Utrecht University

Department of Information and Computing Science

Applied Data Science master thesis

Beyond Words: Unmasking the Relationship between Emotions and Personality Traits

Candidate: Pinelopi Theotoki

Supervisor:

Anastasia Giachanou

Co-supervisor:

Qixiang Fang

June 29, 2023

Abstract

Technology is evolving quickly, and Text-based Personality Computing (TPC) and automatic emotion assessment have attracted the interest of researchers. This study explores their effectiveness as an addition to traditional methods in personality and emotion assessment. The relationship between personality and emotion using automated methods has not been thoroughly examined. Thus, this project aims to investigate the correlations between personality traits and emotions in automated detection. To accomplish this aim we used two datasets with annotated emotions and Big Five personality traits, and we implemented supervised methods to predict personality traits and emotions. We have tried different models including SVM, Logistic Regression and variants of BERT. The utilized evaluation metrics are accuracy, precision, recall and F1-score. Correlations between emotions and traits are computed using the Pearson correlation coefficient and visualized using heatmaps. Additionally, the alignment between the correlations from the annotated and the predicted data is assessed using the Pearson correlation coefficient, RMSE, MAE, and the Wilcoxon signed rank test. Moreover, the correlations from the original data are compared to established theories. The findings reveal that the best predictions were obtained by transformer-based models. Specifically, DistilBERT performs the best in predicting Extraversion and Stability while DeBERTa outperforms other models in predicting emotions, Conscientiousness, Openness, and Agreeableness. The correlations from the predicted data are similar to those from the original data but their range is different. Some of the personality-emotion pairs (e.g., joy-Conscientiousness, anger-Conscientiousness) align with established theories while some others do not (e.g., Neutralism-All personality traits). Overall, this research contributes to the understanding and applicability of TPC in personality and emotion assessment.

Table of Contents

Abstract 2
List of Tables 4
List of Figures 5
1.Introduction 6
2. Literature Review 8
2.1 Personality Traits 8
2.2 Automated personality and emotion prediction 8
2.3 Relationship between personality and emotions 9
3. Datasets 11
3.1 Description of the datasets 11
3.2 Preparation of the data 15
4. Methodology 17
4.1 Description of the methods 17
4.2 Experimental settings 19
5. Results and Discussion 20
5.1 Results on predictions 20
5.2 Results on correlations 21
5.3 Results on alignment with theories 22
5.4 Discussion 24
6. Conclusion and Future work 26
6.1 Ethical and legal considerations 26
6.2 Conclusion 26
References 28
Appendices 34
Appendix A: Emotion prediction 34
Appendix B: Personality prediction 35
Appendix C: Models with continuous outcomes 37
Appendix D: Heatmaps with correlations 38
Appendix E: Comparisons of correlations 39
Appendix F: F1-Scores for emotion prediction 40

List of Tables

Table 1: Example data structure of WASSA 2022 12
Table 2: Summary of text characteristics for WASSA 2022 13
Table 3:Example data structure of PELD 14
Table 4: Summary of text characteristics for PELD 15
Table 5: Train/val/test sets 15
Table 6: Predictions' results 20
Table 7: Comparison of correlations from annotated data and from best predictions 22
Table 8: Alignment with theories 24
Table 9: Emotion prediction, WASSA 2022 34
Table 10: Emotion prediction, PELD 34
Table 11: Conscientiousness prediction, WASSA 2022 35
Table 12: Openness prediction, WASSA 2022 35
Table 13: Extraversion prediction, WASSA 2022 35
Table 14: Stability prediction, WASSA 2022 36
Table 15: Agreeableness prediction, WASSA 2022 36
Table 16: Linear Regression, WASSA 2022 37
Table 17: Neural Network model, WASSA 2022 37
Table 18: Comparisons of correlations 39
Table 19:F1-Scores for emotion prediction, WASSA 2022 40
Table 20: F1-Scores for emotion prediction, PELD 40

List of Figures

Figure 1:Emotions' distribution, WASSA 2022 12
Figure 2: Big Five Traits' distributions, WASSA 2022 13
Figure 3: Emotion's and Speakers' Distributions, PELD 15
Figure 4: The model architecture of DeBERTa (Scale up DeBERTa (iclr.cc)) 18
Figure 5: Heatmaps of correlations from original data and best predictions 22
Figure 6: Correlations of annotated data for WASSA 2022 and PELD datasets 23
Figure 7: Heatmaps with correlations 38

1. Introduction

Personality traits and emotions are two constructs that are fundamental to an individual's psychological constitution (Plomin et al., 2014). Personality traits are characterized by patterns of thoughts, feelings, and behaviors that are relatively consistent over time and across different situations (Vinciarelli et al., 2014). On the other hand, emotions are temporary, subjective experiences that are typically triggered by internal or external stimuli and can range from pleasant to unpleasant. They can vary widely in intensity and duration and play a critical role in social behavior and decision-making (Dalgleish et al., 2000).

The above descriptions indicate that personality traits and emotions share a lot of similarities and a few differences. As mentioned by Revelle et al., (2009) “A helpful analogy is to consider that personality is to emotion as climate is to weather”. Although personality traits describe more stable characteristics and emotions more temporary feelings, both are influenced by a combination of biological and environmental factors (Reisenzein et al., 2020).

Understanding emotions and personality is important and helpful in various societal (e.g., school success) and scientific (e.g., neuroscience) domains. Within the scientific world, an exploration of emotions finds utility in fields like neuroscience (Adolphs, 2017). Additionally, the application of emotional knowledge extends to practical aspects, such as facilitating early school success among preschoolers. Denham et al., (2015) found out that age-appropriate emotion knowledge can make a socially competent child pay more attention to academic tasks and communicate better with peers. Additionally, the knowledge of personality yields its own set of practical implications. The correct interpretation of personality can enhance individuals in making informed decisions about suitable career paths (Kern et al., 2019). At the same time, clinical psychologists can use personality to better understand psychological disorders (Khan et al., 2005). Also, the identification of changes in personality can contribute to health-related matters, such as the early diagnosis of Alzheimer's disease (Robins Wahlin et al., 2011). It is vital to assess personality and emotions since their correct interpretation can aid research in many fields.

Traditionally, the assessment of personality traits has relied on questionnaires that were filled out by individuals themselves or others, a process that can be time-consuming and costly (Fang et al., 2022). However, with the emergence of user-generated data such as texts, audio, and images, coupled with advancements in machine learning algorithms, automatic personality assessment has become increasingly common as an efficient and cost-effective alternative (Phan et al., 2021). This automated approach is referred to as Personality Computing (PC). Our focus is specifically on Text-based Personality Computing (TPC), which requires the application of PC techniques to text data (Mushtaq et al. 2023).

TPC has benefits as well as limitations. The several advances of automated approaches include the potential for large-scale data analysis because of increased computing power and data storage (Vinciarelli et al., 2014). However, challenges such as privacy concerns, data and measurement quality should be carefully considered (Fang et al., 2022). We are aware from questionnaire methods for personality and emotion assessment (Cattell, 1973) that correlations between personality traits and emotions exist, but in TPC researchers rarely report them. Therefore, this research paper aims to examine TPC by analyzing the correlations between personality traits and emotions. Moreover, we hope that the findings will result to a better understanding of automated personality and emotion detection that can help in the development of PC and its applications in various fields, including psychology, social media analysis, and other areas where personality assessment is essential.

To complete this aim, we focus on exploring the relationship between personality traits and emotions in TPC answering the following research question: “What are the correlations between personality traits and emotions in automated personality and emotion detection, and how well do they align with established theories?” To answer these questions, some sub questions need to be answered first.

[RQ 1]: *How accurately can automated models predict personality traits and emotions from text data?*

[RQ 2]: *How related are the correlations between annotated and predicted personality traits and emotions?*

[RQ 3]: *How consistent are the correlations from annotated data with the correlations reported in established personality and emotion theories?*

The research questions are addressed through a series of steps. The initial phase involves the collection and preprocessing of data from two datasets: WASSA 2022[1] and PELD[2]. The personality traits of the individuals are identified utilizing the widely recognized Big Five taxonomy, encompassing the dimensions of Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Goldberg, 1982).

To address **RQ1** we applied supervised machine learning algorithms to predict both personality traits and emotions. The implemented algorithms include Distilled Bidirectional Encoder Representations from Transformers (DistilBERT), Robustly Optimized BERT Pretraining Approach (RoBERTa), Decoding-enhanced BERT with Disentangled Attention (DeBERTa), Support Vector Machines (SVM), and Logistic Regression and the predictions were assessed using Accuracy, Precision, Recall, and F1-score. Subsequently, to answer **RQ2** we computed the correlations between predictions of traits and emotions. Then, to compare the correlations from predictions to the ones obtained from annotated data we used evaluation metrics including Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). This comparative analysis offers insights into the effectiveness of the models in predicting emotions and personality traits. Finally, to address **RQ3** the correlations from annotated data are compared with widely accepted theories reporting the relationship between Big Five traits and emotions to examine their alignment.

The results of the research demonstrate that DeBERTa performs the best in detecting all emotions. Similarly, for Conscientiousness, Agreeableness and Openness, DeBERTa outperforms the rest of the models, while for Extraversion and Stability DistilBERT does. It is worth mentioning that all the models predict emotions and personality traits a lot better than random. The correlations between predicted emotions and personality traits agree with the correlations from annotated data because the Pearson correlation coefficient was above 0.5 in most cases (DeBERTa: 0.725, RoBERTa: 0.508, DistilBERT: 0.504). Also, the correlations from annotated data agree with established theories for certain emotion-personality trait pairs (e.g., anger-conscientiousness, surprise-agreeableness). However, for other pairs, there is a lack of alignment between the correlations and established theories.

This research paper is organized into various sections, which are as follows: Section 2 presents the related work about Big Five personality trait and emotion detection and correlations between them, while Section 3 focuses on the two datasets that are used. Next is Section 4, which reviews the implemented methods. Section 5 analyzes the results and Section 6 concludes the paper and discusses the ethical considerations.

---

[1] CodaLab - Competition (upsaclay.fr)

[2] GitHub - preke/PELD: Personality EmotionLine Dataset

2. Literature Review

This Section discusses the existing theory about Big Five personality traits in 2.1 and the literature about automated systems for personality and emotion prediction in 2.2. Section 2.3 presents widely accepted theories about the relationship between personality and emotions.

2.1 Personality Traits

According to (John et al., 1999), personality taxonomy serves as a framework for describing personality traits. In order to predict personality, we must first define potential personality traits. The Big Five personality taxonomy is used because it is realistic and accurate (Fang et al., 2022). Big Five is based on independent studies (Goldberg, 1982; Tupes et al., 1992; Costa et al., 1992) and identifies five significant personality dimensions: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

1. The personality trait of **Openness** refers to how open someone is to a diverse range of experiences or how abstractly they think about concrete issues. Individuals characterized by high levels of Openness often hold unconventional values and exhibit creative thinking abilities.
2. **Conscientiousness** describes a person's level of self-discipline, organizational skills, and goal-directed behavior. Those who exhibit high levels of Conscientiousness tend to excel in planning and prefer structured approaches over spontaneous decision-making.
3. **Extraversion** denotes a personality trait characterized by an individual's sociability, energy levels, and warmth in interpersonal interactions. Individuals with high levels of Extraversion have a tendency to be outgoing, talkative, and actively engage with others.
4. **Agreeableness** pertains to a personality trait that reflects an individual's propensity for kindness, sympathy, and cooperation. Those with high levels of Agreeableness incline to maintain friendly dispositions towards others, to be helpful and less competitive.
5. **Neuroticism** describes a personality trait associated with emotional instability, nervousness, distress, and fearfulness. Individuals scoring high on Neuroticism are more likely to experience heightened levels of worry (Abdullahi et al., 2020).

