played a little role in lifespan personality development. He concluded that personality has a stable component over time, and the traits are only getting mature as they age.

### 6.2.2 Cross-Domain Personality Assessment

For the cross-domain personality assessment task, we looked into the research question regarding if the same user writes and uses linguistic words on two different social platforms in the same pattern. Compared to traditional methods for determining users' personalities, social media footprints for predicting personality promises straightforward and direct insights. Farnadi et al. [128] examined a variety of univariate and multivariate regression methods on datasets from Facebook, Twitter, and YouTube. The multivariate models often exceeded the univariate ones, but the variations were insignificant. They found that no common features can be recognized that operate properly on all social media datasets. Even extending a model with training samples from another social network could not improve their regressors. They concluded that the context of the data plays a major role in training phases. Their dataset from YouTube was labeled by impressions, whereas their Facebook and Twitter labels were self-reported throughout psychometric questionnaires.

Hall et al. [129] studied the effects of self-representation of Facebook users to study their social phenomena within social media networks. Users of social media platforms consciously or subconsciously represent themselves in a way that is suitable for their readers. The shortage of proper methods to identify and control these effects restricts research findings. They handled a case study concerning 509 paid Amazon Mechanical Turk workers. They provided psychometric survey results and Facebook footprints to the researchers. The data was employed to predict the user's personality according to the (FFM) using (LIWC)-only features. The study pointed out that self-representation is an existing phenomenon in social media and that personality is still detectable even when self-representation is present.

Both research efforts ([128] and [129] ) utilized supervised machine learning approaches to predict user's personality according to the (FFM) model. [130] proposed a new approach using linear semi-supervised regression to improve prediction results. Their study is based on data with 1792 users collected from Sina Microblog, the most popular social platform in mainland China. They stated that their empirical results support their thesis that unlabeled data could improve prediction results.

Bai et.al [131] proposed an approach using multi-task regression and incremental regression for online behaviors to predict from the Sina microblogging platform the Big-Five personality. Their study is based on survey data of 444 users and shows that the correlation factors are significant between different personality dimensions. They stated that their training data set is reliable enough, and multi-task regression performs better than other modeling algorithms.

The research effort of Jaidka et al. [132] focused on comparing self-disclosure on Facebook platform versus Twitter platform collected and processed social media data for the same users for both platforms, and this empowered them to perform a comparative analysis under a proper scientific setup. The results indicate that users prefer to self-disclose more on Facebook than on Twitter ( Twitter recently increased tweet character limit to 280 characters, up from 140) as platform affordances play a big role in determining users' self-disclosure behavior. They also concluded that Facebook and Twitter are equally good at estimating user traits if the same-sized language representations are utilized for training the language models.

In respect to brand personality analysis, two research papers from the business domain are relevant. Aaker et al. [133] constructed a theoretical framework of the brand personality developed by learning the kind of dimensions of brand personality (Sincerity, Excitement, Competence, Sophistication, and Ruggedness). Geuens et al. [134] also developed a new brand personality measure consisting of a dimension mapping to (FFM) personality items (Responsibility = Conscientiousness, Activity = Extraversion, Aggressiveness = Agreeableness, Simplicity = Openness, Emotionality = Neuroticism) in contrast to other models [133]. In this work, we will examine if the same public entities hold the same personality patterns on two different social networks or not.

## 6.3 Implementation

The implementation section is divided as follow: In the first part, we will describe the training dataset from the well-known Facebook application called the MyPersonality project, and we will discuss in detail the second textual dataset we collected from the Twitter microblogging platform associated with timestamps to answer the final research question of personality stability (6.3.1). In the second part, we will discuss the Stability Coefficient Measurements value between all the estimated personality scores tracked for 10-years in (6.4) and we will further analyze the output from the perspective of Inter-Individual Change Trajectories in (6.4.1) and the Intra-Individual Change Trajectories in (6.4.2).

### 6.3.1 Data Acquisition

To solve the research question and quantify personality stability, we decided to approach the problem from an NLP (Natural Language Processing) perspective instead of addressing the problem as theoretical psychologists. Therefore, we used the well-known and the golden dataset MyPersonality's Facebook project as training data to train our personality prediction models that take textual data as input and finally generate the five personality trait scores as discussed in detail below.

As the MyPersonality Facebook dataset does not have timestamps associated with the linguistic posts, we collected another dataset that contains textual inputs for users for the last (1-10 years) with timestamps; each user eventually was represented by a sufficient amount of tweets to allow year by year personality assessment. We implemented a data crawler to collect publicly available data from Twitter that we handled as privately as possible for this research. The final results are anonymized as presented in the results section.

The collected Twitter dataset is not labeled yet with the needed personality scores and, therefore, cannot train machine learning algorithms. Therefore, we used the Facebook golden samples from MyPersonality's Facebook project to build the final personality prediction models in order to use it as a tool to label the Twitter samples. For this reason, we used the prediction models to build personality profiles for each user (year by year) to better analyze and understand personality stability and change over the studied years. The rest of this section describes in-depth the validity of the proposed study and how we used the stability coefficient to conclude the final results.

### 6.3.2 Training Dataset

MyPersonality dataset as we described before, is a Facebook application offered by [15] in 2007. The application enables its users to engage in psychological research by filling in a personality questionnaire similar to the revised NEO Personality Inventory from [62]. The project resulted in more than six million psychometric test results where participants donated their data for academic research purposes as shown in Table 2.1. The project was terminated in 2012 due to a lack of time to maintain it from the original authors. We managed to secure a full copy of the dataset after they anonymized the whole data points and identity of users.

### 6.3.3 Testing Dataset

#### 6.3.3.1 Twitter Time-Stamped Dataset

Using Twitter REST API access tokens and the Tweepy Python library, we managed to gather random public users' data by only selecting public profiles to collect their full history of tweets. We filtered the outcome to make sure we only include in our analysis the users who have a long history of tweeting behavior. This filter criterion resulted in 750 unique Twitter users (gender distribution was not given or specified by most samples) as shown in Table 6.1. For each of the identified users, we crawled almost 7800 tweets on average in order to examine whether their personality stays stable over the past 10 years and if it changed; to which extent it got changed.

**Table 6.1:** Characteristics of Twitter Collected Dataset [20] ©2020 IEEE.

| **Characteristics** | |
|---|---:|
| # Tracked Period | 2010-2019 |
| # Samples = # Users | 750 |
| # Tweets | 9.5 m |
| # AVG Tweets Per Month Per User | 106 |
| # Generated Labels by their Tweets | 5 |

Several feature extraction techniques are examined using two unique methods which are the Pearson Correlation to capture the linearly correlated feature set with each of the personality traits and used the Boosting Decision Trees with relative importance to capture nonlinearity as we already have shown in our previous publication [12]. In the same research paper, we evaluated a set of multiple machine learning classifiers as a decision-tree-based ensemble machine learning algorithm that uses a gradient boosting framework. We investigated other algorithms like Support Vector Regressor and Feed Forward Neural Network for the task of personality prediction through textual inputs.

We adopted the generated Facebook prediction models we trained on Mypersonality dataset using the features that are extracted by two extraction methods from the LIWC tool (linguistic inquiry and word count). We predicted all Big Five personality scores and averaged by a one to two years time window for the whole Twitter collected samples. After various experiments as we showed in Table 2.2, the best performing personality prediction models are trained by XGB algorithm using features extracted by Boosting approach.

#### 6.3.3.2 Cross-Domain Datasets

To answer the research question and estimate the personality on two different social platforms, we decided to collect samples for the same individuals or entities from two social media platforms. Using the Facebook graph API, we collected public status updates from 32popular brands performing in the Top 50 of Forbes' *The World's Most Valuable Brands* list. Some brands are missing in the final dataset because they did not have a representation or appearance on Facebook or Twitter. Collectively, we accumulated 85.347 status updates for the whole lifetime of the brands' Facebook pages until January 2018.

In addition, using the Twitter API, we crawled tweets updates for the same previous Facebook list. Altogether, we collected 103.053 tweet updates. Table 6.3 provides statistics about the crawled Twitter pages. We did not examine further comments on brand pages' posts because the data found in the comments was noisy with a lot of links to other Facebook profiles. Manual inspection of the remaining posts revealed a lot of Spam. Some brands have only very few regular status updates, so we decided only to examine brands with at least 1.000 posts. Table 6.2 represent the statistics about the crawled facebook dataset.

## 6.4 Stability Coefficient Measurements

To measure personality change and stability for the five personality traits trajectories, we estimated the consistency between the bivariate correlation coefficient scores at the aggregated level for the whole sample considering the five traits. The results suggest that some individuals may shift in personality trajectories, and some remain stable from social media data. In this chapter, we decided to investigate the stability and change of users' personality traits scores from two different perspectives: inter-individual differences and intra-individual change.

Previous research in the personality development domain already noticed that personality traits are both stable and changeable. Personality traits are the characteristic set of behaviors, cognitions, and emotional patterns as well as thinking patterns [28]. This five-factor model is built on the relationship between terms and words users use and their reactions and experiences to specific events. Psychologists concluded that human personality could be defined and classified based on the five global factors: openness, Conscientiousness, Extraversion, agreeableness, and Neuroticism.

The theory behind personality traits is viewed as stable or changeable traits show methodical differences that researchers adapt in order to evaluate stability and change in human traits [135]. The two most outstanding and popular methods for this evaluation are called rank-order stability and mean-level change. To comprehend the differences between the two approaches, rank-order stability quantifies the stability of a specific personality trait for an individual within a group of individuals over time. The rank order stability metric answers the question of how an individual-specific personality trait changes over time compared to other individuals' same traits within the same group. Several factors may affect the personality rank order development, such as internal factors as the genetic system and external factors as environmental factors as specified by [127].

On the other hand, Mean-level change measures to which extent all individuals in a specific trait get developed or changed when it is measured over different time spans compared against the whole sample as described by [136]. This type of measurement did not give researchers the ability to quantify changes and development on the whole sample space instead of individual levels. The optimal measurements for individual development should be able to answer and explain individual differences. Therefore, the person-centered approach and the individual-level change metric should answer if a person's personality is highest in a specific trait, will his/her personality profile after ten years also be highest in the same trait.

### 6.4.1 Inter-individual Change Trajectories

The term inter-individual difference or the term mean-level change represents the variations of personality trait scores among the whole sample. This sort of personality transformation reflects to which extent a trait score decreases or increases over long-time measurements. A decrease or increase in trait-level among the samples triggers what so-called significant mean-level change. This type of measurement can be effective for psychologists in understanding people's personality changes over their life span without asking the individuals to fill up a personality questionnaire. Analyzing individuals' writing styles over a long period allows researchers also to observe how personality traits get developed in any given time window from tracking and interpreting their language usage.

We calculate the bivariate correlation coefficients to measure the inter-individual change for a given sample. We also calculated mean and standard deviation scores for the five personality traits generated from models trained to estimate the five personality dimensions from social textual inputs. To understand the results, we show below two different possible scenarios. In the studied sample, the consciousness personality trait showed a pattern of a change in scores over the studied period, whereas neuroticism personality traits showed clear stability in scores over the years for the same sample. As shown in Figure 6.1, the figure reflects the stability coefficients for the neuroticism trait between all samples for ten years studied period.

The measurements of the neuroticism personality trait scores are yielded by feeding our machine learning algorithms the full tweet history, year by year, for the whole aggregated sample to generate the personality score and calculate the bivariate correlations finally. The figure below represents the situation of neuroticism traits trajectories over time between all 750 users. The stability coefficient shows a clear variation for the personality scores in the first four years when the interclass correlation is the lowest compared for the whole investigated period. This instability can be interpreted as humans tend to be less stable in neuroticism traits in their adulthood and increase over time. Where after 2013, the stability coefficient presents a significant stability level for the neuroticism trait between all samples, reflecting gradual stability evolved from 2013 until 2019.