2.2 Automated personality and emotion prediction

In today's digital age, with the huge availability of online data, it has become necessary to understand the complex interplay between different personality traits of individuals and emotional experiences. Models for personality and emotion assessment are essential to help us gain valuable insights from the huge amount of data. Various tools that measure emotions and personality (Marengo et al., 2019) have been used and evaluated in previous studies (e.g., Cattell, 1973). We focus on technologically developed methods such as PC (Phan et al., 2021). PC has emerged as a promising method for assessing personality using digital footprints and computational approaches. By analyzing data from online platforms and social media, researchers have tried to extract individuals' personality traits.

TPC is an automated method to generate personality predictions using Natural Language Processing (NLP). It represents an innovative approach for understanding and analyzing personality traits on a broader scale (Mushtaq et al., 2023). The task of TPC can be framed as either regression or classification, depending on the nature of the personality measurements, meaning whether they are expected to be continuous or discrete. Within the field of TPC, supervised learning methods have predominantly been employed. Initially, simpler methods have been applied such as Support Vector Machines (SVM) (Pratama et al., 2015), and recently, more complex methods that use deep-learning and transformers architecture (Ren et al., 2021) have been employed.

For the prediction of emotions, various approaches have been implemented (e.g., Hasan et al., 2019, Maithri et al., 2022). Hasan et al., (2019) present a supervised learning system separated in two tasks: an offline training task and an online classification task. In the study conducted by Barriere et al., (2022), the authors present the results of the WASSA 2022 Shared Task, which is the primary dataset utilized in the current research. The paper provides an overview of the models employed by participating teams for emotion prediction. Most teams leveraged variants of the Bidirectional Encoder Representations from Transformers (BERT) model, such as RoBERTa and DeBERTa, to accomplish this task. In addition, Wen et al., (2021) implemented the RoBERTa model, among others, for emotion prediction in the PELD dataset, which is also utilized in the present study.

In the context of personality prediction, SVM (Pratama et al., 2015), and Regression (He et al., 2021) are usually implemented, in addition to BERT, its variants (Christian et al., 2021), and Recurrent Neural Networks (RNN) (Liu et al., 2017). These algorithms have been applied to take advantage of textual messages for personality trait detection. Moreover, Maharani et al., (2022), a paper that shares similar interests with our study, applied SVM among others to predict the Big Five personality traits. Our project differs from the aforementioned paper because we implemented variants of BERT for the predictions, and we investigate the alignment between correlations from annotated data and correlations reported in established personality and emotion theories. In the WASSA 2022 paper by Ghosh et al., (2022), SVM was applied using demographic information as input, instead of essays as we did, for personality trait prediction. Furthermore, Christian et al., (2021) employed BERT and its variants for the Big Five prediction.

2.3 Relationship between personality and emotions

Numerous academic papers (e.g., Farnadi et al., 2014) have examined the complex relationship between personality traits and emotions. While some studies (e.g., Maharani et al., 2022) have undertaken a comprehensive analysis of all potential correlations between these constructs, others (e.g., Druschel et al., 1999) have focused on specific pairs of personality traits and emotions presented in the next paragraphs.

Significant contributions to the examination of all possible correlations between personality traits and emotions have been made by studies such as Maharani et al., (2022) and Penley et al., (2002). In Maharani et al., (2022), the findings demonstrate significant associations between Openness and various emotions, including sadness, fear, anger, and disgust. Moreover, Conscientious, Extroverted, and Agreeable individuals have a higher chance of expressing joy, and surprise. Penley et al., (2002), on the other hand, explore the relationship between personality, emotions, and stress. The outcomes of this research unveil insightful information, such as that Agreeableness is not significantly correlated with any of the emotion features. It is worth acknowledging that our approach differs from those studies in terms of the implemented models; we applied variants of BERT for predictions while they used simpler models.

Numerous academic papers investigated the relationship between joy and all Big Five personality traits (e.g., Mitte et al., 2008; Berenbaum et al., 2016; Shiota et al., 2006). According to these studies, joy exhibits a positive association with all Big Five personality traits, except for Neuroticism. In Mitte et al., (2008) and Shiota et al., (2006) two distinct categories of joy—self-measured and peer-measured are included. In the domain of anger, fear, and sadness, and their connections with Big Five personality traits, a comprehensive meta-analysis conducted by Marengo et al., (2021) provides useful information. Their study showed that anger, fear, and sadness tend to reveal negative associations with all Big Five personality traits, except for Neuroticism.

The relationship between disgust and personality has been a point of interest in some papers (Druschel et al., 1999; Tybur et al., 2013). In (Druschel et al., 1999) a positive relationship between disgust and all personality traits, except for Openness, is reported. Investigations, like (Tybur et al., 2013), tend to separate disgust into distinct categories, such as moral, sexual, and pathogen disgust, and examine their individual relationships with personality traits. These studies offer beneficial perspectives on emotion-personality pairs.

It is important to acknowledge the widely accepted theories about the relationship between personality traits and emotions in addition to presenting the results of earlier studies. Prominent among the well accepted relationships is the strong one observed between joy and extraversion, which has been widely supported in literature (e.g., Shiota et al. 2006; Costa et al., 1980; Smillie et al, 2015). Furthermore, the negative correlation between anger and agreeableness is widely accepted (e.g., Marengo et al., 2021; Graziano, 1996). Another well-established theory is the positive relationship between neuroticism and negative emotions, including anger, fear, and sadness. Additionally, it is worth noting that there are studies that have showed that the association between anger and neuroticism is weaker compared to the one between fear/sadness and neuroticism (e.g., Costa et al., 1992; De Young, 2007; Judge 1999).

These established theories provide a foundation for understanding the relationship between personality and emotions, and they support this research by contributing to the theoretical framework. By using them to validate our results, this study aims to further examine this relationship.

3. Datasets

Section 3.1 provides a description of the datasets utilized in this study, along with an explanation of the original annotation process and the most outstanding data exploration results. Additionally, in Section 3.2 the preprocessing of the data during this research project is presented.

3.1 Description of the datasets

Multiple datasets have been created to help research in the field of TPC. Considering the time constraints of this project, the selection of appropriate datasets requires prioritizing those that have already annotated emotions and personality traits. Two datasets, WASSA 2022 and PELD, that meet these requirements are utilized in this project. Furthermore, the selection of these datasets offers the opportunity to examine the correlations between personality traits and emotions across diverse dataset types. The PELD dataset mainly consists of informal and concise dialogues, while the WASSA 2022 dataset contains more formal and extensive essays.

The **WASSA 2022** dataset is the primary dataset utilized in this research, initially used in the WASSA 2022 Shared Task Competition. Only the training data with gold standards for emotion and personality characteristics is employed because the test dataset was not available. The dataset given by Buechel et al., (2018), which includes 418 news articles in total, served as the basis for the creation of WASSA 2022. Each news article is related to essays generated by participants in response to distressing news concerning individuals, groups, or situations. Although there is a short description of the collection process in the next paragraph, a detailed overview on the collection process can be found in Buechel et al., (2018).

To acquire the corpus, Buechel et al., (2018) organized a crowdsourcing task on MTurk.com[3], redirecting participants to a Qualtrics.com[4] questionnaire. Multiple participants completed background measurements on demographics, the Big Five personality traits, empathy, distress, and Interpersonal Reactivity Index (IRI). After that, they read and commented on 5 randomly selected news articles. This dataset also contains emotion labels, the 6 basic Ekman emotion labels, namely anger, disgust, fear, joy, sadness, and surprise, with the addition of neutral category. The emotions were first predicted automatically (RNN, RoBERTa) and then manually verified. For the manual verification Another Amazon Mechanical Turk task was established, for which annotators were recruited (Barriere et al., 2022).

Table 1 shows a sample of the dataset WASSA 2022. The dataset has 1860 entries in total and includes information including demographics, the Big Five personality traits, emotions as well as measures of empathy, distress, and IRI. Although some of this information is not relevant to our research, it was necessary for certain tasks of the competition. Furthermore, the Big Five personality traits are represented by values ranging from one to seven with an incremental step of 0.5. This representation provides a thorough understanding of the personality traits, emotions, and essays in the WASSA 2022 dataset.

[3] Amazon Mechanical Turk (mturk.com)

[4] Qualtrics XM: The Leading Experience Management Software

Message_ID	Writer_ID	Essay	Emotion	Openness	Stability
R_2VxdAcFzlcpbk9U_4	R_2VxdAcFzlcpbk9U	She lost my sympathy really quick. Seems like she likes attention. This does make me look at Paris differently. These all seem like inside jobs. I would be safe there. But, I still don't plan on visiting. Not a fan of people endorsing candidates in countries that they can not vote. Seems like an ego thing.	anger	4	6
R_1CDCVd3BLzRFvYW_3	R_1CDCVd3BLzRFvYW	Im shocked at the possibility, and now at the direct knowledge that the opiate crisis in america not only affects older people (adults), but also very young children as well. Opiate overdoses are sad and brutal but the same thing for a child stirs shock and sadness within me to think of the reasons it occurs and that could certainly be prevented.	surprise	6.5	7

Table 1: Example data structure of WASSA 2022

The distributions of emotions and Big Five personality traits in the WASSA 2022 dataset are presented in Figure 1 and Figure 2. These plots reveal noticeable imbalances in both categories. The occurrence of sadness exceeds that of other emotions, while joy appears infrequently. Similarly, the distribution of personality traits exhibits imbalances, with the exception of extraversion which is balanced. Remarkably, Conscientiousness, Openness, Agreeableness, and Stability demonstrate an increasing trend as the corresponding trait values increase.