On the other hand, the conscientiousness personality trait revealed a different pattern compared to the neuroticism personality trait. As shown in Figure 6.2, the stability

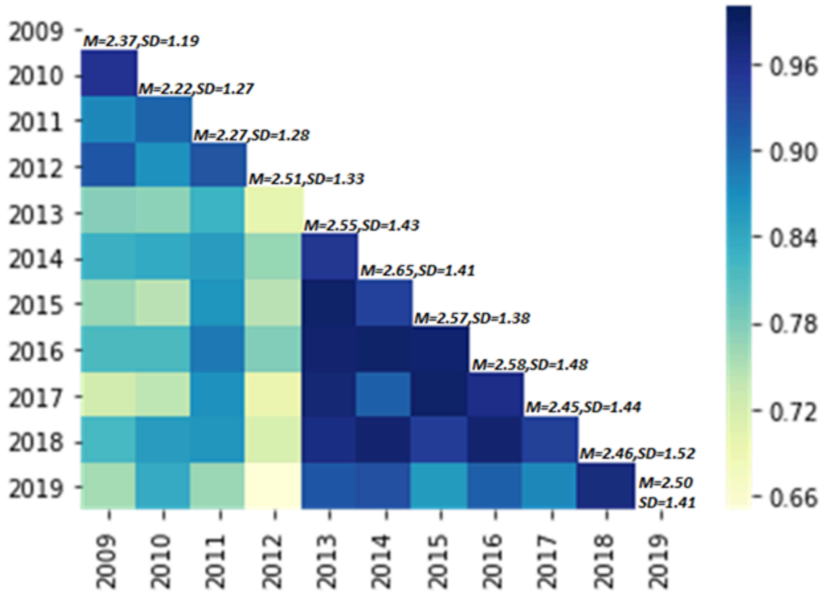



**Figure 6.1:** Stability and Change Coefficients, Means and SD for Neuroticism Personality Trait for the Last Ten Years [21] ©2020 IEEE.

coefficient shows that there was stability for that personality score for the total period from 2009 until 2019, where interclass correlation and between-persons variance is the highest at 0.9. The remaining 0.10 is considered as a within-person variation. The analysis shows among all the studied samples, the conscientiousness personality trajectory was one of the most stable traits over ten years with a slight deviation in the stability coefficient score at 0.7 in 2013 and 2014. The stability coefficients generated by aggregated textual inputs are comparable with previous studies that took physical and clinical setups by psychiatrists, as mentioned in the related work section.

### 6.4.2 Intra-individual Change Trajectories

As inter-individual change examines the stability and change at the aggregated users level within the whole sample, the intra-individual change term applies to the patterns of personality development at an individual level. Psychologists have addressed this kind of change in personality development within-person level by multiple factors. One possible factor is the ontogenetic factor that is dependent on environmental factors as mentioned by [137].

The other possible factor is mentioned by [117] which is the sociogenic factors that credited the personality development to the external environmental influence such as life events and certain life experiences that can have a direct effect on human personality

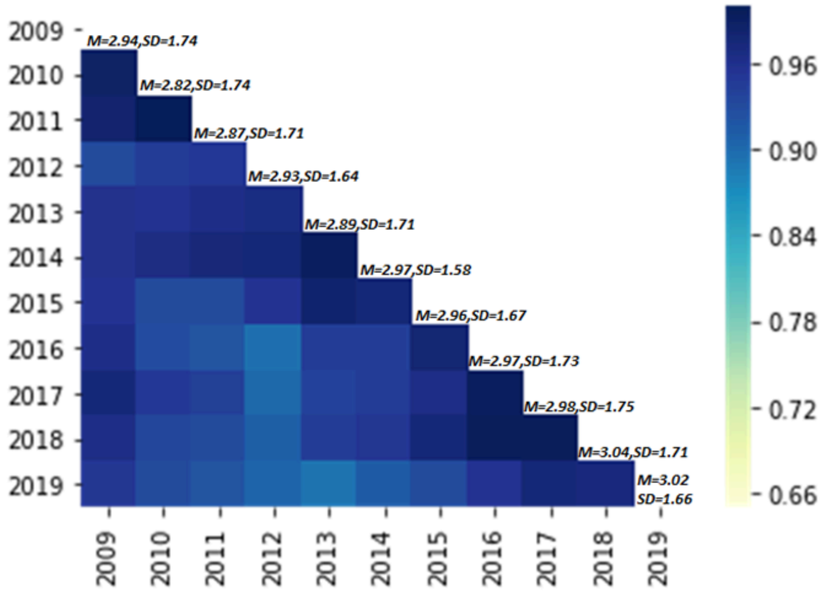



**Figure 6.2:** Stability and Change Coefficients, Means and SD for Conscientiousness Personality Trait for the Last Ten Years [21].

development. As we are concerned in addressing how personality change is formed based on linguistic posts, we decided to study the personality change at individual levels without considering the whole sample space as we showed in the previous subsection. In order to understand how such development of individual personality traits occur, we illustrate in Figure 6.3 the big five personality traits development for a randomly selected user from the sample space where personality scores got entirely estimated based on his/her own tweets shared publicly in Twitter for the last ten years.

We utilized the inter-individual change measurement to address the stability and change in personality scores for the whole sample, and we adopted the intra-individual change analyses to address the change or stability at the individual level. For example, in Figure.13 in the appendix, the stability coefficient for agreeableness shows that there is clear stability for the personality score for the period from 2012 until 2019 wherein between-persons variance varies between 75% and 90%. However, from the years 2009 till 2011, we observed a noticeable within-person variation between 15% and 65% indicating that the changes are happening at the user level instead of as a whole. Consequently, there was both stability for the sample as a whole ( between-persons variation) and change at the individual level (within-person variation) for some samples over the years in agreeableness traits.

As all individuals vary in their personality scores development, some individuals show less change while others do higher compared to the average. As we are examining personality development over time, we noticed that the used samples in this study show the two

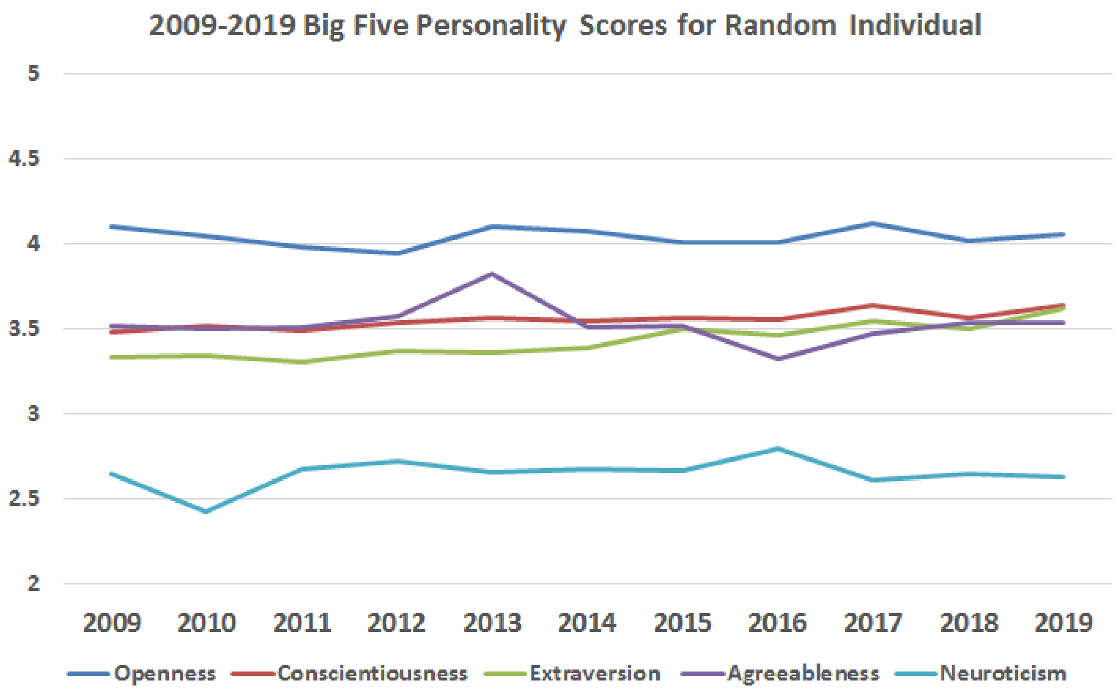



**Figure 6.3:** 10-Years Personality Development Progress for a Random User Selected from the Collected Sample [21] ©2020 IEEE.

possible patterns in human personality development. Psychologists address individuals' personality changes to environmental and genetic factors, as discussed above. In this study, the collected sample is missing for data that can be further engineered and leveraged to address the environmental and genetic effects on personality development from textual inputs.

## 6.5 Results and Discussion

### 6.5.1 Personality Evolvement

Concerning personality stability, the stability coefficients calculation resulted in understanding the overall stability and change patterns in the collected sample. For example, the conscientiousness personality traits for all samples showed a significantly high inter-correlation between the 750 samples indicating almost perfect stability around 95% between 2017 and 2019. While on the other hand, the openness to experience trait showed very limited stability coefficients ranging only between 15% and 30% for all samples from 2010 to 2012. Even though there is clear proof of stability in personality traits scores across ten years for some individuals, some individuals also have evidence of change. For this reason, we decided to visualize an example of personality development the whole ten years in two-year time windows for two random users to address the findings and understand the results.

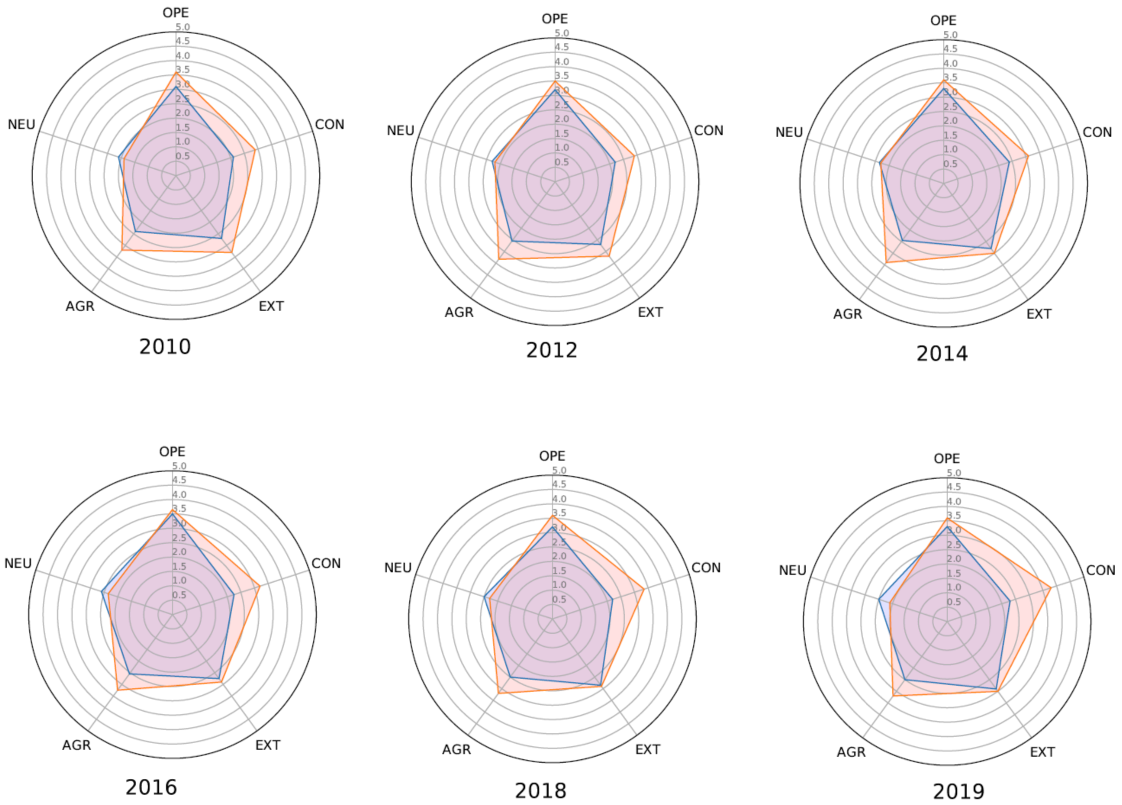



**Figure 6.4:** Personality Development Showcase for Two Random Users from the Twitter Microblogging Platform. Each Radar Chart Represents the Accumulated Personality Traits Scores Commencing from the Year 2010 (Top-Left) Till 2019 (Bottom-Right) [21].

In Figure 6.4, we showed the personality development trajectories for two random users from 2010 until 2019. The Orange radar chart represents user1 while the blue radar chart exemplifies user2. Historical tweets are grouped by two years intervals to calculate each personality trait score to show a perception of How personality gets developed over time. The Orange Radar Chart Represent Random User1 and the Blue one Embody Random User2I observed that in 2010, the tweets that we used to calculate the personality scores for user1 resulted in a dominant score of 3.6 out of 5 for the openness trait, while user2 showed a similar pattern with a dominant score of 3.1 out of 5 for the same trait. At the same time, user1 tweets resulted in a low score in neuroticism, around 1.85 out of 5, whereas user2 showed a higher score for the same trait with 2.1 out of 5. After nine years, we again gathered all tweets generated by user1 and user 2 to build their personality profile to examine to what extent it differs from the last nine years. We did the same from 2010 to 2012 to 2014 to 2016 to 2018, and finally to 2019. For 2019, we found that user1 and user2 indeed showed a variation within the final personality traits trajectories development. For instance, the user1 personality is no longer the highest and dominant in openness trait. Instead, the conscientious personality trait showed an impulse change in the score with a peak from 2.7 in 2010 to 3.7 in 2019, resulting in a noticeable and dominant change in personality trait scores for user1.

### 6.5.2 Cross-Domain Personality Assessment

We used the Facebook brand pages and Twitter brand pages specified earlier (in 6.3.1) to predict the Big Five personality traits with the proposed SVR model in chapter two. To compare whether our general personality prediction is accurate on brand data, we used the API from ApplyMagicSauce [138] to predict the traits on the same data. ApplyMagicSauce is a research project from the University of Cambridge, utilizing datasets like the MyPersonality project and questionnaires, Tweets, browsing data, and open text to recognize various psychological parameters. The five-factor model scores are presented in percentiles in ratio to the average of each trait in the whole dataset.

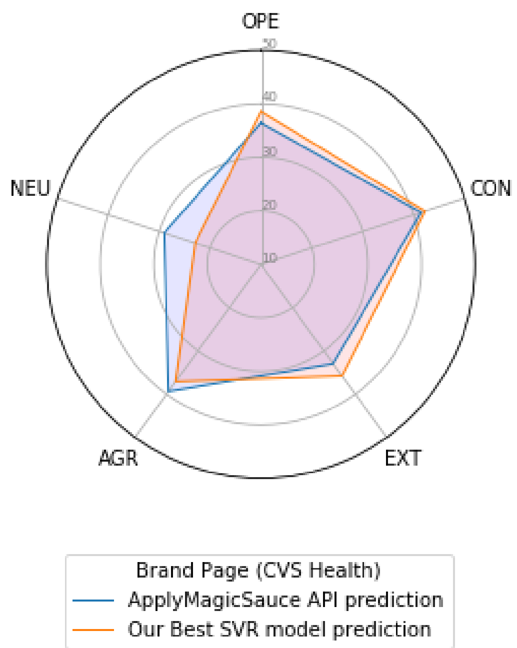



**Figure 6.5:** Personality Traits Scores Prediction for consumer Brand called (*CVS Health*) at ApplyMagicSauce API Versus the Proposed SVR Prediction Model using their Public Facebook Posts [20].

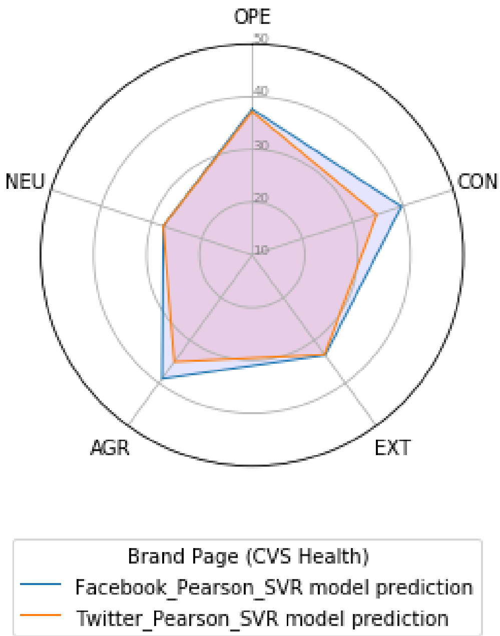



**Figure 6.6:** Personality Prediction for Consumer Brand called (*CVS Health*) using the Proposed SVR Model with Features Extracted by *Pearson-Correlation*. Prediction is Made Based on their Facebook Public Posts (Blue) VS Twitter Public Tweets (Orange) [20].

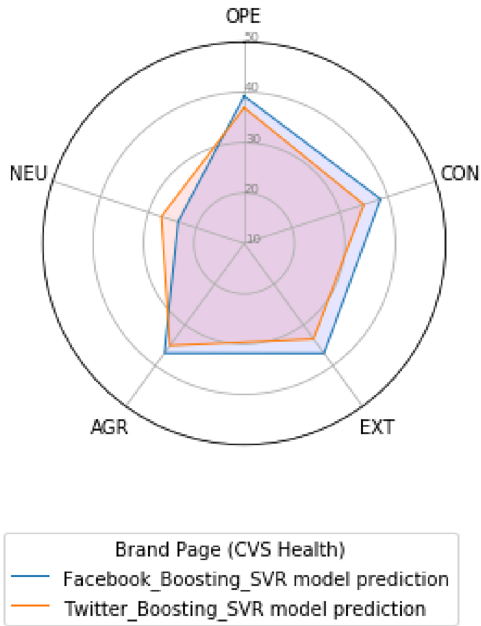



**Figure 6.7:** Personality Prediction for Consumer Brand called (*CVS Health*) using the Proposed SVR Model with Features Defined by *Gradient Boosted Regression Tree*. Prediction is Made Based on their Facebook Public Posts (Blue) VS Twitter Public Tweets (Orange) [20].

As noticed in the radar diagrams in Figure 6.5, the generated model is capable of detecting the five personality traits of Facebook pages on Facebook and reported significant improvements in detecting personality traits over the other by extracting and engineering several textual features from online available social fingerprints. Figure 6.6 and 6.7 represent the predicted personality scores for the same brand page by analyzing their public posts and public tweets using the proposed models, and it shows how feature extraction approaches at the training phase (Pearson versus Boosting trees) can influence the final predicted outcomes.

We assessed how supervised language models trained on Facebook samples are able to detect personality traits from Twitter samples. The results suggest that Facebook users tend to use more psycho-linguistic conceptual emotion categories words than Twitter users, leading to better personality prediction at the Facebook platform. The results are comparable to the state-of-the-art language models provided by [132] where they conclude that Facebook users prefer to use Facebook social platform for posting content about their personal relationships and personal concerns. In contrast, Twitter users tend to use the Twitter microblogging platform to post their psychological needs and derive.

Furthermore, the absence of limitations on the length of the post on the Facebook platform can be considered a major factor in Facebook's detection of Twitter in predicting personality dimensions. Figure 6.8 reveals how brands' (SAP) personality is diversified when language models are used to estimate personality traits based on their published Facebook posts and Twitter tweets.

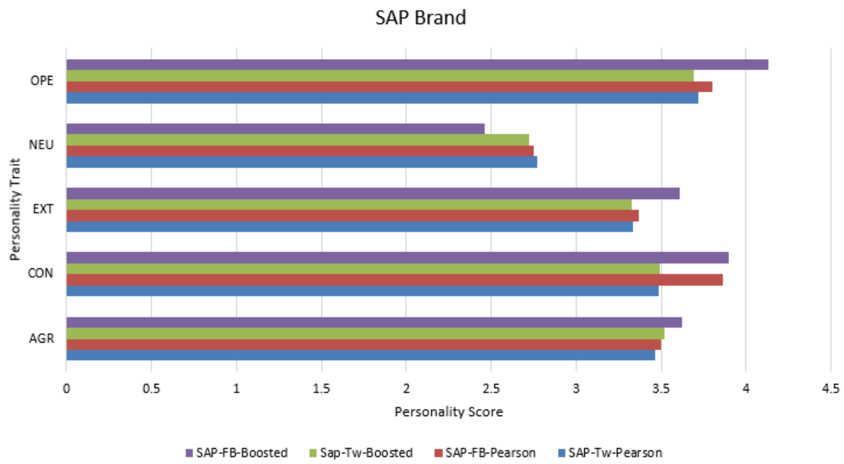



**Figure 6.8:** Personality Prediction for Consumer Brand called (*SAP*) using the Proposed SVR Model with Features Defined by *Gradient Boosted Regression Tree* and *Pearson*. Prediction is Made Based on their Facebook Public Posts Versus Twitter Public Tweets [20].

The evaluation of personality in online space is not a straightforward task. Building a gold standard for online brand personalities is possible by conducting interviews and questionnaires with employees as well as marketing and enterprise managers. A point worth examining is whether followers' personalities match the brand personality. A brand can take advantage and employ reverse psychology in marketing campaigns to attract similar or even totally contrary personality types. This knowledge would also greatly help public relations to identify the target audience of a brand over various Social Media networks. The same analysis is conceivable for employee personalities compared to the brand and could support human resources in a company or help new applicants find an appropriate job position.

## 6.6 Chapter Summary

Psychologists manifest that individuals are different in their personality development patterns, and it was not clear whether these phenomena were also observed on social media platforms. To our knowledge, this study is the first attempt of using social media data to advance the psychologists' understanding of how humans' personality traits develop over time without asking individuals to fill up a time-consuming questionnaire. Our research endeavors are implemented with full respect for ethical and privacy concerns. The goal is to show evidence from the available public big data publicly available on social media platforms about the stability and change patterns in users' personality development.

The chapter tracked personality development by analyzing individual writing behavior in public social media platforms. In the future, the optimal outline would be a consideration of wider and diverse samples from all over the world instead of only considering individuals whose textual inputs are in the English language. Even though some social platforms have become alive in 2006, we are considering collecting multiple types of inputs from users out of various social platforms to expand the features space diversifying from textual inputs to like objects and photo posting data to better understand individual's personality development from various perspectives and data types.

The final analysis examines the bivariate correlation coefficients of personality scores generated by machine learning models that can accurately estimate individual personality traits based on textual inputs. The analysis demonstrates the usefulness of the two main subcategories of the stability coefficient measurements: inter-individual and intra-individual changes. We showed multiple cases where personality stays stable and when it got changed based on users' public linguistic features. We hope that our findings encourage other researchers to establish collaboration at a bigger scale from data science and personality psychology to leverage social media platforms' fingerprints to address various unanswered questions from all behavioral domains.

The cross-domain study intends to predict entities' personalities from online social fingerprints with machine learning algorithms trained on labeled data from user self-report personalities tests at Facebook and Twitter platforms. It utilizes two separate approaches to pick the feature sets and assesses three different types of machine learning algorithms. The final model is able to accurately distinguish between personality dimensions of Facebook and Twitter pages by investigating a wide set of combinations between the extracted features with state-of-the-art machine learning classifiers. In terms of the implications for the machine learning domain, the analyses suggest that the source of the language samples can considerably influence the ability to capture users' personalities.