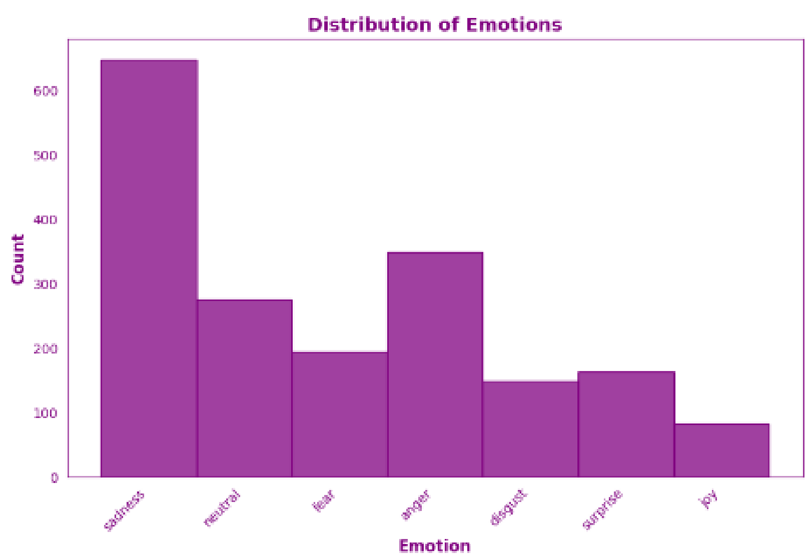

Figure 1:Emotions' distribution, WASSA 2022

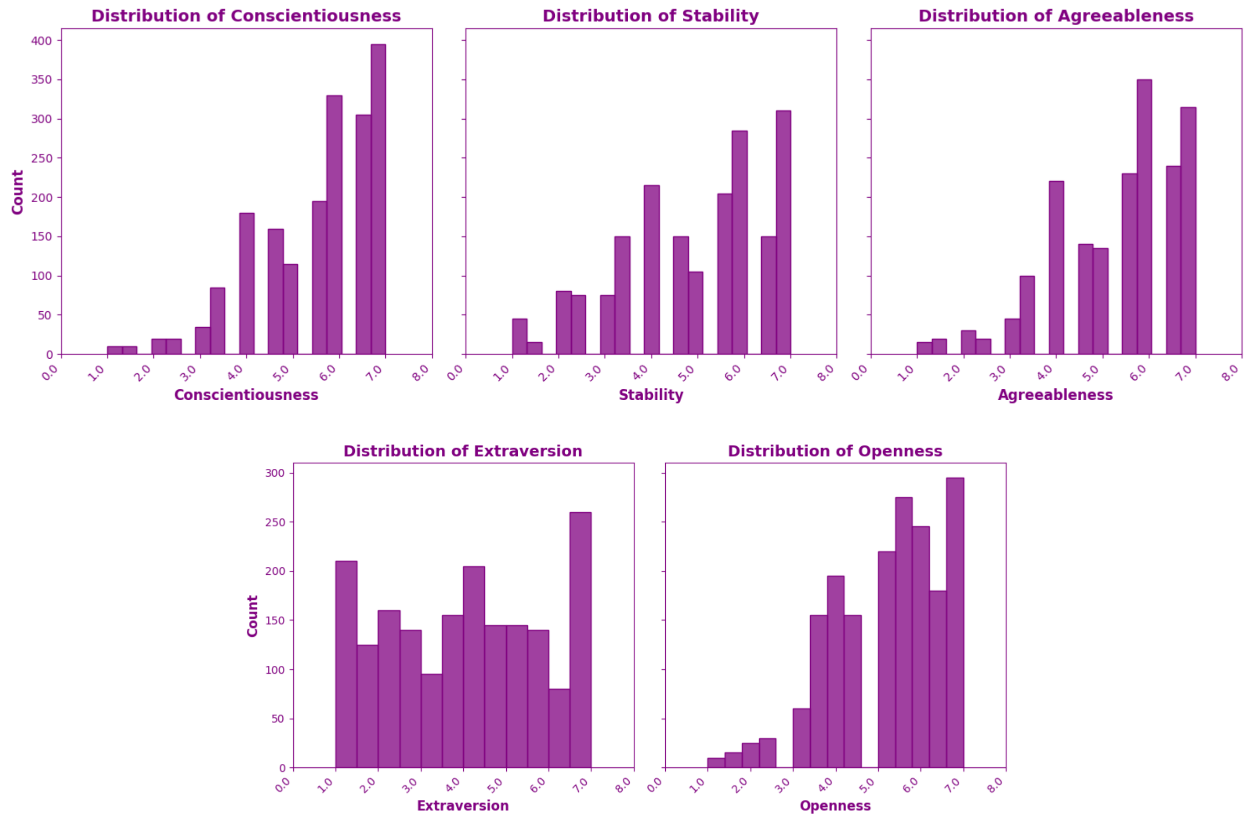

Figure 2: Big Five Traits' distributions, WASSA 2022

Concerning the characteristics of essays within the WASSA 2022 dataset, their length aligns with expectations. The token counts of comments range from 49 to 189, with an average of approximately 92 tokens per essay as shown in Table 2; this does not influence the implemented methods in a negative nor a positive way. It is worth mentioning that all comments are unique, except for one instance where an essay appears twice. However, the remaining characteristics of the comment differ, thus it is not considered as a duplicate entry. Lastly, the dataset does not have any missing values.

Maximum number of tokens in one essay	189
Maximum number of tokens in one essay	49
Average number of tokens per essay	91.64

Table 2: Summary of text characteristics for WASSA 2022

While the WASSA 2022 dataset served as the primary dataset for the current study, it is essential to acknowledge its limitations. Its most remarkable limitation is the presence of imbalanced data, as previously discussed. This imbalance may negatively impact the performance of prediction models because they might not learn to predict certain classes. Additionally, the dataset is characterized by a small number of entries, which might hinder us from achieving high model accuracy; the models are trained more effectively with larger datasets.

The **PELD** (Personality and Emotion in Language Dataset) dataset is a text-based emotional dialog collection that includes personality traits for speakers. The dialogues within PELD are sourced from the emotional dialogues found in MELD and EmoryNLP. The Multimodal EmotionLines Dataset (MELD)[5] dataset in addition to text data contains audio and visual modality distinguishing it from the EmotionLines dataset. The personality trait annotations in PELD are adopted from the personality annotations in FriendsPersona[6]. It comprises 711 short conversations and annotations from the first four seasons of Friends TV Show transcripts. Only the personality features of the six main characters are preserved as they occur more frequently.

The structure of the PELD is depicted in Table 3. There are 6509 rows in total; each one represents a complete dialogue consisting of three utterances, and thus, enabling the reader to instantly understand the context rather than just analyzing isolated sentences. The personality column in the dataset identifies the Big Five personality traits for each speaker, arranged in the following order: Openness, Conscientiousness, Extraversion, Agreeableness, and Stability. Additionally, the dataset includes information about the emotions caused by each utterance, including the following categories: anger, disgust, fear, joy, neutrality, sadness, and surprise. Apart from that, it contains information related to the sentiment of the utterances. Consequently, this structured representation allows an understanding of the personality traits, emotions, and dialogues within the PELD dataset.

Speaker_1	Personality	Utterance_1	Utterance_2	Utterance_3	Emotion_1
Monica	[0.713, 0.457, 0.457, 0.66, 0.511]	Do you love her?	We said it was only going to be two weeks, y know?	You love her!	neutral
Joey	[0.574, 0.614, 0.297, 0.545, 0.455]	Dude, I am sorry about what I said!	No, no, you re right, it is a ridiculous name!	It s not that bad.	sadness
Monica	[0.713, 0.457, 0.457, 0.66, 0.511]	You kissed another woman!	Call it even?!	Okay!	anger
Rachel	[0.635, 0.354, 0.521, 0.552, 0.469]	Really?	Yeah! Look!	I ve never lived like this before.	surprise

Table 3:Example data structure of PELD

The emotion and speaker distributions of the PELD dataset are shown in Figure 3. The analysis of emotion distribution reveals imbalanced patterns; the occurrence of neutral emotion is significantly higher compared to the other emotions, while the presence of disgust is low. This observation differs from our expectations, considering that the PELD dataset includes personal dialogues. Notably, neutral emotion appears in approximately half of the total entries, amounting to 2910 occurrences. On the other hand, the distribution of speakers is nearly evenly distributed, with each speaker making appearances relatively equally.

[5] GitHub - declare-lab/MELD: MELD: A Multimodal Multi-Party Dataset for Emotion Recognition in Conversation

[6] GitHub - emorynlp/personality-detection: Personality detection on multiparty dialogue

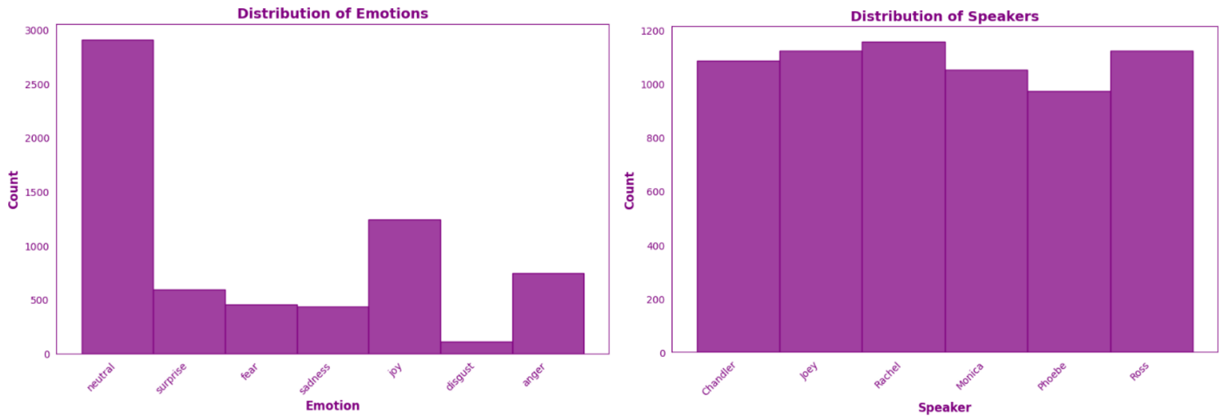

Figure 3: Emotion's and Speakers' Distributions, PELD

Regarding the characteristics of utterances in the PELD dataset as shown in Table 4, they tend to be short due to their nature as parts of dialogues. The length of phrases varies, with the maximum token count reaching 91 tokens, while the minimum consists of only one token. It is worth mentioning that the most frequently appearing utterance is the word "What?", which occurs 85 times. Moreover, the average token count per phrase is approximately 11 tokens. Finally, the dataset does not comprise missing values or duplicates.

Maximum number of tokens in one utterance	91
Maximum number of tokens in one utterance	1
Average number of tokens per utterance	11.24

Table 4: Summary of text characteristics for PELD

There are several limitations related to the PELD dataset which hindered its use for our study. Despite containing relevant information, the personality traits in the PELD dataset were attributed to the speakers rather than the individual utterances. This means that there are only six speakers and consequently, only six groups of personality traits. Therefore, this dataset failed to provide meaningful insights into the relationship between emotions and personality traits, rendering it mainly suitable for investigation of speakers' emotions. The research of this dataset focused on preliminary predictions, and this dataset did not serve as the primary focus of our study.

3.2 Preparation of the data

The initial stage of preprocessing for both datasets involved splitting them into separate sets for training, validation, and testing. The distributions of the sets for both datasets are presented in Table 5.

Dataset	Total	Train	Val	Test
WASSA 2022	1860	1339	149	372
PELD	6509	4686	521	1302

Table 5: Train/val/test sets

In both datasets, the textual data in the form of essays did not require additional preprocessing, such as lemmatization, punctuation or stop word removal. The models employed in this study performed better when only tokenization was applied to the text. However, preprocessing was needed for the emotion and personality traits columns to align with the specific format required by the models. Furthermore, certain columns of the datasets such as sentiment for PELD and IRI for WASSA 2022 were irrelevant to our research objectives and, therefore, were removed.

The preprocessing procedure of **WASSA 2022** involved the modification of the personality traits and emotion columns. First, we tried to approach the task of personality prediction as a classification task with 13 labels and then, as a regression one but finally, we rounded the values corresponding to the personality traits to the nearest integer resulting in a reduction from 13 labels to 7. This rounding process outperformed and enhanced the accuracy of the predictions. In addition, to satisfy the demands of the prediction models, the values of the personality traits were decreased by one, forcing the labels to begin from zero rather than one.

Regarding the emotion column, it underwent an encoding process, where the seven different emotions were replaced by integers ranging from zero to six, as follows: 0 for anger, 1 for disgust, 2 for fear, 3 for joy, 4 for neutral, 5 for sadness, and 6 for surprise. Moreover, binary columns were created, with each emotion being assigned its respective column. This preprocessing step served the computation of the original correlations between the emotions and the personality traits. These procedures were necessary to prepare the WASSA 2022 dataset appropriately, and to ensure that the models perform optimally.

In the **PELD** dataset, the Big Five personality traits were initially in a single column. For their separate investigation, the personality traits were divided into five distinct columns. However, due to the limited number of speakers in the dataset which hindered us from generating personality predictions, the individual personality trait columns were not useful during this study.

For the emotion column in PELD we followed the same preprocessing steps as for WASSA 2022. This implies the replacement of the seven different emotions with integer labels ranging from zero to six. Specifically, the encodings were as follows: 0 for anger, 1 for disgust, 2 for fear, 3 for joy, 4 for neutral, 5 for sadness, and 6 for surprise. These preprocessing actions benefited the subsequent research regarding the PELD dataset.

4. Methodology

This section discusses the diverse methodologies that were implemented in the current study to satisfy the tasks of prediction, evaluation, computation of correlation, and comparison.

4.1 Description of the methods

The research process consists of several steps to answer the three research questions. The first step focuses on personality and emotion predictions using state-of-the-art supervised and transformer-based algorithms including variants of BERT, SVM and Logistic Regression. These predictions were then evaluated using numerous metrics such as accuracy and F1-score. The next step involves the computation of correlations between personality traits and emotions for annotated and predicted data and their comparison. The last step is theoretical in nature, requiring theoretical research to identify established theories.