**Table 6.2:** Statistics for the Crawled Public Posts Dataset for Top Consumer Brands from Facebook Social platform [20].

| **Characteristics** | |
|---|---|
| # Samples = #Brands | 32 |
| # Features | 93 |
| Avg. #Posts | 2460 |

| **Brand** | **#Posts** | **Period** | **Brand** | **#Posts** | **Period** |
|---|---|---|---|---|---|
| ESPN | 4817 | 3199 | CVS | 2348 | 2697 |
| Cisco | 4596 | 3619 | Home Depot | 2247 | 3151 |
| Accenture | 3960 | 2913 | Wells Fargo | 2173 | 2700 |
| Amazon | 3912 | 3582 | UPS | 2004 | 2682 |
| Mercedes Benz | 3501 | 2320 | Verizon | 1753 | 2704 |
| Toyota | 3421 | 2970 | Google | 1706 | 3097 |
| HP | 3389 | 2923 | Siemens | 1595 | 1951 |
| Disney | 3292 | 3167 | H&M | 1538 | 3958 |
| GE | 3110 | 2475 | Microsoft | 1493 | 1911 |
| Intel | 3034 | 3473 | SAP | 1482 | 3705 |
| Gucci | 2942 | 2532 | Audi | 1470 | 1981 |
| AT&T | 2700 | 3515 | IBM | 1330 | 2268 |
| Ford | 2659 | 3331 | Nescafe | 1329 | 3107 |
| Walmart | 2548 | 3008 | Frito-Lay | 1266 | 2574 |
| Oracle | 2518 | 3073 | L'Oreal | 1193 | 2228 |
| BMW | 2394 | 1943 | Pampers | 1003 | 3199 |

**Table 6.3:** Statistics for the Crawled Public Tweets Dataset for Top Consumer Brands from Twitter Micro Blogging Platform [20].

| **Characteristics** | | | | | |
|---|---|---|---|---|---|
| # Samples = #Brands | | | | | 32 |
| # Features | | | | | 81 |
| Avg. #Posts | | | | | 3220 |
| **Brand** | **#Posts** | **Period** | **Brand** | **#Posts** | **Period** |
| ESPN | 3248 | 3199 | CVS | 3231 | 2697 |
| Cisco | 3224 | 3619 | Home Depot | 3234 | 3151 |
| Accenture | 3227 | 2913 | Wells Fargo | 3228 | 2700 |
| Amazon | 3204 | 3582 | UPS | 3208 | 2682 |
| Mercedes Benz | 3201 | 2320 | Verizon | 3224 | 2704 |
| Toyota | 3220 | 2970 | Google | 3201 | 3097 |
| HP | 3211 | 2923 | Siemens | 3224 | 1951 |
| Disney | 3248 | 3167 | H&M | 3224 | 3958 |
| GE | 3202 | 2475 | Microsoft | 3238 | 1911 |
| Intel | 3215 | 3473 | SAP | 3216 | 3705 |
| Gucci | 3216 | 2532 | Audi | 3201 | 1981 |
| AT&T | 3227 | 3515 | IBM | 3208 | 2268 |
| Ford | 3247 | 3331 | Nescafe | 3213 | 3107 |
| Walmart | 3198 | 3008 | Frito-Lay | 3242 | 2574 |
| Oracle | 3209 | 3073 | L'Oreal | 3235 | 2228 |
| BMW | 3221 | 1943 | Pampers | 3208 | 3199 |

# 7 Conclusion and Outlook

In this thesis, multiple studies that test the interaction and trade-offs between the choice of algorithms, feature collections, and their combinations for the task of predicting and modeling personality are investigated. While previous studies are adopting the same pair of classifiers and feature selection for all Big Five traits, in this thesis, we showed how this might potentially lead to miss-classifications for some of the personality dimensions. We showed how individual differences expressed in language, likes, profile pictures, and reactions require a unique modeling approach for each personality trait. Multiple experiments for personality computational tasks show how prediction performances are affected by increasing the number of attributes or changing the type of algorithms used.

We examined the literature on the effectiveness of using public social media data to study the relationship between users' personality traits against their written and published public posts on the Facebook social media platform. For this task, we estimate individuals' personality traits by adopting the MyPersonality dataset to investigate a wide set of linguistic features that impersonate an essential part in determining complex personality traits. We noticed that utilizing (LIWC) the linguistic dictionary, can assist and improve cross-domain research by cracking the path for more data scientists and psychologists to collaborate synchronically. We realized that the 82 linguistic features are important for the personality prediction task. Therefore, we decided to examine two different feature extraction methods: Pearson correlation and gradient boosting between the available personality traits scores and the linguistic features extracted by linguistic inquiry and word count tool. After we extracted all possible features using the proposed closed vocabulary approach, we adopted three different machine learning algorithms (support vector regression, gradient boosting, and feed-forward- neural network) to predict the personality scores from texts. The greatest personality prediction scores were achieved by training the XGBoost machine learning models using the defined Union features between the Common and Own sets of features extracted by boosting trees. Overall, the combination of all linguistic features we extracted reveals a high potential of using linguistic social network posts features for personality estimation tasks. The final models outperforms the baseline model and can adequately distinguish between personality dimensions by investigating various combinations between the extracted linguistic features with state-of-the-art machine learning classifiers.

The concept of using metadata as features for training machine learning models for the personality computation task is introduced. Like metadata is the data associated with each like object stated by the Facebook API. While the metadata of the like objects is a small proportion to the total information a user produces on social networks, the learning models that we developed confirmed that they can correctly predict an individual's personality entirely by accessing his/her likes list. Traditionally, entities were able to access an individual's personality by filling out psychological questionnaires. This chapter presents a method that indicates that a person's Big Five personality score can be easily predicted by leveraging the information about the pages a person likes on Facebook.

The results validate the significance of the correlations between users' personality and their Facebook Likes history and Likes categories. Predicting users' characteristics and preferences can be used to enhance numerous product services as well as reveal new opportunities for personalizing interfaces. We applied the hierarchy that Facebook uses to categorize pages as features to train the models and estimate users' Big Five scores for each personality trait. we also visualized that the more data the model has, the more accurate the final models will be. Ultimately, the Like-based trained models can be assembled and integrated into a bigger knowledge ensemble method that considers other relevant information (profile pictures, posts, and reactions) to better build personality prediction systems that can predict social media users' personalities by fair approximation.

A novel framework for predicting emotions and reactions distributions for any given post within the Facebook social media platform is presented. In 2016, Facebook introduced a new reactions feature that allows users to express their psychological emotions regarding published content using so-called Facebook reactions. We examined the potential of using Facebook reactions to identify and distinguish human emotions. For this task, we gathered an enormous amount of Facebook posts associated with their reaction labels using the introduced scalable Facebook crawler. The training process utilizes 3 million labeled posts for more than 64,000 unique Facebook pages from diverse categories. The evaluation at standard benchmarks using the proposed features shows promising results compared to previous research. The final model can predict the reaction distribution on Facebook posts with a recall score of 0.90 for "Joy" emotion. The proposed framework is trained on a dataset that was collected using our scalable Facebook posts crawler. While there has been plenty of research on sentiment interpretation in general, emotion classification is still mostly undiscovered, and this work contributes to this direction and aims to create a universal model using a single training set.

Future work for for predicting emotions and reactions involves examining several mapping techniques for emotion labels. As described in the data normalization section in the chapter, we joined the labels "laugh" and "love" to "joy" labels. Intuitively, these labels are considerably comparable, but joining them under one label might moderately twist the results. Furthermore, a more balanced dataset could be crawled to examine whether the performance of more under-represented labels like "surprise" and "anger" can be fixed. Nevertheless, as the collected and crawled datasets declared, aside from "Like", "laugh" and "love" are the most commonly and frequently used reactions.

We investigated how social media profile pictures differ based on the users' personality posting them. We handled profile images from the Twitter platform whose personalities are predicted based on 1.7 million data points in the experiment. we analyzed users' faces and extracted 50 unique facial features to measure the relationship between personality traits and profile pictures. The results reveal notable distinctions in user profile picture selection choices for different personality traits. Various machine learning approaches were investigated to test the effectiveness of these facial features in predicting users' psychological traits. To our knowledge, this work is one of the first attempts for using ensemble learning techniques for personality prediction tasks from the users' profile pictures.

Because a significant effort on cross-domain research in data science and psychology has been achieved to grasp the human online characteristics, estimating individuals' personality scores from their face and appearance are still mostly not fully explored, and this research effort contributes to this direction, and the goal is to investigate the best possible feature engineering to build automatic personality prediction models. Final models utilize two distinct approaches to select feature sets and evaluate four different types of machine learning algorithms. The final models can accurately estimate users' personality scores by analyzing a huge set of combinations among their facial features with state-of-the-art machine learning models. The evaluation of the best Random Forest models results reveals that there is a considerable connection between users' personalities and the photo they want as profile pictures.

We examined personality continuity and stability in the form of social media platforms. Individuals are evaluated and observed for over ten years in a longitudinal study to reveal how personality traits develop over time. Seven hundred fifty public users were tracked to cover the Big Five personality traits development when they first posted on the platforms and then ten years later. Examinations using advanced machine learning and stability coefficient measurements show two patterns that human personality follows in their online identity development across the life span: the inter-individual development pattern and intra-individual development patterns. We showed multiple cases of stability and continuity based on users' public linguistic features. We believe that the findings will encourage other researchers to establish collaboration at a larger scale from both the data science and the personality psychology domain to leverage social media platforms' fingerprints to address various unanswered questions from all behavioral domains. To my knowledge, this work is the first attempt to use social media data to answer how stable or changeable the personality is.

Although the domain of psychology science manifests that individuals are diverse in their personality development patterns, it was still not clear whether these phenomena were also observed in social media platforms or not. Our research endeavors are implemented with full respect for ethical and privacy concerns. The goal is to show evidence from the available public big data publicly available on social media platforms about the stability and change patterns in users' personality development. The final analysis examines the bivariate correlation coefficients of personality scores generated by machine learning models that can accurately estimate individual personality traits based on textual inputs. The analysis demonstrates the usefulness of the two main subcategories of the stability coefficient measurements (inter-individual and intra-individual) changes. We showed multiple cases where personality stays stable and how it changed based on users' public linguistic features. Predictive evaluation on brands' and entities' accounts reveals that the Facebook platform provides a slight advantage over the Twitter platform in offering more self-disclosure for users' to express their emotions, especially their demographic and psychological traits. Results also confirm that the same social media account carries similar and comparable personality scores over different social media platforms.

As some researchers have pointed out previously, conceptualizing thoughts and feelings in a standard objective manner remains unsolved, presenting the problem of considering self-reported reports as ground truth datasets for machine learning. We hope a better mechanism to capture personality in real life is introduced in the future. In general, digital footprints still lack essential evaluations of their psychometric properties, and we believe more studies and observations are needed to evaluate the reliability and validity of such samples. Social media data may introduce unique dependencies based on cultural factors, as likes of Facebook pages may have different meanings in different cultures.

Capturing online users' personalities is a double edge sword. Collecting and understanding users' generated online content allows companies to understand their users better to design and implement more customized services and interfaces. On the other hand, other parties can use it to deliberate users' opinions by targeting them through misleading content and suggestions. In general, online personality detection helped many parties and companies understand their end users correctly, e.g., by creating sophisticated personality-based recommender systems, early detection of mental disorders and depressions, fake news detection, improve online experience by customizing online advertisements and marketing, public safety as predictive and preventive systems, matching apps and team building services.

Many ethical considerations can be highlighted for the process of collecting and studying social media users' fingerprints. Users should be aware that their data can be collected and used by third parties for hidden purposes. Automated and comprehensive data collection methods can be utilised to determine if a user will get a specific job or graduate from college, leading to intentional shifts in online users' behaviour to avoid being judged based on what they post or like on social media platforms. Researchers, policymakers, engineers, and online users should collaborate to create more acceptable social media platforms and raise awareness of all involved parties.

Future research redirection can be enormous. To mention some, the big five personality model is a well-established model that has been used for a while now. By exploring users' massive social media content, we found that the big five models might be restricted to only five traits, and new traits and dimensions of personality can be introduced. For example, we can have new traits introduced automatically from the analysis by understanding the pattern of online behaviours. Another point we think is interesting is to collect and recruit more users to expand the current sample and features space and include a new type of fingerprints as sensing and biomarkers data. Furthermore, personality evolution studies from social media data can be investigated from likes records and profile picture contents over time instead of only using linguistic content as the main factor. Another suggestion we can mention to be included in future research efforts is to investigate the new MOE (misspelling oblivious embeddings) introduced recently by Facebook as a feature vector to train a deep neural network for the automatic personality prediction task.

# Appendix

## .1 Textual Features Extracted by LIWC Tool

**Table .1:** Feature sets Selected by Pearson Correlation Coefficient and Gradient Boosting Features Importance. The Boldly Printed Features for the Boosting Approach are the Common Defined Features (Feature set *common-B*). There are nine Features in this set.

| Set | Count | Features |
|---|---|---|
| **Pearson** | | |
| O-own-P | 17 | *Apostro, Sixltr, Tone, WC, affect, affiliation, article, death, drives, family, informal, insight, netspeak, percept, posemo, reward, time* |
| C-own-P | 30 | *Clout, Dic, Tone, achieve, affiliation, anger, article, body, death, drives, family, focusfuture, function., i, informal, negemo, netspeak, posemo, prep, quant, relativ, relig, reward, sexual, social, space, swear, time, we, work* |
| E-own-P | 9 | *Apostro, Sixltr, Tone, WC, affect, affiliation, informal, netspeak, posemo* |
| A-own-P | 19 | *Authentic, Clout, Dic, Tone, affiliation, anger, conj, drives, focusfuture, function., negemo, posemo, prep, relativ, sexual, social, swear, time, we* |
| N-own-P | 5 | *Analytic, Tone, WC, i, negemo* |
| common-P | 37 | *Analytic, Apostro, Clout, Sixltr, Tone, WC, achieve, affect, affiliation, anger, article, auxverb, bio, body, cogproc, conj, death, drives, female, focuspresent, friend, informal, leisure, male, motion, negate, percept, posemo, ppron, relativ, reward, sexual, space, swear, time, work, you* |
| **Boosting** | | |
| O-own-B | 39 | *AllPunc, Apostro, Clout, Comma, Dic,* ***Exclam****, OtherP,* ***Period****, QMark,* ***Sixltr****,* ***Tone****,* ***WPS****, article, assent, certain, conj,* ***death****, drives,* ***family****, focusfuture, focuspast, focuspresent, home, i, informal, insight, negate,* ***negemo****, netspeak, number, power, relig, reward, sad, sexual, social, space, time, you* |

*continued*

| Set | Count | Features |
|---|---|---|
| C-own-B | 38 | *AllPunc, Colon, Comma, Dic,* ***Exclam****, OtherP,* ***Period****, QMark,* ***Sixltr****,* ***Tone****,* ***WPS****, adverb, anger, article, assent, conj,* ***death****, drives,* ***family****, function., hear, i, informal, ipron, leisure, money, motion,* ***negemo****, prep, quant, relativ, sexual, swear, tentat, they, time, we, work* |
| E-own-B | 41 | *AllPunc, Apostro, Clout, Colon, Comma,* ***Exclam****, OtherP, Parenth,* ***Period****,* ***Sixltr****,* ***Tone****,* ***WPS****, adverb, affiliation, bio, certain, conj,* ***death****, discrep, drives,* ***family****, female, focusfuture, focuspast, friend, function., home, informal, leisure, motion,* ***negemo****, netspeak, nonflu, sad, sexual, social, space, tentat, they, we, work* |
| A-own-B | 35 | *Analytic, Apostro, Colon, Dic,* ***Exclam****,* ***Period****, QMark,* ***Sixltr****,* ***Tone****,* ***WPS****, affiliation, anger, article, bio,* ***death****,* ***family****, female, focuspast, function., home, ingest, ipron, male, money, motion, negate,* ***negemo****, number, power, relativ, relig, sexual, swear, they, time* |
| N-own-B | 38 | *Apostro, Clout,* ***Exclam****, OtherP,* ***Period****, QMark,* ***Sixltr****,* ***Tone****,* ***WPS****, anx, article, assent, compare,* ***death****, discrep, drives,* ***family****, female, focusfuture, health, home, informal, ingest, interrog, ipron, leisure, male, motion,* ***negemo****, number, reward, sad, see, shehe, space, swear, verb, you* |

## .2 Facial Features Extracted by Face++ Tool

**Table .2:** Feature sets for Males and Females Selected by Pearson Correlation Coefficient and Gradient Boosting Feature Importance. Facial Feature are extracted by *Face++*.

| Set | Type | Features' Name |
|---|---|---|
| HPP_A | continuous | *A Headpose_Pitch_Angle* |
| Mb_V | continuous | *Motionblur_Value* |
| Face_Q | continuous | *Face_Quality* |
| GB_V | continuous | *Gaussianblur_Value* |
| BM_S | continuous | *Beauty_Male_Score* |
| RES_O | continuous | *Right_Eye_Status_Occlusion* |
| LESNGE_O | continuous | *Left_Eye_Status_Normal_Glass_Eye_Open* |
| RESNGE_C | continuous | *Eight_Eye_Status_Normal_Glass_Eye_Close* |
| FR_L | continuous | *Face_Rectangle_Left* |
| FR_H | continuous | *Face_Rectangle_Height* |
| FR_W | continuous | *Face_Rectangle_Width* |
| LESNGE_C | continuous | *Left_Eye_Status_Normal_Glass_Eye_Close* |
| LEGV_C | continuous | *Left_Eye_Gaze_Vector_Y_Component* |
| REGVY_C | continuous | *Right_Eye_Gaze_Vector_y_Component* |
| RESNGE_O | continuous | *Right_Eye_Status_Normal_Glass_Eye_Open* |
| SS_S | continuous | *SkinStatus_Stain* |
| REGVZ_C | continuous | *Right_Eye_Gaze_Vector_Z_Component* |
| SSD_C | continuous | *SkinStatus_Dark_Circle* |
| SS_H | continuous | *SkinStatus_Health* |

*continued*

| **Set** | **Type** | **Features** |
|---|---|---|
| LESD_G | continuous | *Left_Eye_Status_Dark_Glasses* |
| REGPX_C | continuous | *Right_Eye_Gaze_Position_X_Coordinate* |
| MSO_O | continuous | *MouthStatus_Other_Occlusion* |
| REGPY_C | continuous | *Right_Eye_Gaze_Position_Y_Coordinate* |
| LESNoGE_C | continuous | *Left_Eye_Status_No_Glass_Eye_Close* |
| RESNoGE_C | continuous | *Right_Eye_Status_No_Glass_Eye_Close* |
| REGVX_C | continuous | *Right_Eye_Gaze_Vector_X_Component* |
| HPY_A | continuous | *Headpose_Yaw_Angle* |
| MS_O | continuous | *MouthStatus_Open* |
| LEGVX_C | continuous | *Left_Eye_Gaze_Vector_X_Component* |
| BF_S | continuous | *Beauty_Female_Score* |
| E_Surprise | continuous | *Emotion_Surprise* |
| E_Neutral | continuous | *Emotion_Neutral* |
| E_Anger | continuous | *Emotion_Anger* |
| E_Happiness | continuous | *Emotion_Happiness* |
| E_Fear | continuous | *Emotion_Fear* |
| E_Disgust | continuous | *Emotion_Disgust* |
| Age_value | continuous | *Age_Value* |
| MS_C | continuous | *MouthStatus_Close* |
| MSSM_R | continuous | *MouthStatus_Surgical_Mask_or_Respirator* |
| LEGVZ_C | continuous | *Left_Eye_Gaze_Vector_Z_Component* |
| LEGPY_C | continuous | *Left_Eye_Gaze_Position_y_coordinate* |
| FQuality_V | continuous | *FaceQuality_Value* |

*continued*

| **Set** | **Type** | **Features** |
|---|---|---|
| LESNoGE_O | continuous | *Left_Eye_Status_No_Glass_Eye_Open* |
| SS_A | continuous | *SkinStatus_Acne* |
| LES_O | continuous | *Left_Eye_Status_Occlusion* |
| RESNoGE_O | continuous | *Right_Eye_Status_No_Glass_Eye_Open* |
| FR_T | continuous | *Face_Rectangle_Top* |

## .3 The Extracted Facial Features Correlations

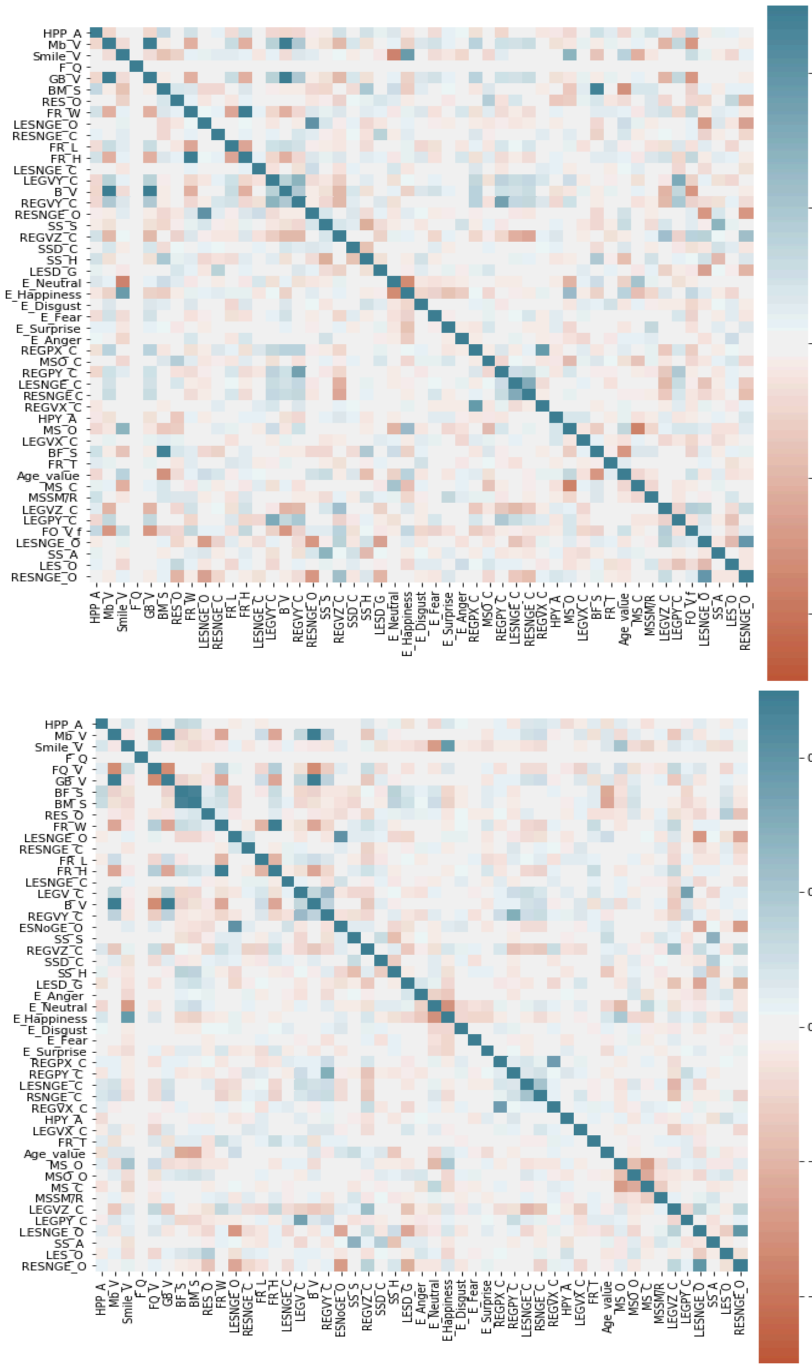



**Figure .1:** Inter-Correlation Coefficient Heatmap Scaled for the Facial Features Samples. The Higher part represent the Females Facial Features and the lower part represent Male Facial Features. Features full descriptions is available in table .2.2 in the Appendix. Blue = High Coefficient, Red = Low Coefficient.

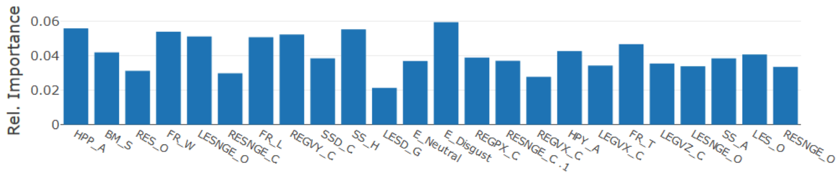



**Figure .2:** Significance and Relative Importance for Male Extraversion: The Diagram Contains all Features with a Relative Importance Higher than 0.011 value.

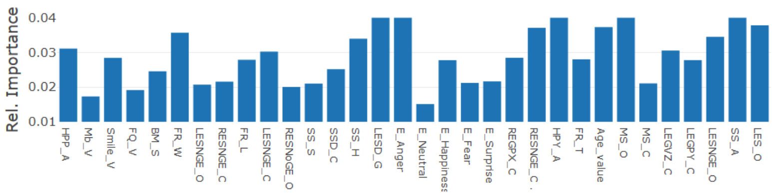



**Figure .3:** Significance and Relative Importance for Male Neuroticism: The Diagram Contains all Features with a Relative Importance Higher than 0.011 value.