Predictions' generation in our case is a classification task instead of a regression one. Methods such as linear regression were applied to check whether it is a regression task, but the R squared was negative. A positive R squared is desired, so the problem is better suited for classification. We applied supervised machine learning methods to predict the labels for emotions and Big Five traits. Supervised methods are designed to predict or classify an outcome of interest; its goal is to forecast or classify a specific outcome (Jiang et al., 2020).

The methods of SVM and Logistic Regression were applied for personality prediction; both are simple yet powerful algorithms. The SVM algorithm initially proposed by Boser et al. (1992) has since become a widely adopted (e.g., Maharani et al., 2022) machine learning technique for classification and regression tasks. It identifies an optimal hyperplane that effectively separates different classes by maximizing the margin between them. Similarly, logistic regression is a widely utilized (e.g., Alavi et al., 2017; Lane et al., 2011) statistical modeling technique primarily employed for binary classification tasks (Peng et al., 2002) and determines the probability that an instance belongs to a particular class. Notably, these two methods were implemented once by using raw counts of words and once by using Term Frequency - Inverse Document Frequency (TF-iDF) vectors.

Three variants of BERT, DistilBERT, RoBERTa, and DeBERTa, were applied for emotion and personality prediction. BERT originally proposed by Devlin et al., (2019) is a pre-trained transformer-based model that can perform very well in various NLP tasks such as sentiment analysis. To learn contextual word representations, it makes use of a transformer architecture and a masked language model objective. By incorporating bidirectional context information, BERT significantly improves the understanding of word meaning and context (Devlin et al., 2019). DistilBERT, a compact variant of BERT introduced by Sanh et al. (2019), is especially designed to offer decreased size and improved computational efficiency, while still operating accurately. This is achieved through a technique known as knowledge distillation, enabling the compression of the original BERT model while preserving most of its predictive capabilities.

Furthermore, RoBERTa, proposed by Liu et al. (2019), uses the fundamental architecture of BERT while incorporating advancements in the training process, enhancing its performance. The main modifications in the BERT model are the following: a longer training period with larger batches and a larger dataset, the removal of the next sentence prediction target, training on longer sequences, and dynamic alterations in the masking pattern used on the training data (Angin et al., 2022). Moreover, DeBERTa, by He et al. (2020), introduces changes in the attention mechanism of the original BERT model. The key innovation is the addition of separated attention mechanisms, allowing the model to concentrate on different aspects of the input text. By doing so, DeBERTa captures more relationships, resulting in improved performance. It was the best performing model for most cases and thus, its architecture is presented in Figure 4. The variants of BERT have demonstrated effectiveness in various NLP tasks such as text-based classification (e.g., Büyüköz et al., 2020; Adel et al., 2022; Huang et al., 2021; Prasad et al., 2022) making them valuable tools in computational linguistics.

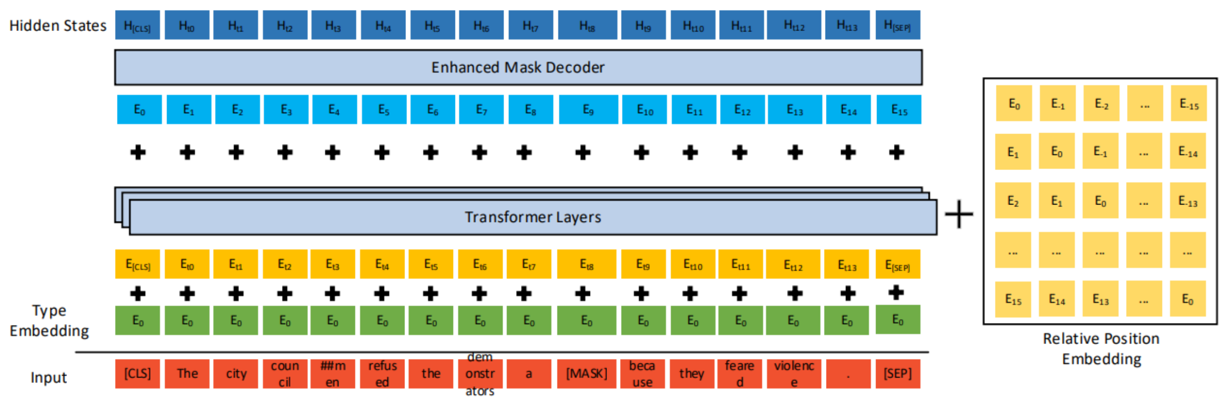

Figure 4: The model architecture of DeBERTa (Scale up DeBERTa (iclr.cc))

In addition, both Linear Regression and Neural Networks (NN) are two effective methodologies (Rastogi et al., 2020, Liu et al., 2017) for personality predictions, although in this project their results were not further used. Linear Regression, being a simpler approach, estimates the relationship between independent and dependent variables through linear combinations (Su et al., 2012). In contrast, NN represent a more complex approach, inspired by the structure and functioning of the human brain, consisting of interconnected layers of artificial neurons (Bishop, 1994). The type of NN that was implemented in this project is a feedforward network, named Multilayer Perceptron (MLP) (Pinkus, 1999).

Evaluation of the employed methodologies is crucial to assess the effectiveness of predicting personality traits and emotions from text data. We used state-of-the-art classification metrics, namely accuracy, weighted precision, recall, and F1-score, and macro precision, recall, and F1-score. The weighted average version of these metrics considers class imbalance, while their macro average version treats each class equally (Grandini et al., 2020). Both versions are important in our study to examine a broad evaluation of the predictions.

Correlations between personality traits and emotions were calculated to address the second research question namely *how related the correlations between annotated and predicted personality traits and emotions are*. For emotions, binary variables with one and zero were created for each emotion. For personality traits, the predicted values were increased by one because they were reduced by one before training the models. The Pearson correlation coefficient is employed as a statistical measure to quantify the strength and direction of these relationships. It ranges from -1 (perfect negative correlation) to 1 (perfect positive correlation), with a value of 0 indicating no linear correlation (Ly et al., 2018). To enhance the interpretability of the findings, heatmaps are utilized as visual representations of the correlation results. By using different colors to represent the magnitude of the correlations, heatmaps provide a visual summary of patterns and relationships of the data (Bojko, 2009).

A **comparative analysis** is conducted by examining the differences and similarities between the correlations of predicted and original data. The metrics that we used are the Pearson correlation coefficient, the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the p-value from the Wilcoxon signed-rank test. This comparative assessment serves to identify the reliability of TPC.

To answer the third research question namely *how consistent the correlations from annotated data with the correlations reported in established personality and emotion theories are*, an analysis is conducted to compare the correlations from annotated data with the ones in established theories. This research question relates to a theoretical investigation. We looked in literature and found numerous studies that investigate the relationship between all the Big Five traits and emotions (e.g., Penley et al., 2002), and others that investigate the relationship between one Big Five trait and emotions (e.g., Berenbaum et al., 2016).

4.2 Experimental settings

Python programming language serves as the primary tool for implementing the research methodology, and all the code developed for this study is available on GitHub[7]. The application of BERT variants for emotion and personality analysis involved a straightforward process with simple modifications to the BERT model. To ensure reproducibility, we set random seeds to the code for random number generation. Two essential steps were needed to implement the models: tokenization of the textual data and model training.

For tokenization, the input texts were tokenized using specific pretrained models: 'microsoft/deberta-base', 'distilbert-base-cased', and 'roberta-base' for DeBERTa, DistilBERT, and RoBERTa, respectively. A maximum sequence length of 128 was set in all cases. After that, the corresponding models received the tokenized inputs.

Fine-tuning of the models was performed for 10 epochs using the Adam optimizer (Kingma et al., 2017) with a learning rate of 1e-5 and a batch size of 16. These hyperparameters were chosen empirically, following common BERT fine-tuning procedures (Devlin et al., 2019). The training process was conducted on Google Colab with a T4 GPU as a hardware accelerator, taking approximately 6 minutes for each model.

On the other hand, for SVM the kernel 'LinearSVC' is used empirically, while for Logistic Regression the solver algorithm 'lbfgs' is utilized. Lastly, MLP comprised two hidden layers with 100 and 50 neurons, respectively. The Adam optimization algorithm was used again empirically, and the model was trained for a maximum of 1000 iterations.

[7] GitHub - pinelopiiiii/Text_Personality_Emotion_Computing

5. Results and Discussion

This section presents the findings of the study. Section 5.1 presents the results of the predictions from the different models. Then, Section 5.2 discusses the results of correlations between personality and emotions and Section 5.3 presents the alignment of the established theories with the correlations obtained from the original data. Finally, Section 5.4 comprises the discussion on the results while answering the research questions.

5.1 Results on predictions

The prediction of emotions in the WASSA 2022 and PELD datasets involved the utilization of DistilBERT, RoBERTa, and DeBERTa models, with random predictions serving as a baseline. Random predictions were generated by randomly assigning one emotion to each essay. DeBERTa consistently demonstrated the best performance for emotion prediction across evaluation metrics including accuracy in both datasets. Appendix A provides a detailed overview of the performances of all the models we implemented, while Table 6 presents an overview of the accuracy of some of the applied models. Remarkably, personality predictions were conducted only for the WASSA 2022 dataset, as discussed in Section 3.

Model's Accuracy	Emotion WASSA 2022	Emotion PELD	Conscientiousness	Openness	Extraversion	Stability	Agreeableness
DistilBERT Accuracy	0.56	0.59	0.38	0.33	**0.27**	**0.26**	0.37
RoBERTa Accuracy	0.6	0.60	0.40	0.27	0.20	0.25	0.37
DeBERTa Accuracy	**0.62**	**0.61**	**0.37**	**0.33**	0.25	0.20	**0.37**
SVM Accuracy	-	-	0.29	0.26	0.20	0.22	0.33
Random Accuracy	0.12	0.13	0.13	0.13	0.13	0.12	0.12

Table 6: Predictions' results

For emotion prediction in the WASSA 2022 dataset, DeBERTa achieved an accuracy of 0.62, a weighted F1-score of 0.62, and a macro F1-score of 0.64. There were seven classes, and the random model achieved an accuracy of 0.12, a weighted F1-score of 0.12 and a macro F1-score of 0.12 as well. Moreover, for emotion prediction in the PELD dataset, DeBERTa achieved an accuracy of 0.61, a weighted F1-score of 0.57, and a macro F1-score of 0.36, while the metrics for the random predictions were 0.13, 0.15 and 0.10 respectively. Thus, all the models highly outperformed random predictions in predicting emotions in both datasets. Another fact that verifies that the predictions were satisfying is that in WASSA 2022 Shared Task the emotion predictions had an accuracy around 0.65 (Barriere et al., 2022) which is slightly higher than our results. The F1-scores from all the models for each emotion in both datasets are presented in Appendix F.

The models employed for the personality prediction task included DistilBERT, RoBERTa, DeBERTa, SVM with raw counts, SVM with TF-IDF counts, Logistic Regression with raw counts, and Logistic Regression with TF-IDF counts. Additionally, Linear Regression and NN were applied, but the R squared values were all negative, while positive values were desired, and the results are presented in Appendix C. Each model was run five times, once for each of the Big Five personality traits. Random predictions were also generated to serve as a baseline reference. The complete results for personality predictions are shown in Appendix B, while the performance of the best models is presented in Table 6.

From Table 6, we notice that DeBERTa demonstrated the highest performance in predicting Conscientiousness, Openness, and Agreeableness among the applied models. Specifically, for Conscientiousness, DeBERTa achieved an accuracy of 0.37, a macro F1-score of 0.17, and a weighted F1-score of 0.31. Regarding Openness prediction, DeBERTa outperformed other models across all evaluation metrics, except for macro and weighted F1-score, with an accuracy of 0.33. For Agreeableness, DeBERTa achieved an accuracy of 0.37, a macro F1-score of 0.18, and a weighted F1-score of 0.31.