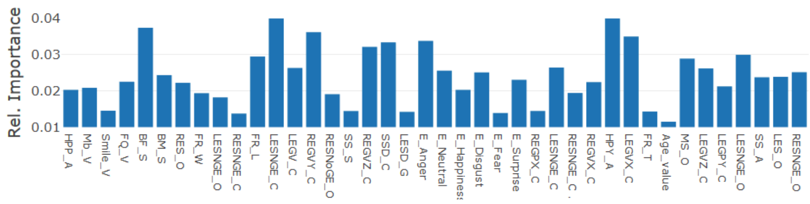



**Figure .4:** Significance and Relative Importance for Male Openness: The Diagram Contains all Features with a Relative Importance Higher than 0.011 value.

## .4 Mean Squared Error for Prediction Models

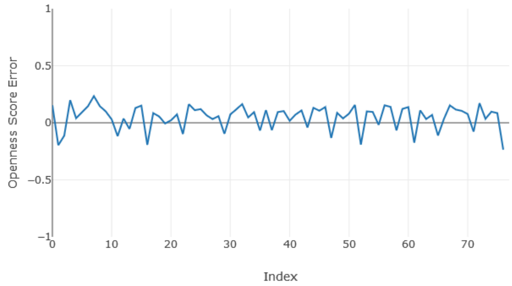



**Figure .5:** Mean Squared Error Results for Openness Personality Trait With Support Vector Regressor among RBF Kernel at the Males Samples.

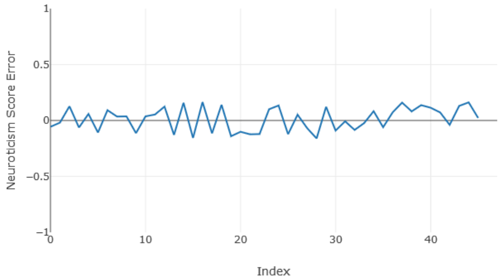



**Figure .6:** Mean Squared Error Results for Neuroticism Personality Trait With Support Vector Regressor among RBF Kernel at the Females Samples.

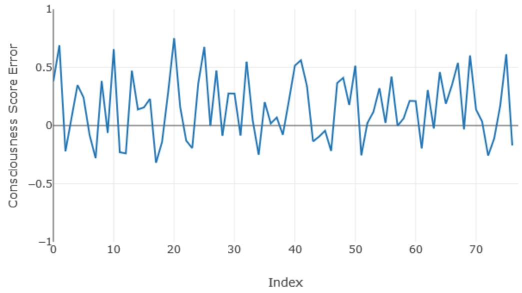



**Figure .7:** Mean Squared Error Results for Conscientiousness Personality Trait With Adaptive Boosting Algorithm at Males Samples.

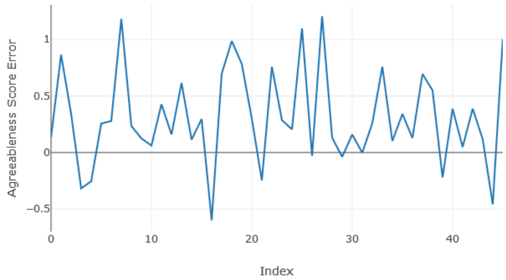



**Figure .8:** Mean Squared Error Results for Agreeableness Personality Trait With Adaptive Boosting Algorithm at Females Samples.

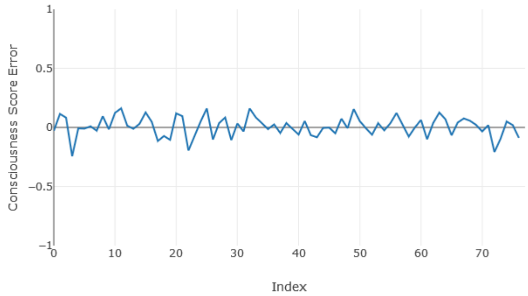



**Figure .9:** Mean Squared Error Results for Conscientiousness Personality Trait With Random Forest Algorithm at Males Samples.

## .5 Stability and Change Coefficients for Personality Scores

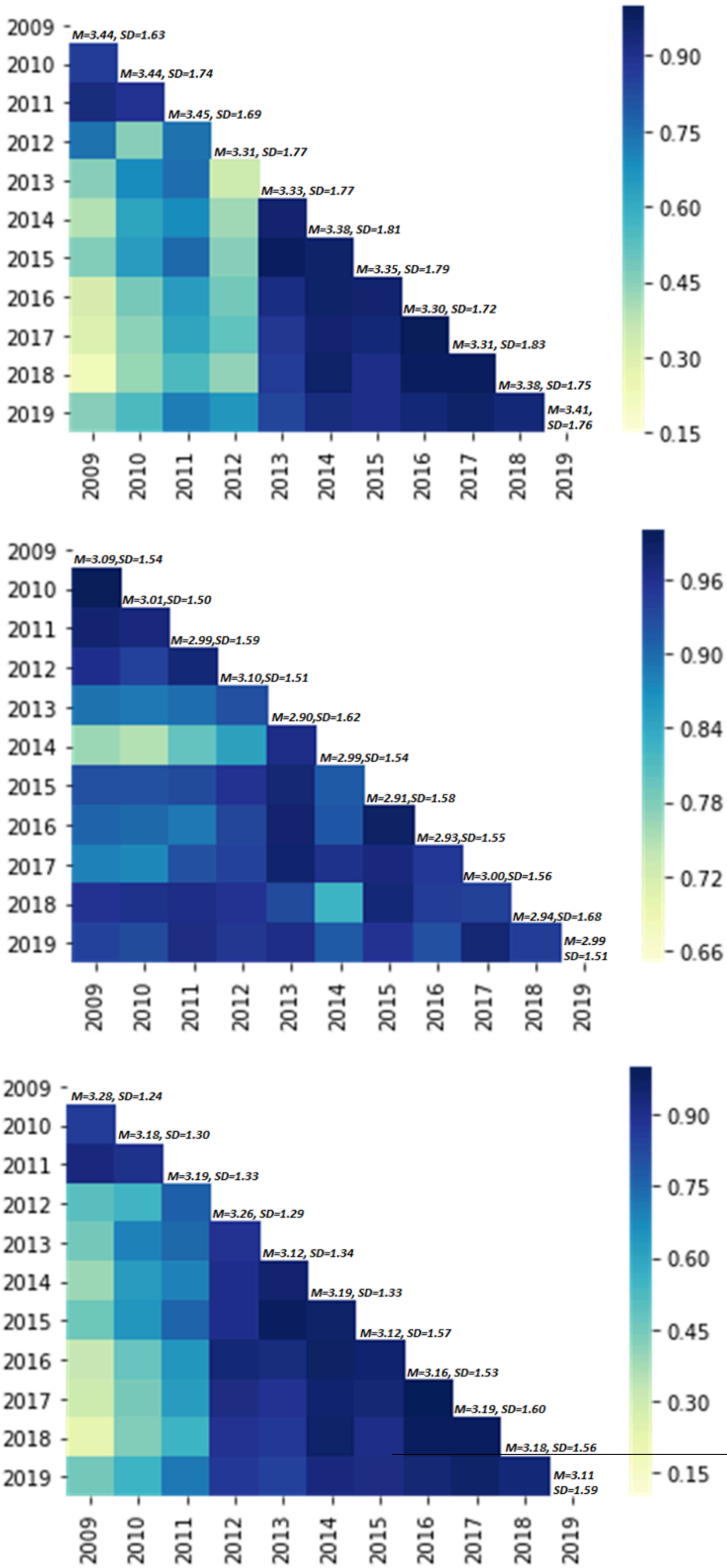



**Figure .10:** Top-Down Stability and Change Coefficients, Means and SD for Openness to Experience, Agreeableness and Extraversion Personality Trait for the Last Ten Years Measurements.

# References


[1] Facebook Inc. *Facebook Company Info*. `https://about.fb.com/company-info/`. [Online; accessed 21-Jan-2021].

[2] Facebook Inc. *Facebook Reports First Quarter 2020 Results*. `https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-First-Quarter-2020-Results/default.aspx`. [Online; accessed 21-Jan-2021].

[3] Twitter Inc. *Twitter Company Info*. `https://investor.twitterinc.com/contact/faq/default.aspx`. [Online; accessed 21-Jan-2021].

[4] SimilarWeb. *Twitter Traffic Overview*. `https://www.similarweb.com/website/twitter.com/#overview`. [Online; accessed 21-Jan-2021].

[5] CL Sneison. „Qualitative and mixed methods social media research“. In: *International Journal of Qualitative Methods* 15.1 (2016), pages 1–15.

[6] Kathleen Beullens and Adriaan Schepers. „Display of alcohol use on Facebook: a content analysis“. In: *Cyberpsychology, Behavior, and Social Networking* 16.7 (2013), pages 497–503.

[7] Sean D Young. „Behavioral insights on big data: using social media for predicting biomedical outcomes“. In: *Trends in microbiology* 22.11 (2014), pages 601–602.

[8] Sarah Catrin Titgemeyer and Christian Patrick Schaaf. „Facebook Support Groups for Rare Pediatric Diseases: Quantitative Analysis“. In: *JMIR pediatrics and parenting* 3.2 (2020), e21694.

[9] Elaf Ali Alsisi, Ahmed Al-Ashaab, and Wadhah Ahmed Abualfaraa. „The Development of a Smart Health Awareness Message Framework Based on the Use of Social Media: Quantitative Study“. In: *Journal of medical Internet research* 22.7 (2020), e16212.

[10] Xin Zhang, Liang Ma, Bo Xu, and Feng Xu. „How social media usage affects employees' job satisfaction and turnover intention: An empirical study in China“. In: *Information & Management* 56.6 (2019), page 103136.

[11] Robert E Wilson, Samuel D Gosling, and Lindsay T Graham. „A review of Facebook research in the social sciences“. In: *Perspectives on psychological science* 7.3 (2012), pages 203–220.

[12] Raad Bin Tareaf, Philipp Berger, Patrick Hennig, and Christoph Meinel. „Personality Exploration System for Online Social Networks: Facebook Brands As a Use Case“. In: *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*. IEEE. 2018, pages 301–309.

[13] Robin Effing, Jos Van Hillegersberg, and Theo Huibers. „Social media and political participation: are Facebook, Twitter and YouTube democratizing our political systems?“ In: *International conference on electronic participation*. Springer. 2011, pages 25–35.

[14] Ahmed Shams, Raad Bin Tareaf, Jan Renz, and Christoph Meinel. „Smart MOOC-Social Computing for Learning and Knowledge Sharing.“ In: *CSEDU (2)*. 2018, pages 391–396.

[15] David J Stillwell and Michal Kosinski. „myPersonality project: Example of successful utilization of online social networks for large-scale social research“. In: *American Psychologist* 59.2 (2004), pages 93–104.

[16] Michal Kosinski, David Stillwell, and Thore Graepel. „Private traits and attributes are predictable from digital records of human behavior“. In: *Proceedings of the national academy of sciences* 110.15 (2013), pages 5802–5805.

[17] Raad Bin Tareaf, Seyed Ali Alhosseini, Philipp Berger, Patrick Hennig, and Christoph Meinel. „Towards Automatic Personality Prediction Using Facebook Likes Metadata“. In: *14th IEEE International Conference on Intelligent Systems and Knowledge Engineering, ISKE 2019, Dalian, China, November 14-16, 2019*. Edited by Li Zou, Lingling Fang, Bo Fu, and Panpan Niu. IEEE, 2019, pages 714–719. DOI: 10.1109/ISKE47853.2019.9170375. URL: https://doi.org/10.1109/ISKE47853.2019.9170375.

[18] Bin Tareaf Raad, Berger Philipp, Hennig Patrick, and Meinel Christoph. „Aseds: Towards automatic social emotion detection system using facebook reactions“. In: *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*. IEEE. 2018, pages 860–866.

[19] Raad Bin Tareaf, Seyed Ali Alhosseini, and Christoph Meinel. „Facial-Based Personality Prediction Models for Estimating Individuals Private Traits“. In: *2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*. IEEE. 2019, pages 1586–1594.

[20] Raad Bin Tareaf, Philipp Berger, Patrick Hennig, and Christoph Meinel. „Cross-platform personality exploration system for online social networks: Facebook vs. Twitter“. In: *Web Intelligence*. Preprint. IOS Press. 2020, pages 1–17.

[21] Raad Bin Tareaf, Seyed Ali Alhosseini, and Christoph Meinel. „Does Personality Evolve? A Ten-Years Longitudinal Study from Social Media Platforms“. In: *2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*. IEEE. 2020, pages 1205–1213.

[22] Raad Bin Tareaf, Philipp Berger, Patrick Hennig, and Christoph Meinel. „Malicious behaviour identification in online social networks“. In: *IFIP International Conference on Distributed Applications and Interoperable Systems*. Springer. 2018, pages 18–25.

[23] Seyed Ali Alhosseini, Raad Bin Tareaf, Pejman Najafi, and Christoph Meinel. „Detect Me If You Can: Spam Bot Detection Using Inductive Representation Learning“. In: *Companion Proceedings of The 2019 World Wide Web Conference*. 2019, pages 148–153.

[24] Raad Bin Tareaf, Philipp Berger, Patrick Hennig, Sebastian Koall, Jan Kohstall, and Christoph Meinel. „Information Propagation Speed and Patterns in Social Networks: A Case Study Analysis of German Tweets.“ In: *JCP* 13.7 (2018), pages 761–770.

[25] Raad Bin Tareaf, Philipp Berger, Patrick Hennig, Jaeyoon Jung, and Christoph Meinel. „Identifying audience attributes: predicting age, gender and personality for enhanced article writing“. In: *Proceedings of the 2017 International Conference on Cloud and Big Data Computing*. 2017, pages 79–88.

[26] Hanadi Traifeh, Raad Bin Tareaf, and Christoph Meinel. „E-Learning Experiences from the Arab World“. In: *In proceedings of International Conference on Advanced Research in Education, Sorbonne University, Paris, France*. 2019.

[27] Seyed Ali Alhosseini, Raad Bin Tareaf, and Christoph Meinel. „Engaging with Tweets: The Missing Dataset On Social Media“. In: *Proceedings of the Recommender Systems Challenge 2020*. 2020, pages 34–37.

[28] Daniel J Ozer and Veronica Benet-Martinez. „Personality and the prediction of consequential outcomes“. In: *Annu. Rev. Psychol.* 57 (2006), pages 401–421.

[29] John M Digman. „Personality structure: Emergence of the five-factor model“. In: *Annual review of psychology* 41.1 (1990), pages 417–440.

[30] Robert R McCrae and Paul T Costa. *Personality in adulthood: A five-factor theory perspective*. Guilford Press, 2003.

[31] Lewis R Goldberg, John A Johnson, Herbert W Eber, Robert Hogan, Michael C Ashton, C Robert Cloninger, and Harrison G Gough. „The international personality item pool and the future of public-domain personality measures“. In: *Journal of Research in personality* 40.1 (2006), pages 84–96.

[32] Robert R McCrae and Oliver P John. „An introduction to the five-factor model and its applications“. In: *Journal of personality* 60.2 (1992), pages 175–215.

[33] Klaus Rainer Scherer and Howard Giles. *Social markers in speech.* Volume 6. Cambridge University Press, 1979.

[34] James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. „Psychological aspects of natural language use: Our words, our selves“. In: *Annual review of psychology* 54.1 (2003), pages 547–577.

[35] Jon Oberlander and Alastair J Gill. „Language with character: A stratified corpus comparison of individual differences in e-mail communication“. In: *Discourse Processes* 42.3 (2006), pages 239–270.

[36] Jon Oberlander and Scott Nowson. „Whose thumb is it anyway? Classifying author personality from weblog text“. In: *Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions.* 2006, pages 627–634.

[37] Philip J Stone, Dexter C Dunphy, and Marshall S Smith. „The general inquirer: A computer approach to content analysis.“ In: (1966).

[38] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. „Recursive deep models for semantic compositionality over a sentiment treebank“. In: *Proceedings of the 2013 conference on empirical methods in natural language processing.* 2013, pages 1631–1642.

[39] Syeda Asra and DC Shubhangi. „Personality Trait Identification Using Unconstrained Cursive and Mood Invariant Handwritten Text“. In: *International Journal of Education and Management Engineering* 5.5 (2015), page 20.

[40] Firoj Alam, Evgeny A Stepanov, and Giuseppe Riccardi. „Personality traits recognition on social network-facebook“. In: *Proceedings of the International AAAI Conference on Web and Social Media.* Volume 7. 1. 2013.

[41] Srilakshmi Bharadwaj, Srinidhi Sridhar, Rahul Choudhary, and Ramamoorthy Srinath. „persona traits identification based on myers-briggs type indicator (MBTI)-a text classification approach“. In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI).* IEEE. 2018, pages 1076–1082.

[42] Dilini Sewwandi, Kusal Perera, Sajith Sandaruwan, Oshani Lakchani, Anupiya Nugaliyadde, and Samantha Thelijjagoda. „Linguistic features based personality recognition using social media data“. In: *2017 6th National Conference on Technology and Management (NCTM)*. IEEE. 2017, pages 63–68.

[43] Fabio Celli. „Unsupervised personality recognition for social network sites“. In: *Proc. of sixth international conference on digital society*. 2012, pages 59–62.

[44] Fabio Celli and Luca Rossi. „The role of emotional stability in Twitter conversations“. In: *Proceedings of the workshop on semantic analysis in social media*. 2012, pages 10–17.

[45] Xiangguo Sun, Bo Liu, Qing Meng, Jiuxin Cao, Junzhou Luo, and Hongzhi Yin. „Group-level personality detection based on text generated networks“. In: *World Wide Web* (2019), pages 1–20.

[46] Robin SS Kramer, James E King, and Robert Ward. „Identifying personality from the static, nonexpressive face in humans and chimpanzees: evidence of a shared system for signaling personality“. In: *Evolution and Human Behavior* 32.3 (2011), pages 179–185.

[47] Shuotian Bai, Tingshao Zhu, and Li Cheng. „Big-five personality prediction based on user behaviors at social network sites“. In: *arXiv preprint arXiv:1204.4809* (2012).

[48] RK Hernandez and I Scott. „Predicting Myers-Briggs type indicator with text“. In: *31st Conference on Neural Information Processing Systems (NIPS 2017)*. 2017.

[49] Navonil Majumder, Soujanya Poria, Alexander Gelbukh, and Erik Cambria. „Deep learning-based document modeling for personality detection from text“. In: *IEEE Intelligent Systems* 32.2 (2017), pages 74–79.

[50] Fei Liu, Julien Perez, and Scott Nowson. „A language-independent and compositional model for personality trait recognition from short texts“. In: *arXiv preprint arXiv:1610.04345* (2016).

[51] Jennifer Golbeck and Derek Hansen. „Computing political preference among twitter followers“. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2011, pages 1105–1108.

[52] Daniele Quercia, Michal Kosinski, David Stillwell, and Jon Crowcroft. „Our twitter profiles, our selves: Predicting personality with twitter“. In: *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE. 2011, pages 180–185.

[53] Dejan Markovikj, Sonja Gievska, Michal Kosinski, and David Stillwell. „Mining facebook data for predictive personality modeling“. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Volume 7. 1. 2013.

[54] Ben Verhoeven, Walter Daelemans, and Tom De Smedt. „Ensemble methods for personality recognition“. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Volume 7. 1. 2013.

[55] François Mairesse, Marilyn A Walker, Matthias R Mehl, and Roger K Moore. „Using linguistic cues for the automatic recognition of personality in conversation and text“. In: *Journal of artificial intelligence research* 30 (2007), pages 457–500.

[56] Francisco Iacobelli, Alastair J Gill, Scott Nowson, and Jon Oberlander. „Large scale personality classification of bloggers“. In: *international conference on affective computing and intelligent interaction*. Springer. 2011, pages 568–577.

[57] Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. „Predicting personality from twitter“. In: *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE. 2011, pages 149–156.

[58] Hans Christian, Derwin Suhartono, Andry Chowanda, and Kamal Z Zamli. „Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging“. In: *Journal of Big Data* 8.1 (2021), pages 1–20.

[59] Zhancheng Ren, Qiang Shen, Xiaolei Diao, and Hao Xu. „A sentiment-aware deep learning approach for personality detection from text“. In: *Information Processing & Management* 58.3 (2021), page 102532.

[60] H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Lukasz Dziurzynski, Stephanie M Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin EP Seligman, et al. „Personality, gender, and age in the language of social media: The open-vocabulary approach“. In: *PloS one* 8.9 (2013), e73791.

[61] Michal Kosinski, Sandra C Matz, Samuel D Gosling, Vesselin Popov, and David Stillwell. „Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines.“ In: *American Psychologist* 70.6 (2015), page 543.

[62] Paul T Costa Jr and Robert R McCrae. *The Revised NEO Personality Inventory (NEO-PI-R)*. Sage Publications, Inc, 2008.

[63] Yla R Tausczik and James W Pennebaker. „The psychological meaning of words: LIWC and computerized text analysis methods“. In: *Journal of language and social psychology* 29.1 (2010), pages 24–54.

[64] Jean-Philippe Vert, Koji Tsuda, and Bernhard Schölkopf. „A primer on kernel methods“. In: *Kernel methods in computational biology* 47 (2004), pages 35–70.

[65] Tera D Letzring. „The good judge of personality: Characteristics, behaviors, and observer accuracy". In: *Journal of research in personality* 42.4 (2008), pages 914–932.

[66] Yoram Bachrach, Michal Kosinski, Thore Graepel, Pushmeet Kohli, and David Stillwell. „Personality and patterns of Facebook usage". In: *Proceedings of the 4th annual ACM web science conference*. 2012, pages 24–32.

[67] Wu Youyou, Michal Kosinski, and David Stillwell. „Computer-based personality judgments are more accurate than those made by humans". In: *Proceedings of the National Academy of Sciences* 112.4 (2015), pages 1036–1040.

[68] Menasha Thilakaratne, Ruvan Weerasinghe, and Sujan Perera. „Knowledge-driven approach to predict personality traits by leveraging social media data". In: *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*. IEEE. 2016, pages 288–295.

[69] Chris Pool and Malvina Nissim. „Distant supervision for emotion detection using Facebook reactions". In: *arXiv preprint arXiv:1611.02988* (2016).

[70] Soumaya Chaffar and Diana Inkpen. „Using a heterogeneous dataset for emotion analysis in text". In: *Canadian conference on artificial intelligence*. Springer. 2011, pages 62–67.

[71] Lillian Gray. „Gender Bias Detection Using Facebook Reactions". In: (2020).

[72] Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. „Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm". In: *arXiv preprint arXiv:1708.00524* (2017).

[73] Florian Krebs, Bruno Lubascher, Tobias Moers, Pieter Schaap, and Gerasimos Spanakis. „Social emotion mining techniques for Facebook posts reaction prediction". In: *arXiv preprint arXiv:1712.03249* (2017).

[74] Angelo Basile, Tommaso Caselli, and Malvina Nissim. „Predicting Controversial News Using Facebook Reactions." In: *CLiC-it*. 2017.

[75] Carlo Strapparava and Rada Mihalcea. „Semeval-2007 task 14: Affective text". In: *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*. 2007, pages 70–74.

[76] Carlo Strapparava and Rada Mihalcea. „Learning to identify emotions in text". In: *Proceedings of the 2008 ACM symposium on Applied computing*. 2008, pages 1556–1560.

[77] Sunghwan Mac Kim, Alessandro Valitutti, and Rafael A Calvo. „Evaluation of unsupervised emotion models to textual affect recognition“. In: *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. 2010, pages 62–70.

[78] Andrew McCallum, Kamal Nigam, et al. „A comparison of event models for naive bayes text classification“. In: *AAAI-98 workshop on learning for text categorization*. Volume 752. 1. Citeseer. 1998, pages 41–48.