DistilBERT demonstrated the best performance in predicting Stability and Extraversion compared to other models. For Stability prediction, DistilBERT achieved an accuracy of 0.26, and there were no substantial differences in performance observed among the models. Concerning Extraversion prediction, DistilBERT outperformed other models across all evaluation metrics, except for weighted precision where DeBERTa showed better performance. Specifically, DistilBERT achieved an accuracy of 0.27, a macro F1-score of 0.21, and a weighted F1-score of 0.22. These findings are not very good, but studies focused on personality predictions using the same dataset report similar results (Ghosh et al., 2022); the reasons about this performance are discussed in Section 6.

5.2 Results on correlations

For the next step of our study, we calculated the correlations between the original Big Five traits and emotions, as well as between the predicted Big Five traits and emotions for each individual model and compared them. The results reveal a strong association between the correlations derived from the predicted data and those observed in the original data, as explained in more detail in the following paragraphs.

The analysis of the comparison between correlations obtained from the original data and the best predictions reveals an outstanding similarity. The heatmaps in Figure 5 allow for a clear visualization of the correlations, while the statistical measures reported in Table 7 provide evidence of the correlations' similarity. The high correlation coefficient of 0.725 indicates a strong positive correlation between the two sets of correlations, while the values of the errors suggest differences in the range of correlations. Additionally, the p-values obtained from the Wilcoxon signed-rank test support the conclusion that there is no statistically significant difference between the two sets of correlations. These findings provide a statistical reassurance of the close relationship between the correlations derived from the original data and the best predictions.

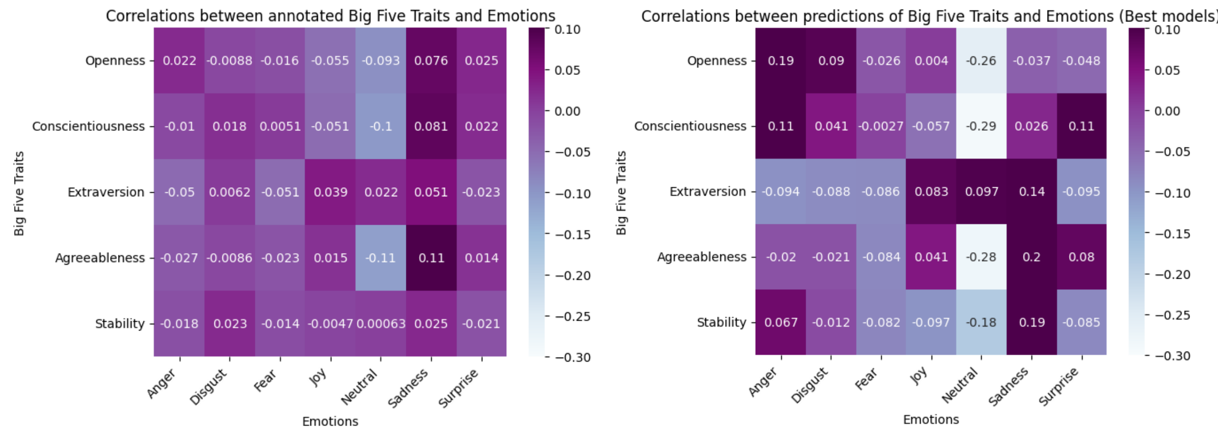

Figure 5: Heatmaps of correlations from original data and best predictions

Evaluation metrics	Values
Correlations coefficient	0.725
MAE	0.079
RMSE	0.095
Wilcoxon	No significant difference

Table 7: Comparison of correlations from annotated data and from best predictions

Appendix D includes all the heatmaps with correlations and Appendix E contains the comparison of the correlations between the annotated data and the predictions from each individual model. Overall, the comparisons consistently indicate that the correlations are not statistically different, confirming that the predicted correlations align well with those obtained from the annotated data. However, it should be noted that the degree of similarity varies among the different models; some models including correlations using DeBERTa predictions exhibit a closer relationship to the annotated correlations compared to others such as correlations using Logistic regression predictions.

5.3 Results on alignment with theories

Figure 6 displays the correlations between annotated emotions and personality traits for both the WASSA 2022 and PELD datasets. In this Section, the alignment of these correlations with the established theories outlined in Section 2 will be examined and the results are presented in Table 8. It is noteworthy to mention that although the theory explores Neuroticism as a personality trait, our datasets consider its opposite counterpart, which is Stability. Consequently, the emotions should have an inverse relationship with Stability compared to their relationship with Neuroticism.

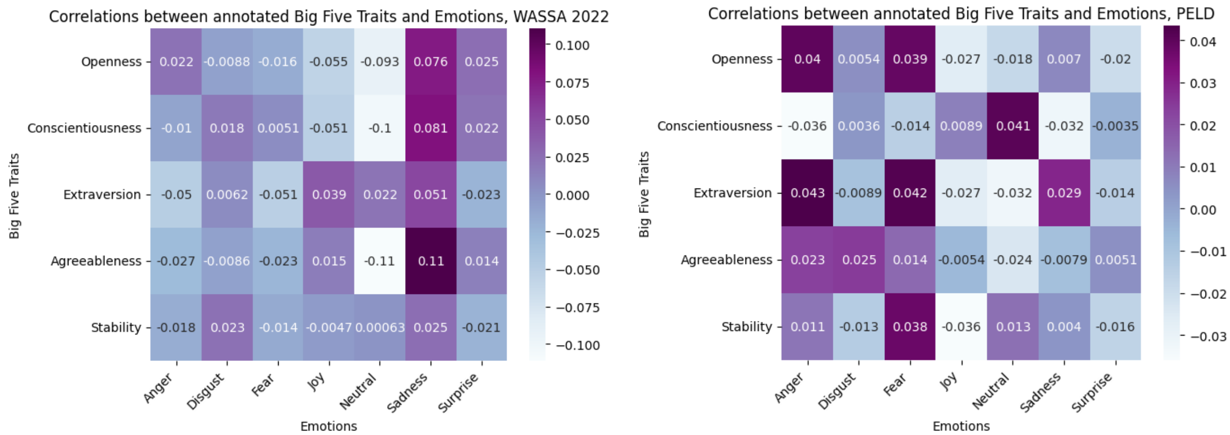

Figure 6: Correlations of annotated data for WASSA 2022 and PELD datasets

Based on established theories (Marengo et al., 2021), anger, fear, and sadness are expected to exhibit negative correlations with all the personality traits. In the case of WASSA 2022, this theory is partially supported for anger and fear, apart from the anger-Openness (0.022) and fear-Conscientiousness (0.0051) pairs. However, this theory does not hold for any of the correlations between personality traits and sadness in this dataset. As for the PELD dataset, the theory aligns with only half of the correlation pairs (e.g., anger-Conscientiousness, sadness-Agreeableness). Furthermore, established theories suggest that the association between anger and Stability should be weaker compared to the associations between fear and sadness with Stability. However, this pattern is not observed in either of the datasets, indicating a deviation from the expectations proposed by the established theory.

In addition, established theories (e.g., Mitte et al., 2008; Berenbaum et al., 2016; Shiota et al., 2006). propose that joy should have positive correlations with all the Big Five traits but Figure 5 reveals that this is not the case for most of the personality-emotion pairs in both datasets. For WASSA 2022, joy exhibits a positive correlation only with Extraversion (0.039) and Agreeableness (0.015). On the other hand, for the PELD dataset, joy demonstrates a positive correlation only with Conscientiousness (0.0089). Moreover, it is widely accepted (e.g., Shiota et al. 2006; Costa et al., 1980; Smillie et al, 2015) that joy and Extraversion have a strong positive relationship. While this is true for the WASSA 2022 dataset, it does not align with the results obtained for the PELD dataset.

The neutral emotion, in both the WASSA 2022 and PELD datasets, exhibits a lack of alignment with established theories. According to theories, neutral emotion should have no correlation with any of the Big Five personality traits. However, in both datasets, the neutral emotion demonstrates extreme values, indicating a difference from the expected pattern. In terms of disgust and Agreeableness, established theories (Druschel et al., 1999) suggest a highly positive correlation. This expectation is met in the PELD dataset, but not in WASSA 2022.

WASSA 2022		PELD	
Pairs that agree with theories	**Pairs that disagree with theories**	**Pairs that agree with theories**	**Pairs that disagree with theories**
Anger-All personality traits except Openness	Neutrality, Surprise, Sadness-All personality pairs	Joy, Anger, Fear - Conscientiousness	Neutrality-All personality pairs
Fear-All personality traits except Conscientiousness	Anger-Openness	Sadness-Conscientiousness, Agreeableness	Anger-Extraversion, Agreeableness, Openness, Stability
Joy-Agreeableness, Extraversion	Disgust-Agreeableness	Disgust-Agreeableness	Joy-Extraversion, Agreeableness, Openness, Stability

Table 8: Alignment with theories

5.4 Discussion

- **[RQ 1]:** *How accurately can automatic models predict personality traits and emotions from text data?*

Section 5.1 provides the results that answer the first research question. In summary, for both tasks, the pre-trained models achieved a good performance compared with the results from the WASSA 2022 Shared Task (Barriere et al., 2022). Emotion predictions have an accuracy of around 0.6 while personality predictions have an approximate accuracy of 0.3 improving the random baseline by around 350 and 150% respectively. Moreover, emotions were predicted a lot more accurately than personality traits. As shown in Table 7 Conscientiousness and Agreeableness were the best to predict, while Stability and Extraversion had the worst results. Additionally, because of imbalanced data, when implementing most models, the lowest values of each personality trait do not occur on the predictions since they do not appear often on the training set. The models that achieve the best performance for predicting emotions, Conscientiousness, Openness, and Agreeableness is DeBERTa, while for Extraversion and Stability predictions DistilBERT performed the best.

- **[RQ2]:** *How similar are the correlations between the annotated and the predicted personality traits and emotions?*

The results concerning the second research question are presented in Section 5.2, highlighting the outcomes of the correlation analysis. Overall, these results show the ability of various models to capture the general patterns of correlations, with certain models such as DeBERTa providing more accurate representations of the annotated data correlations than others such as Logistic regression. It is expected that models producing more accurate predictions would exhibit correlations that relate closer to those of the original data and the results support this expectation. Specifically, the comparison between the correlations of the original data and the ones from the best predictions reveals a strong positive correlation coefficient of 0.725. Additionally, the fact that there is no statistical difference further supports the conclusion that the predicted correlations follow a similar pattern to the one observed in the original data but the relatively high values of MAE and RMSE indicate that there is a different range in the correlations from annotated and predicted data.

- **[RQ 3]:** *How well do the correlations from original data align with the correlations reported in established personality and emotion theories?*

Section 5.3 presents the findings related to the third research question, focusing on the alignment of correlations with established theories. Regarding the WASSA 2022 dataset, the relationships between joy and Agreeableness, as well as joy and Extraversion, align with established theories. Additionally, the correlations involving anger conform to theories, except for its relationship with Openness. Similarly, the correlations involving fear agree with widely accepted research findings, with the exception of its relationship with Conscientiousness. However, certain emotions in the WASSA 2022 dataset, such as sadness and neutrality, do not align with established theories in terms of their relationships with personality traits. Moreover, the pairs involving disgust and Agreeableness, joy and Stability, joy and Conscientiousness, joy and Openness do not align with established theories.

For the PELD dataset, the correlations between disgust and Agreeableness align with established theories, as do the relationships of Conscientiousness with joy, anger, fear, and sadness, but the relationship between Stability and the same set of emotions disagree with widely accepted theories. In addition, the relationship between sadness and Agreeableness aligns with theories. However, the correlations involving neutral emotion and personality traits do not align with well-accepted research findings. Furthermore, the pairs involving joy and Extraversion, joy and Agreeableness, joy and Openness, anger and Extraversion, anger and Agreeableness, and anger and Openness do not agree with established theories. The reasons for the partial disagreement with theories are discussed in Section 6.

6. Conclusion and Future work

6.1 Ethical and legal considerations

The investigation of personality and emotions in this research project requires the inclusion of ethical considerations. As far as we are aware of, there are no ethical concerns associated with the data used in this study. It is crucial to emphasize that the data utilized is publicly accessible and does not include any personal user information.

Nevertheless, it is essential to recognize the potential for misuse of automated systems designed for personality computing. One case of misuse is the discrimination based on classes like gender or race (Buolamwini et al., 2018). Apart from that, the methodology proposed holds the potential for application in more delicate areas, such as mental health assessment. It is strongly advised that these systems be made available in such a way that personalized information cannot be derived so that potential risks are mitigated.

6.2 Conclusion

In this study we addressed the questions of *how accurately can automated methods predict personality traits and emotions from text data* and *how related the correlations between annotated and predicted personality traits and emotions are*. Finally, we investigated *how consistent the correlations from annotated data with the correlations reported in established personality and emotion theories are*.

Many models including RoBERTA, and SVM are applied but in the end the BERT variants were the ones with the highest performance. Specifically, DistilBERT and DeBERTa performed the best in predicting emotions and Big Five personality traits and the results show that emotions are predicted with an accuracy of around 0.6, while personality traits with an accuracy of around 0.3. These are a lot higher than the accuracy of the baseline random predictions and are similar with the results from the WASSA 2022 Task (Barriere et al., 2022). The correlations of predicted and original data are similar as indicated by the Pearson correlation coefficient of 0.725 and no significant difference is observed based on the Wilcoxon signed-rank test. Finally, it is revealed that certain emotion-personality pairs, such as fear and anger in relation to most Big Five traits, align to a great extent with established theories. However, pairs involving sadness/neutrality and all personality traits disagree with widely accepted theories.

Two limitations should be considered in this research. The primary limitation stems from the imbalanced and small datasets. Due to the imbalanced data, the models utilized for predictions may not have been trained to reach their fullest abilities. The relatively small dataset size (WASSA 2022: 1860, PELD: 6509) may have affected the generalization of the relationship between personality traits and emotions, leading to potential limitations in the alignment with established theories. Furthermore, the constrained time frame for conducting this research hindered the additional analyses that could have provided further insights. These challenges serve as an opportunity for future research.

To tackle the limitation of small, imbalanced datasets, the recommendations for future research include the enrichment of the datasets. By generating additional data, the models can be trained on a larger and more diverse set of examples enabling them to learn more patterns and potentially achieve higher prediction accuracy. Additionally, efforts should be made to enhance the representation of minority classes within the dataset to resolve the problem of data imbalance. In that way the original data might be more representative of reality and thus, more consistent with established theories. Lastly, this study could be extended by performing multi-task classification for personality trait predictions; this approach would involve simultaneously predicting multiple personality traits increasing the degree of agreement between the relationship of original data and predictions.

References

Abdullahi, Aisha Muhammad, Orji Rita, Rabiu Abbas Muhammad, and Kawu Abdullahi Abubakar. "Personality and Subjective Well-Being: Towards Personalized Persuasive Interventions for Health and Well-Being." *Online Journal of Public Health Informatics* 12, no. 1 (May 16, 2020): e1. https://doi.org/10.5210/ojphi.v12i1.10335.

Adel, Hadeer, Dahou Abdelghani, Mabrouk Alhassan, Elaziz Mohamed Abd, Kayed Mohammed, El-Henawy Ibrahim Mahmoud, Alshathri Samah, and Ali Abdelmgeid Amin. "Improving Crisis Events Detection Using DistilBERT with Hunger Games Search Algorithm." *Mathematics* 10, no. 3 (January 2022): 447. https://doi.org/10.3390/math10030447.

Adolphs, Ralph. "How Should Neuroscience Study Emotions? By Distinguishing Emotion States, Concepts, and Experiences." *Social Cognitive and Affective Neuroscience* 12, no. 1 (January 1, 2017): 24–31. https://doi.org/10.1093/scan/nsw153.

Alavi, Seyyed Salman, Mohammadi Mohammad Reza, Souri Hamid, Kalhori Soroush Mohammadi, Jannatifard Fereshteh, and Sepahbodi Ghazal. "Personality, Driving Behavior and Mental Disorders Factors as Predictors of Road Traffic Accidents Based on Logistic Regression." *Iranian Journal of Medical Sciences* 42, no. 1 (January 2017): 24–31.

Angin, Merih, Taşdemir Beyza, Yılmaz Cenk Arda, Demiralp Gökcan, Atay Mert, Angin Pelin, and Dikmener Gökhan. "A RoBERTa Approach for Automated Processing of Sustainability Reports." *Sustainability* 14, no. 23 (January 2022): 16139. https://doi.org/10.3390/su142316139.

Barriere, Valentin, Tafreshi Shabnam, Sedoc João, and Alqahtani Sawsan. "WASSA 2022 Shared Task: Predicting Empathy, Emotion and Personality in Reaction to News Stories." In *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, 214–27. Dublin, Ireland: Association for Computational Linguistics, 2022. https://doi.org/10.18653/v1/2022.wassa-1.20.

Berenbaum, Howard, Chow Philip I., Schoenleber Michelle, and Flores Luis E. "Personality and Pleasurable Emotions." *Personality and Individual Differences* 101 (October 2016): 400–406. https://doi.org/10.1016/j.paid.2016.06.023.

Bishop, Chris M. "Neural Networks and Their Applications." *Review of Scientific Instruments* 65, no. 6 (June 1, 1994): 1803–32. https://doi.org/10.1063/1.1144830.

Bojko, Agnieszka (Aga). "Informative or Misleading? Heatmaps Deconstructed." In *Human-Computer Interaction. New Trends*, edited by Julie A. Jacko, 30–39. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2009. https://doi.org/10.1007/978-3-642-02574-7_4.

Boser, Bernhard E., Guyon Isabelle M., and Vapnik Vladimir N. "A Training Algorithm for Optimal Margin Classifiers." In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 144–52. Pittsburgh Pennsylvania USA: ACM, 1992. https://doi.org/10.1145/130385.130401.

Buechel, Sven, Buffone Anneke, Slaff Barry, Ungar Lyle, and Sedoc João. "Modeling Empathy and Distress in Reaction to News Stories." In *Proceedings of the 2018 Conference on Empirical Methods in Natural*

Language Processing, 4758–65. Brussels, Belgium: Association for Computational Linguistics, 2018. https://doi.org/10.18653/v1/D18-1507.

Buolamwini, Joy, and Gebru Timnit. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91. PMLR, 2018. https://proceedings.mlr.press/v81/buolamwini18a.html.

Büyüköz, Berfu, Hürriyetoğlu Ali, and Özgür Arzucan. "Analyzing ELMo and DistilBERT on Socio-Political News Classification." In *Proceedings of the Workshop on Automated Extraction of Socio-Political Events from News 2020*, 9–18. Marseille, France: European Language Resources Association (ELRA), 2020. https://aclanthology.org/2020.aespen-1.4.

Cattell, Raymond B. *Personality and Mood by Questionnaire*. Personality and Mood by Questionnaire. Oxford, England: Jossey-Bass, 1973.

Christian, Hans, Suhartono Derwin, Chowanda Andry, and Zamli Kamal Z. "Text Based Personality Prediction from Multiple Social Media Data Sources Using Pre-Trained Language Model and Model Averaging." *Journal of Big Data* 8, no. 1 (December 2021): 68. https://doi.org/10.1186/s40537-021-00459-1.

Costa, Paul, and McCrae Robert. "Costa Jr PT, McCrae RR. Influence of Extraversion Neuroticism on Subjective Well-Being: Happy Unhappy People. J Pers Soc Psychol 38: 668-678." *Journal of Personality and Social Psychology* 38 (May 1, 1980): 668–78. https://doi.org/10.1037/0022-3514.38.4.668.

Costa, Paul T., and McCrae Robert R. "Four Ways Five Factors Are Basic." *Personality and Individual Differences* 13, no. 6 (June 1992): 653–65. https://doi.org/10.1016/0191-8869(92)90236-I.

Costa, Paul, and Mccrae R. "Neo PI-R Professional Manual." *Psychological Assessment Resources* 396 (January 1, 1992).

Dalgleish, Tim, and Power Mick. *Handbook of Cognition and Emotion*. John Wiley & Sons, 2000.

De Raad, Boele, and Marja Kokkonen. "Traits and Emotions: A Review of Their Structure and Management." *European Journal of Personality* 14, no. 5 (2000): 477–96. https://doi.org/10.1002/1099-0984(200009/10)14:5<477::AID-PER396>3.0.CO;2-I.

Denham, Susanne A, Bassett Hideko H, Brown Chavaughn, Way Erin, and Steed Jessica. "'I Know How You Feel': Preschoolers' Emotion Knowledge Contributes to Early School Success." *Journal of Early Childhood Research* 13, no. 3 (October 2015): 252–62. https://doi.org/10.1177/1476718X13497354.

Devlin, Jacob, Chang Ming-Wei, Lee Kenton, and Toutanova Kristina. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–86. Minneapolis, Minnesota: Association for Computational Linguistics, 2019. https://doi.org/10.18653/v1/N19-1423.

Druschel, B.a, and Sherman M.F. "Disgust Sensitivity as a Function of the Big Five and Gender." *Personality and Individual Differences* 26, no. 4 (March 1999): 739–48. https://doi.org/10.1016/S0191-8869(98)00196-2.

Fang, Qixiang, Giachanou Anastasia, Bagheri Ayoub, Boeschoten Laura, van Kesteren Erik-Jan, Kamalabad Mahdi Shafiee, and Oberski Daniel L. "On Text-Based Personality Computing: Challenges and Future Directions." arXiv, December 14, 2022. http://arxiv.org/abs/2212.06711.

Farnadi, Golnoosh, Sitaraman Geetha, Rohani Mehrdad, Kosinski Michal, Stillwell David, Moens Marie-Francine, Davalos Sergio, and De Cock Martine. "How Are You Doing? : Emotions and Personality in Facebook." In *EMPIRE2014 : 2nd Workshop on Emotions and Personality in Personalized Services*, 45–56, 2014. http://hdl.handle.net/1854/LU-7100029.

Ghosh, Soumitra, Maurya Dhirendra, Ekbal Asif, and Bhattacharyya Pushpak. "Team IITP-AINLPML at WASSA 2022: Empathy Detection, Emotion Classification and Personality Detection." In *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, 255–60. Dublin, Ireland: Association for Computational Linguistics, 2022.https://doi.org/10.18653/v1/2022.wassa-1.26.

Goldberg, Lewis R. "From Ace to Zombie: Some Explorations in the Language of Personality." *Advances in Personality Assessment* 1 (1982): 203–34.

Grandini, Margherita, Bagli Enrico, and Visani Giorgio. "Metrics for Multi-Class Classification: An Overview." arXiv, August 13, 2020. http://arxiv.org/abs/2008.05756.

Graziano, William, Jensen-Campbell Lauri, and Hair Elizabeth. "Perceiving Interpersonal Conflict and Reacting to It: The Case for Agreeableness." *Journal of Personality and Social Psychology* 70 (May 1, 1996): 820–35. https://doi.org/10.1037/0022-3514.70.4.820.

Hasan, Maryam, Rundensteiner Elke, and Agu Emmanuel. "Automatic Emotion Detection in Text Streams by Analyzing Twitter Data." *International Journal of Data Science and Analytics* 7, no. 1 (February 1, 2019): 35–51. https://doi.org/10.1007/s41060-018-0096-z.