[79] Jason D Rennie, Lawrence Shih, Jaime Teevan, and David R Karger. „Tackling the poor assumptions of naive bayes text classifiers“. In: *Proceedings of the 20th international conference on machine learning (ICML-03)*. 2003, pages 616–623.

[80] Ebba Cecilia Ovesdotter Alm. *Affect in* text and speech*. Citeseer, 2008.

[81] Alexander Kachur, Evgeny Osin, Denis Davydov, Konstantin Shutilov, and Alexey Novokshonov. „Assessing the Big Five personality traits using real-life static facial images“. In: *Scientific reports* 10.1 (2020), pages 1–11.

[82] Alireza Souri, Shafigheh Hosseinpour, and Amir Masoud Rahmani. „Personality classification based on profiles of social networks' users and the five-factor model of personality“. In: *Human-centric Computing and Information Sciences* 8.1 (2018), pages 1–15.

[83] Raïssa Yapan Dougnon, Philippe Fournier-Viger, and Roger Nkambou. „Inferring user profiles in online social networks using a partial social graph“. In: *Canadian Conference on Artificial Intelligence*. Springer. 2015, pages 84–99.

[84] Leqi Liu, Daniel Preotiuc-Pietro, Zahra Riahi Samani, Mohsen E Moghaddam, and Lyle Ungar. „Analyzing personality through social media profile picture choice“. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Volume 10. 1. 2016.

[85] Bernhard Fink, Nick Neave, John T Manning, and Karl Grammer. „Facial symmetry and the 'big-five'personality factors“. In: *Personality and individual differences* 39.3 (2005), pages 523–529.

[86] F Richter. „Snapchat crowned number 1 by American teens“. In: *Statista*. 2017.

[87] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. „Predicting personality traits with instagram pictures“. In: *Proceedings of the 3rd Workshop on Emotions and Personality in Personalized Systems 2015*. 2015, pages 7–10.

[88] I Megvii. *Face++ research toolkit*. 2013.

[89] Marcin Skowron, Marko Tkalčič, Bruce Ferwerda, and Markus Schedl. „Fusing social media cues: personality prediction from twitter and instagram“. In: *Proceedings of the 25th international conference companion on world wide web*. 2016, pages 107–108.

[90] Jie Nie, Lei Huang, Peng Cui, Zhen Li, Yan Yan, Zhiqiang Wei, and Wenwu Zhu. „Social media profiler: Inferring your social media personality from visual attributes in portrait“. In: *Pacific Rim Conference on Multimedia*. Springer. 2016, pages 640–649.

[91] Marco Cristani, Alessandro Vinciarelli, Cristina Segalin, and Alessandro Perina. „Unveiling the multimedia unconscious: Implicit cognitive processes and multimedia content analysis“. In: *Proceedings of the 21st ACM international conference on Multimedia*. 2013, pages 213–222.

[92] Chanchal Suman, Sriparna Saha, Aditya Gupta, Saurabh Kumar Pandey, and Pushpak Bhattacharyya. „A multi-modal personality prediction system“. In: *Knowledge-Based Systems* 236 (2022), page 107715.

[93] Gregory J Meyer, Stephen E Finn, Lorraine D Eyde, Gary G Kay, Kevin L Moreland, Robert R Dies, Elena J Eisman, Tom W Kubiszyn, and Geoffrey M Reed. „Psychological testing and psychological assessment: A review of evidence and issues.“ In: *American psychologist* 56.2 (2001), page 128.

[94] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. „Unsupervised learning“. In: *The elements of statistical learning*. Springer, 2009, pages 485–585.

[95] Jerome H Friedman. „Greedy function approximation: a gradient boosting machine“. In: *Annals of statistics* (2001), pages 1189–1232.

[96] Yoav Freund and Robert E Schapire. „A decision-theoretic generalization of on-line learning and an application to boosting“. In: *Journal of computer and system sciences* 55.1 (1997), pages 119–139.

[97] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

[98] Philip J Corr and Gerald Matthews. *The Cambridge handbook of personality psychology*. Cambridge University Press, 2020.

[99] Alan E Kazdin, American Psychological Association, et al. *Encyclopedia of psychology*. Volume 2. American Psychological Association Washington, DC, 2000.

[100] Brent W Roberts, Nathan R Kuncel, Rebecca Shiner, Avshalom Caspi, and Lewis R Goldberg. „The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes“. In: *Perspectives on Psychological science* 2.4 (2007), pages 313–345.

[101] Margaret L Kern and Howard S Friedman. „Do conscientious individuals live longer? A quantitative review.“ In: *Health psychology* 27.5 (2008), page 505.

[102] Ted Schwaba, Richard W Robins, Emily Grijalva, and Wiebke Bleidorn. „Does Openness to Experience matter in love and work? Domain, facet, and developmental evidence from a 24-year longitudinal study“. In: *Journal of Personality* 87.5 (2019), pages 1074–1092.

[103] Jaap JA Denissen, Wiebke Bleidorn, Marie Hennecke, Maike Luhmann, Ulrich Orth, Jule Specht, and Julia Zimmermann. „Uncovering the power of personality to shape income“. In: *Psychological science* 29.1 (2018), pages 3–13.

[104] Jenny Wagner, Michael Becker, Oliver Lüdtke, and Ulrich Trautwein. „The first partnership experience and personality development: A propensity score matching study in young adulthood“. In: *Social Psychological and Personality Science* 6.4 (2015), pages 455–463.

[105] Oya Celiktutan, Evangelos Sariyanidi, and Hatice Gunes. „Let me tell you about your personality!†: Real-time personality prediction from nonverbal behavioural cues“. In: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. Volume 1. IEEE. 2015, pages 1–1.

[106] Ludwig Küster, Carola Trahms, and Jan-Niklas Voigt-Antons. „Predicting personality traits from touchscreen based interactions“. In: *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE. 2018, pages 1–6.

[107] Brent W Roberts, Richard W Robins, Kali H Trzesniewski, and Avshalom Caspi. „Personality trait development in adulthood“. In: *Handbook of the life course*. Springer, 2003, pages 579–595.

[108] Richard W Robins, R Chris Fraley, Brent W Roberts, and Kali H Trzesniewski. „A longitudinal study of personality change in young adulthood“. In: *Journal of personality* 69.4 (2001), pages 617–640.

[109] Daniel K Mroczek and Avron Spiro III. „Modeling intraindividual change in personality traits: Findings from the Normative Aging Study“. In: *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 58.3 (2003), P153–P165.

[110] Stephen E Finn. „Stability of personality self-ratings over 30 years: Evidence for an age/cohort interaction.“ In: *Journal of Personality and Social Psychology* 50.4 (1986), page 813.

[111] James J Conley. „The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion“. In: *Personality and Individual Differences* 5.1 (1984), pages 11–25.

[112] James J Conley. „Longitudinal stability of personality traits: A multitrait–multimethod–multioccasion analysis.“ In: *Journal of personality and social psychology* 49.5 (1985), page 1266.

[113] Richard M Lerner, Christina Theokas, and Deborah L Bobek. „Concepts and Theories of Human Development: Historical and Contemporary Dimensions.“ In: (2005).

[114] Orville Gilbert Brim and Jerome Kagan. *Constancy and change in human development*. Harvard University Press, 1980.

[115] Ravenna Helson, Virginia SY Kwan, Oliver P John, and Constance Jones. „The growing evidence for personality change in adulthood: Findings from research with personality inventories“. In: *Journal of research in personality* 36.4 (2002), pages 287–306.

[116] Gerard Saucier, Amber Gayle Thalmayer, and Tarik S Bel-Bahar. „Human attribute concepts: Relative ubiquity across twelve mutually isolated languages.“ In: *Journal of Personality and Social Psychology* 107.1 (2014), page 199.

[117] Wiebke Bleidorn, Patrick L Hill, Mitja D Back, Jaap JA Denissen, Marie Hennecke, Christopher J Hopwood, Markus Jokela, Christian Kandler, Richard E Lucas, Maike Luhmann, et al. „The policy relevance of personality traits.“ In: *American Psychologist* 74.9 (2019), page 1056.

[118] Robert B Ewen et al. *An introduction to theories of personality*. Psychology Press, 2014.

[119] Lin Qiu, Han Lin, Jonathan Ramsay, and Fang Yang. „You are what you tweet: Personality expression and perception on Twitter“. In: *Journal of research in personality* 46.6 (2012), pages 710–718.

[120] Jacopo Staiano, Bruno Lepri, Nadav Aharony, Fabio Pianesi, Nicu Sebe, and Alex Pentland. „Friends don't lie: inferring personality traits from social network structure“. In: *Proceedings of the 2012 ACM conference on ubiquitous computing*. 2012, pages 321–330.

[121] Ravenna Helson, Constance Jones, and Virginia SY Kwan. „Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples.“ In: *Journal of personality and social psychology* 83.3 (2002), page 752.

[122] Paul T Costa Jr and Robert R McCrae. „Set like plaster? Evidence for the stability of adult personality.“ In: (1994).

[123] Franz J Neyer and Jens B Asendorpf. „Personality–relationship transaction in young adulthood.“ In: *Journal of personality and social psychology* 81.6 (2001), page 1190.

[124] Norman Haan. *Present and Past in Middle Life*. Academic Press, 1981.

[125] Paul T Costa and Robert R McCrae. „Objective personality assessment“. In: *The clinical psychology of aging*. Springer, 1978, pages 119–143.

[126] Ilene C Siegler, Linda K George, and Morris A Okun. „Cross-sequential analysis of adult personality.“ In: *Developmental Psychology* 15.3 (1979), page 350.

[127] Rodica Ioana Damian, Marion Spengler, Andreea Sutu, and Brent W Roberts. „Sixteen going on sixty-six: A longitudinal study of personality stability and change across 50 years.“ In: *Journal of Personality and Social Psychology* 117.3 (2019), page 674.

[128] Golnoosh Farnadi, Geetha Sitaraman, Shanu Sushmita, Fabio Celli, Michal Kosinski, David Stillwell, Sergio Davalos, Marie-Francine Moens, and Martine De Cock. „Computational personality recognition in social media“. In: *User modeling and user-adapted interaction* 26.2 (2016), pages 109–142.

[129] Margeret Hall and Simon Caton. „Am I who I say I am? Unobtrusive self-representation and personality recognition on Facebook“. In: *PloS one* 12.9 (2017), e0184417.

[130] Dong Nie, Zengda Guan, Bibo Hao, Shuotian Bai, and Tingshao Zhu. „Predicting personality on social media with semi-supervised learning“. In: *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. Volume 2. IEEE. 2014, pages 158–165.

[131] Shuotian Bai, Sha Yuan, Bibo Hao, and Tingshao Zhu. „Predicting personality traits of microblog users“. In: *Web Intelligence and Agent Systems: An International Journal* 12.3 (2014), pages 249–265.

[132] Kokil Jaidka, Sharath Guntuku, and Lyle Ungar. „Facebook versus Twitter: Differences in Self-Disclosure and Trait Prediction“. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Volume 12. 1. 2018.

[133] Jennifer L Aaker. „Dimensions of brand personality“. In: *Journal of marketing research* 34.3 (1997), pages 347–356.

[134] Maggie Geuens, Bert Weijters, and Kristof De Wulf. „A new measure of brand personality“. In: *International journal of research in marketing* 26.2 (2009), pages 97–107.

[135] Avshalom Caspi and Ellen S Herbener. „Continuity and change: Assortative marriage and the consistency of personality in adulthood.“ In: *Journal of personality and social psychology* 58.2 (1990), page 250.

[136] Avshalom Caspi, Brent W Roberts, and Rebecca L Shiner. „Personality development: Stability and change“. In: *Annu. Rev. Psychol.* 56 (2005), pages 453–484.

[137] Robert R McCrae, Paul T Costa Jr, Fritz Ostendorf, Alois Angleitner, Martina Hřebíčková, Maria D Avia, Jesús Sanz, Maria L Sanchez-Bernardos, M Ersin Kusdil, Ruth Woodfield, et al. „Nature over nurture: temperament, personality, and life span development.“ In: *Journal of personality and social psychology* 78.1 (2000), page 173.

[138] University of Cambridge. *ApplyMagicSauce*. `https://applymagicsauce.com/`.