He, Pengcheng, Liu Xiaodong, Gao Jianfeng, and Chen Weizhu. "DeBERTa: Decoding-Enhanced BERT with Disentangled Attention." *ArXiv*, 2020. https://www.semanticscholar.org/reader/05f5f8b2065a520846d89771ebaea2bb1534e9c6.

He, Xiaoli, and de Melo Gerard. "Personality Predictive Lexical Cues and Their Correlations." In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, 514–23. Held Online: INCOMA Ltd., 2021. https://aclanthology.org/2021.ranlp-1.58.

Huang, Bo, Bai Yang, and Zhou Xiaobing. "Hub at SemEval-2021 Task 2: Word Meaning Similarity Prediction Model Based on RoBERTa and Word Frequency." In *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)*, 719–23. Online: Association for Computational Linguistics, 2021. https://doi.org/10.18653/v1/2021.semeval-1.94.

Jiang, Tammy, Gradus Jaimie L., and Rosellini Anthony J. "Supervised Machine Learning: A Brief Primer." *Behavior Therapy* 51, no. 5 (September 1, 2020): 675–87. https://doi.org/10.1016/j.beth.2020.05.002

John, O., and Srivastava. S. "The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives.," 1999. https://www.semanticscholar.org/paper/The-Big-Five-Trait-taxonomy%3A-History%2C-measurement%2C-John-Srivastava/a354854c71d60a4490c42ae47464fbb9807d02bf.

Judge, Timothy A., Higgins Chad A., Thoresen Carl J., and Barrick Murray R. “The Big Five Personality Traits, General Mental Ability, and Career Success Across the Life Span.” *Personnel Psychology* 52, no. 3 (1999): 621–52. https://doi.org/10.1111/j.1744-6570.1999.tb00174.x.

Kern, Margaret L., McCarthy Paul X., Chakrabarty Deepanjan, and Rizoiu Marian-Andrei. “Social Media-Predicted Personality Traits and Values Can Help Match People to Their Ideal Jobs.” *Proceedings of the National Academy of Sciences of the United States of America* 116, no. 52 (December 26, 2019): 26459–64. https://doi.org/10.1073/pnas.1917942116.

Khan, Amir A., Jacobson Kristen C., Gardner Charles O., Prescott Carol A., and Kendler Kenneth S.. “Personality and Comorbidity of Common Psychiatric Disorders.” *The British Journal of Psychiatry: The Journal of Mental Science* 186 (March 2005): 190–96. https://doi.org/10.1192/bjp.186.3.190.

Kingma, Diederik P., and Ba Jimmy. “Adam: A Method for Stochastic Optimization.” arXiv, January 29, 2017. https://doi.org/10.48550/arXiv.1412.6980.

Lane, Wilburn, and Manner Chris. “The Impact of Personality Traits on Smartphone Ownership and Use.” *International Journal of Business and Social Science* 2 (January 1, 2011).

Liu, Fei, Perez Julien, and Nowson Scott. “A Language-Independent and Compositional Model for Personality Trait Recognition from Short Texts.” In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 754–64. Valencia, Spain: Association for Computational Linguistics, 2017. https://aclanthology.org/E17-1071.

Liu, Yinhan, Ott Myle, Goyal Naman, Du Jingfei, Joshi Mandar, Chen Danqi, Levy Omer, Lewis Mike, Zettlemoyer Luke, and Stoyanov Veselin. “RoBERTa: A Robustly Optimized BERT Pretraining Approach.” arXiv, July 26, 2019. http://arxiv.org/abs/1907.11692.

Ly, Alexander, Marsman Maarten, and Wagenmakers Eric-Jan. “Analytic Posteriors for Pearson’s Correlation Coefficient.” *Statistica Neerlandica* 72, no. 1 (2018): 4–13. https://doi.org/10.1111/stan.12111.

Maharani, Warih, and Effendy Veronikha. “Big Five Personality Prediction Based in Indonesian Tweets Using Machine Learning Methods.” *International Journal of Electrical and Computer Engineering (IJECE)* 12, no. 2 (April 1, 2022): 1973. https://doi.org/10.11591/ijece.v12i2.pp1973-1981.

Maithri, M., Raghavendra U., Gudigar Anjan, Samanth Jyothi, Barua Prabal Datta, Murugappan Murugappan, Chakole Yashas, and Rajendra Acharya U. “Automated Emotion Recognition: Current Trends and Future Perspectives.” *Computer Methods and Programs in Biomedicine* 215 (March 2022): 106646. https://doi.org/10.1016/j.cmpb.2022.106646.

Marengo, Davide, and Settanni Michele. “Mining Facebook Data for Personality Prediction: An Overview.” In *Digital Phenotyping and Mobile Sensing*, edited by Harald Baumeister and Christian Montag, 109–24. Studies in Neuroscience, Psychology and Behavioral Economics. Cham: Springer International Publishing, 2019. https://doi.org/10.1007/978-3-030-31620-4_7.

Marengo, Davide, Davis Kenneth L., Gradwohl Gökçe Özkarar, and Montag Christian. “A Meta-Analysis on Individual Differences in Primary Emotional Systems and Big Five Personality Traits.” *Scientific Reports* 11, no. 1 (April 2, 2021): 7453. https://doi.org/10.1038/s41598-021-84366-8.

Mitte, Kristin, and Kämpfe Nicole. "Personality and the Four Faces of Positive Affect: A Multitrait-Multimethod Analysis Using Self- and Peer-Report." *Journal of Research in Personality* 42, no. 5 (October 2008): 1370–75. https://doi.org/10.1016/j.jrp.2008.04.004.

Mushtaq, Sumiya, and Kumar Neerendra. "Text-Based Automatic Personality Recognition: Recent Developments." In *Proceedings of Third International Conference on Computing, Communications, and Cyber-Security*, edited by Pradeep Kumar Singh, Sławomir T. Wierzchoń, Sudeep Tanwar, Joel J. P. C. Rodrigues, and Maria Ganzha, 537–49. Lecture Notes in Networks and Systems. Singapore: Springer Nature, 2023. https://doi.org/10.1007/978-981-19-1142-2_43.

Peng, Chao-Ying Joanne, Lee Kuk Lida, and Ingersoll Gary M. "An Introduction to Logistic Regression Analysis and Reporting." *The Journal of Educational Research* 96, no. 1 (September 1, 2002): 3–14. https://doi.org/10.1080/00220670209598786.

Penley, Julie A, and Tomaka Joe. "Associations among the Big Five, Emotional Responses, and Coping with Acute Stress." *Personality and Individual Differences* 32, no. 7 (May 2002): 1215–28. https://doi.org/10.1016/S0191-8869(01)00087-3.

Phan, Le Vy, and Rauthmann John F. "Personality Computing: New Frontiers in Personality Assessment." *Social and Personality Psychology Compass* 15, no. 7 (2021): e12624. https://doi.org/10.1111/spc3.12624.\

Pinkus, Allan. "Approximation Theory of the MLP Model in Neural Networks." *Acta Numerica* 8 (January 1999): 143–95. https://doi.org/10.1017/S0962492900002919.

Plomin, Arnold H. Buss, Robert. *Temperament (PLE: Emotion): Early Developing Personality Traits*. London: Psychology Press, 2014. https://doi.org/10.4324/9781315745701.

Prasad, Janvi, Prasad Gaurang, and Gunavathi C. "GJG@TamilNLP-ACL2022: Emotion Analysis and Classification in Tamil Using Transformers." In *Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages*, 86–92. Dublin, Ireland: Association for Computational Linguistics, 2022. https://doi.org/10.18653/v1/2022.dravidianlangtech-1.14.

Pratama, Bayu Yudha, and Sarno Riyanarto. "Personality Classification Based on Twitter Text Using Naive Bayes, KNN and SVM." In *2015 International Conference on Data and Software Engineering (ICoDSE)*, 170–74, 2015. https://doi.org/10.1109/ICODSE.2015.7436992.

Rastogi, Rohit, Chaturvedi D. K., Satya Santosh, Arora Navneet, Trivedi Piyush, Singh Akshay Kr., Sharma Amit Kr., and Singh Ambuj. "Intelligent Personality Analysis on Indicators in IoT-MMBD-Enabled Environment." In *Multimedia Big Data Computing for IoT Applications: Concepts, Paradigms and Solutions*, edited by Sudeep Tanwar, Sudhanshu Tyagi, and Neeraj Kumar, 185–215. Intelligent Systems Reference Library. Singapore: Springer, 2020. https://doi.org/10.1007/978-981-13-8759-3_7.

Reisenzein, Rainer, Hildebrandt Andrea, and Weber Hannelore. "Personality and Emotion." In *The Cambridge Handbook of Personality Psychology*, edited by Gerald Matthews and Philip J. Corr, 2nd ed., 81–100. Cambridge Handbooks in Psychology. Cambridge: Cambridge University Press, 2020. https://doi.org/10.1017/9781108264822.009.

Ren, Zhancheng, Shen Qiang, Diao Xiaolei, and Xu Hao. "A Sentiment-Aware Deep Learning Approach for Personality Detection from Text." *Information Processing & Management* 58, no. 3 (May 2021): 102532. https://doi.org/10.1016/j.ipm.2021.102532.

Revelle, William R, and Scherer Klaus. "Personality and Emotion." In *Oxford Companion to Emotion and the Affective Sciences*, edited by David Sander and Klaus Scherer, 304–5. Oxford University Press, 2009.

Robins Wahlin, Tarja-Brita, and Byrne Gerard J. "Personality Changes in Alzheimer's Disease: A Systematic Review." *International Journal of Geriatric Psychiatry* 26, no. 10 (2011): 1019–29. https://doi.org/10.1002/gps.2655.

Sanh, Victor, Debut Lysandre, Chaumond Julien, and Wolf Thomas. "DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter." *ArXiv*, October 2, 2019. https://www.semanticscholar.org/paper/DistilBERT%2C-a-distilled-version-of-BERT%3A-smaller%2C-Sanh-Debut/a54b56af24bb4873ed0163b77df63b92bd018ddc.)

Shiota, Michelle N., Keltner Dacher, and John Oliver P. "Positive Emotion Dispositions Differentially Associated with Big Five Personality and Attachment Style." *The Journal of Positive Psychology* 1, no. 2 (April 2006): 61–71. https://doi.org/10.1080/17439760500510833.

Smillie, Luke D., DeYoung Colin G., and Hall Phillip J. "Clarifying the Relation Between Extraversion and Positive Affect." *Journal of Personality* 83, no. 5 (2015): 564–74. https://doi.org/10.1111/jopy.12138.

Su, Xiaogang, Yan Xin, and Tsai Chih-Ling. "Linear Regression." *WIREs Computational Statistics* 4, no. 3 (2012): 275–94. https://doi.org/10.1002/wics.1198.

Tupes, E. C., and Christal R. E. "Recurrent Personality Factors Based on Trait Ratings." *Journal of Personality* 60, no. 2 (June 1992): 225–51. https://doi.org/10.1111/j.1467-6494.1992.tb00973.x.

Tybur, Joshua M., and De Vries Reinout E. "Disgust Sensitivity and the HEXACO Model of Personality." *Personality and Individual Differences* 55, no. 6 (October 2013): 660–65. https://doi.org/10.1016/j.paid.2013.05.008.

Vinciarelli, Alessandro, and Mohammadi Gelareh. "A Survey of Personality Computing." *IEEE Transactions on Affective Computing* 5, no. 3 (July 2014): 273–91. https://doi.org/10.1109/TAFFC.2014.2330816.

Wen, Zhiyuan, Cao Jiannong, Yang Ruosong, Liu Shuaiqi, and Shen Jiaxing. "Automatically Select Emotion for Response via Personality-Affected Emotion Transition." In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 5010–20. Online: Association for Computational Linguistics, 2021. https://doi.org/10.18653/v1/2021.findings-acl.444.

Appendices

Appendix A: Emotion prediction

Models for emotion prediction, WASSA 2022

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted Precision	Weighted Recall	Weighted F1-score
DistilBERT	0.56	0.51	0.43	0.43	0.55	0.56	0.53
RoBERTa	0.60	0.53	0.49	0.48	0.60	0.60	0.58
DeBERTa	**0.62**	**0.59**	**0.54**	**0.55**	**0.64**	**0.62**	**0.62**
Random	0.12	0.12	0.12	0.11	0.19	0.12	0.14

Table 9: Emotion prediction, WASSA 2022

Models for emotion prediction, PELD

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted Precision	Weighted Recall	Weighted F1-score
DistilBERT	0.59	0.35	0.35	0.33	0.53	0.59	0.55
RoBERTa	0.60	0.36	**0.37**	**0.36**	0.54	0.60	**0.57**
DeBERTa	**0.61**	**0.45**	0.36	**0.36**	**0.58**	**0.61**	**0.57**
Random	0.13	0.12	0.11	0.10	0.25	0.13	0.15

Table 10: Emotion prediction, PELD

Appendix B: Personality prediction

Models for conscientiousness prediction

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted Precision	Weighted Recall	Weighted F1-score
DistilBERT	0.38	0.16	0.15	0.11	0.28	0.38	0.24
RoBERTa	**0.40**	0.16	0.17	0.14	0.28	**0.40**	0.29
DeBERTa	0.37	0.23	**0.19**	0.17	**0.32**	0.37	**0.31**
SVM raw	0.27	0.20	0.18	0.18	0.27	0.27	0.27
SVM TF-iDF	0.29	**0.33**	0.18	**0.19**	0.31	0.29	0.28
Log. Regression raw	0.29	0.31	0.18	**0.19**	0.31	0.29	0.28
Log. Regression TF-iDF	0.34	0.16	0.15	0.13	0.27	0.34	0.26
Random	0.13	0.13	0.10	0.10	0.22	0.13	0.16

Table 11: Conscientiousness prediction, WASSA 2022

Models for openness prediction

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted Precision	Weighted Recall	Weighted F1-score
DistilBERT	**0.33**	0.16	**0.19**	0.15	0.24	**0.33**	0.25
RoBERTa	0.27	0.11	0.16	0.11	0.18	0.27	0.18
DeBERTa	**0.33**	**0.24**	**0.19**	0.17	**0.31**	**0.33**	0.26
SVM raw	0.28	0.19	**0.19**	**0.19**	0.27	0.28	**0.27**
SVM TF-iDF	0.26	0.17	0.17	0.17	0.24	0.26	0.25
Log. Regression raw	0.28	0.18	**0.19**	0.18	0.26	0.28	**0.27**
Log. Regression TF-iDF	0.27	0.18	0.15	0.13	0.24	0.27	0.21
Random	0.13	0.12	0.12	0.11	0.18	0.13	0.15

Table 12: Openness prediction, WASSA 2022

Models for extraversion prediction

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted Precision	Weighted Recall	Weighted F1-score
DistilBERT	**0.27**	**0.25**	**0.24**	**0.21**	0.24	**0.27**	**0.22**
RoBERTa	0.20	0.11	0.18	0.12	0.13	0.20	0.14
DeBERTa	0.25	0.24	0.23	0.20	**0.27**	0.25	0.21
SVM raw	0.20	0.19	0.19	0.18	0.20	0.20	0.19
SVM TF-iDF	0.20	0.21	0.20	0.19	0.21	0.20	0.20
Log. Regression raw	0.22	0.22	0.21	**0.21**	0.22	0.22	0.21
Log. Regression TF-iDF	0.21	0.13	0.18	0.14	0.15	0.21	0.17
Random	0.13	0.13	0.13	0.13	0.14	0.13	0.14

Table 13: Extraversion prediction, WASSA 2022

Models for stability prediction

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted Precision	Weighted Recall	Weighted F1-score
DistilBERT	**0.26**	0.14	**0.19**	0.16	0.19	**0.26**	**0.21**
RoBERTa	0.25	0.19	0.18	0.16	**0.23**	0.25	**0.21**
DeBERTa	0.20	0.14	0.16	0.14	0.18	0.20	0.18
SVM raw	0.22	0.18	0.17	0.17	0.21	0.22	**0.21**
SVM TF-iDF	0.22	0.16	0.17	0.16	0.20	0.22	0.20
Log. Regression raw	0.22	**0.21**	0.18	**0.19**	0.21	0.22	**0.21**
Log. Regression TF-iDF	0.24	0.09	0.15	0.09	0.13	0.24	0.14
Random	0.12	0.11	0.10	0.10	0.13	0.12	0.12

Table 14: Stability prediction, WASSA 2022

Models for agreeableness predictions

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score	Weighted Precision	Weighted Recall	Weighted F1-score
DistilBERT	**0.37**	0.05	0.14	0.08	0.13	**0.37**	0.20
RoBERTa	**0.37**	0.14	0.16	0.12	0.24	**0.37**	0.24
DeBERTa	**0.37**	**0.25**	**0.20**	0.18	**0.34**	**0.37**	**0.31**
SVM raw	0.27	0.17	0.18	0.17	0.27	0.27	0.27
SVM TF-iDF	0.33	0.20	0.19	0.19	0.31	0.33	**0.31**
Log. Regression raw	0.32	0.21	**0.20**	**0.20**	0.31	0.32	**0.31**
Log. Regression TF-iDF	0.34	0.15	0.16	0.14	0.25	0.34	0.27
Random	0.12	0.11	0.10	0.09	0.20	0.12	0.14

Table 15: Agreeableness prediction, WASSA 2022

Appendix C: Models with continuous outcomes

Big Five Personality trait	R squared for Linear Regression model
Conscientiousness	-0.34
Openness	-0.32
Extraversion	-0.32
Agreeableness	-0.36
Stability	-0.29

Table 16: Linear Regression, WASSA 2022

Big Five Personality trait	R squared for Neural Network model
Conscientiousness	-0.14
Openness	-0.16
Extraversion	-0.19
Agreeableness	-0.23
Stability	-0.16

Table 17: Neural Network model, WASSA 2022

Appendix D: Heatmaps with correlations

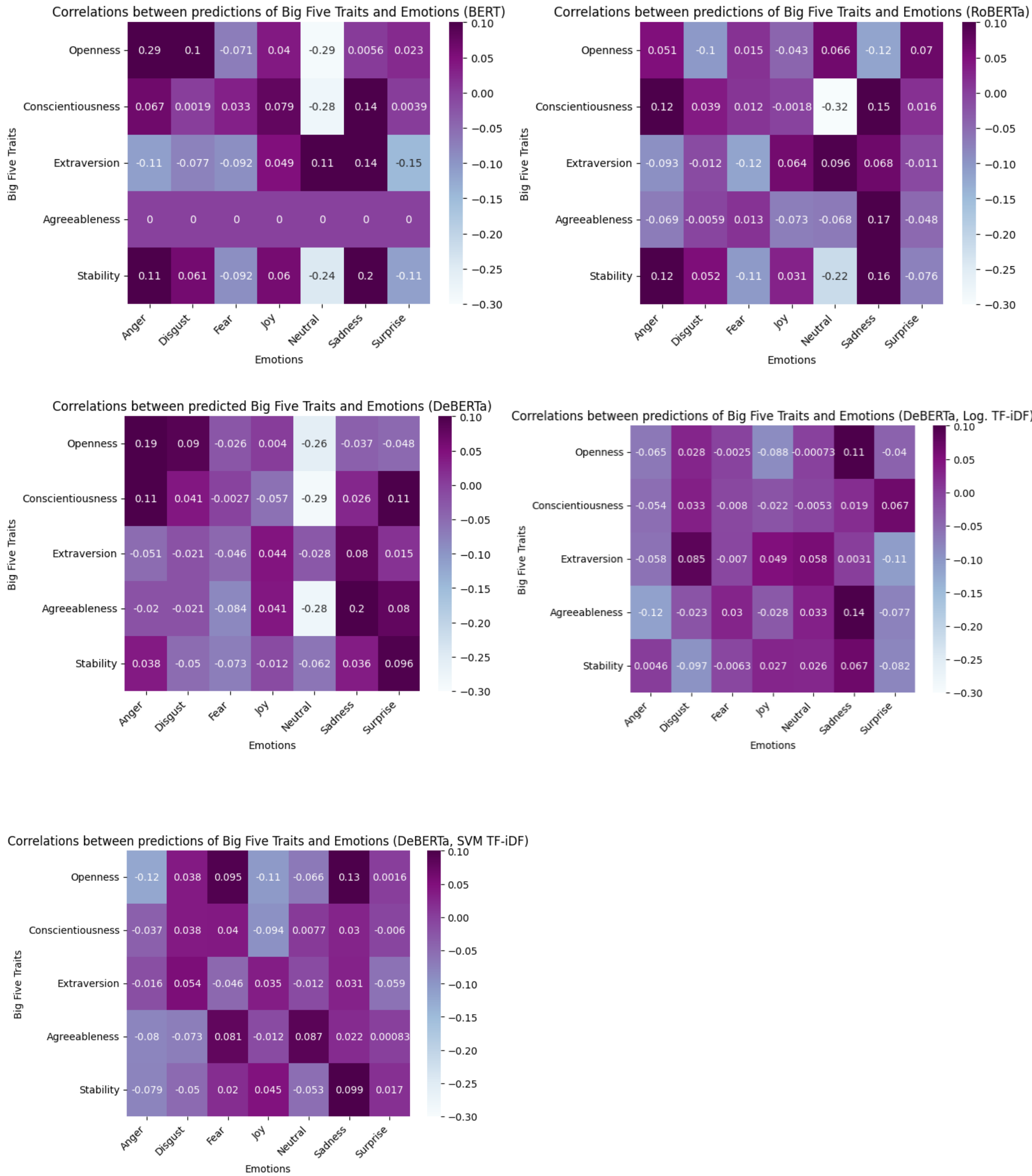

Figure 7: Heatmaps with correlations

Appendix E: Comparisons of correlations

Comparison: original and DistilBERT

Correlations coefficient	0.504
MAE	0.083
RMSE	0.106
Wilcoxon	No significant difference

Comparison: original and RoBERTa

Correlations coefficient	0.508
MAE	0.067
RMSE	0.089
Wilcoxon	No significant difference

Comparison: original and DeBERTa

Correlations coefficient	0.709
MAE	0.061
RMSE	0.008
Wilcoxon	No significant difference

Comparison: original and SVM (TF-iDF)

Correlations coefficient	0.268
MAE	0.054
RMSE	0.067
Wilcoxon	No significant difference

Comparison: original and Logistic Regression (TF-iDF)

Correlations coefficient	0.392
MAE	0.050
RMSE	0.060
Wilcoxon	No significant difference

Table 18: Comparisons of correlations

Appendix F: F1-Scores for emotion prediction

WASSA 2022

Model's F1-Score	Anger	Disgust	Fear	Joy	Neutralism	Sadness	Surprise
DistilBERT F1-Score	0.54	0.12	0.45	0.17	0.54	0.71	0.45
RoBERTa F1-Score	**0.62**	0.06	0.60	0.23	0.56	**0.76**	**0.51**
DeBERTa F1-Score	0.57	**0.33**	**0.64**	**0.47**	**0.64**	**0.76**	0.46

Table 19:F1-Scores for emotion prediction, WASSA 2022

PELD

Model's F1-Score	Anger	Disgust	Fear	Joy	Neutralism	Sadness	Surprise
DistilBERT F1-Score	0.31	0.00	0.00	0.56	**0.76**	0.18	**0.47**
RoBERTa F1-Score	0.37	0.00	0.00	**0.61**	**0.76**	0.30	0.46
DeBERTa F1-Score	**0.38**	0.00	**0.04**	0.59	**0.76**	**0.37**	0.41

Table 20: F1-Scores for emotion prediction, PELD