



OPEN ACCESS

EDITED BY
Fabrizio Riguzzi,
University of Ferrara, Italy

REVIEWED BY
Yoram Bachrach,
DeepMind Technologies Limited,
United Kingdom
Cheng Peng,
University of Florida, United States

*CORRESPONDENCE
Fatima Habib
✉ fatima.habib@nu.edu.pk;
✉ fatima.habib0923@gmail.com

RECEIVED 05 March 2024
ACCEPTED 28 August 2024
PUBLISHED 13 September 2024

CITATION

Habib F, Ali Z, Azam A, Kamran K and Pasha FM (2024) Navigating pathways to automated personality prediction: a comparative study of small and medium language models. *Front. Big Data* 7:1387325. doi: 10.3389/fdata.2024.1387325

COPYRIGHT

© 2024 Habib, Ali, Azam, Kamran and Pasha. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Navigating pathways to automated personality prediction: a comparative study of small and medium language models

Fatima Habib^{1*}, Zeeshan Ali², Akbar Azam¹, Komal Kamran¹ and Fahad Mansoor Pasha³

¹FAST School of Management, National University of Computer and Emerging Sciences, Lahore, Pakistan, ²Oxford Brookes Business School, Oxford Brookes University, Oxford, United Kingdom, ³Faculty of Business Administration, Lahore School of Economics, Lahore, Pakistan

Introduction: Recent advancements in Natural Language Processing (NLP) and widely available social media data have made it possible to predict human personalities in various computational applications. In this context, pre-trained Large Language Models (LLMs) have gained recognition for their exceptional performance in NLP benchmarks. However, these models require substantial computational resources, escalating their carbon and water footprint. Consequently, a shift toward more computationally efficient smaller models is observed.

Methods: This study compares a small model ALBERT (11.8M parameters) with a larger model, RoBERTa (125M parameters) in predicting big five personality traits. It utilizes the PANDORA dataset comprising Reddit comments, processing them on a Tesla P100-PCI-E-16GB GPU. The study customized both models to support multi-output regression and added two linear layers for fine-grained regression analysis.

Results: Results are evaluated on Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), considering the computational resources consumed during training. While ALBERT consumed lower levels of system memory with lower heat emission, it took higher computation time compared to RoBERTa. The study produced comparable levels of MSE, RMSE, and training loss reduction.

Discussion: This highlights the influence of training data quality on the model's performance, outweighing the significance of model size. Theoretical and practical implications are also discussed.

KEYWORDS

automated personality prediction, Big Five personality model, natural language processing, social media text, multi-output regression, large language models, psychometric analysis

1 Introduction

Human personality, a multifaceted account of behavior and traits, has been a fascinating subject for psychology researchers for several decades. Human personality acts as a driving force for steering behavior, emotional responses, and social interactions (Youyou et al., 2015). Digital records of human behavior have been empirically proven to offer effective personality assessments (Quercia et al., 2011; Kosinski et al., 2013), directly correlating with better decision-making (Letzring and Human, 2014). Various studies have highlighted the potential of Natural Language Processing (NLP) (Christian et al., 2021; Berggren et al., 2024) and empirically predicted personality traits using Machine Learning

(ML) (Tadesse et al., 2018) and deep learning models (Tandera et al., 2017; Yu and Markov, 2017). With ongoing advancements in computation, Large Language Models (LLMs) have experienced a significant surge in popularity. Pre-trained LLMs outperform state-of-the-art models in downstream NLP tasks, particularly personality prediction (Kazameini et al., 2020; Theil et al., 2023). Several studies indicate that Automated Personality Prediction (APP) with LLMs can significantly improve the accuracy and responsiveness of computing decisions, resulting in valuable insights (Jukić et al., 2022; Matz et al., 2023; Peters and Matz, 2023).

Contemporary technologies with big data have significantly improved automated predictions of personality (Ihsan and Furnham, 2018). This development continues to reshape the societal processes across diverse domains including social media, online education, business functions, and the electoral process (Alexander et al., 2020). Various researchers have executed APP in diverse contexts from the study of CEO risk-taking personalities, personality-job fit, and brand-follower personality matches to music recommendation systems tailored to personality (Wynekoop and Walz, 2000; Tang et al., 2018; Tomat et al., 2021; Kleć et al., 2023; Theil et al., 2023).

Traditionally, personality has been analyzed using a personality inventory of self-report questionnaires (Zhong et al., 2018). Amidst all potential sources of personality data, social media platforms stand out the most, given the frequency of their use by immensely large and varying demographics (Alexander et al., 2020). In behavioral sciences, self-report scales are often affected by social desirability bias as people tend to give socially acceptable responses that are different from their true feelings (Podsakoff et al., 2003). Conversely, on social media, users exhibit more naturalistic behavior as they are not being observed (Funder, 2012), resulting in a potentially better reflection of their personality. Since the digital footprint allows data collection of a larger population in real-time, individual as well as group-level dynamics can be better understood (Golder and Macy, 2014). In this era of rapid computational advancement, APP with its hidden intricacies, has arisen as a promising frontier in learning human psychology. To support APP, the linguistic information from our everyday language can help us effectively draw inferences about a person's personality (Kulkarni et al., 2018). Scholars have proposed various theoretical underpinnings to explain complex human personality. Among various models, the Big Five model is widely used as a robust and meticulous framework (Poropat, 2009). Table 1 exhibits the Big Five traits, their corresponding social aspects, and specific words used (Yarkoni, 2010; López-Pabón and Orozco-Arroyave, 2022).

Various statistical and Machine Learning (ML) methods have been developed and tested, producing APP knowledge. These include Support Vector Machines (SVM) (Berggren et al., 2024), Naïve Bayes (Cui and Qi, 2017), and booster classifiers (Tadesse et al., 2018). Deep learning models including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Recurrent Neural Networks (RNN) have also been investigated for predicting personality through NLP (Majumder et al., 2017; Tandera et al., 2017; Yu and Markov, 2017; Deilami et al., 2022). With promising improvements in newer technologies, research on APP is continuing to gain momentum.

TABLE 1 Big Five traits with corresponding social aspects and frequently used words.

Big Five trait	Social aspect explained	Frequently used words
Conscientiousness	Careless vs. self-disciplined	Completed, adventure, boring
Openness to experience	Dull vs. intellectual	Folk, humans, art, poetry, culture
Agreeableness	Uncooperative vs. obliging	Wonderful, together, felt, morning
Extraversion	Shy vs. friendly	Restaurant, dancing, shots, bar
Neuroticism	Emotionally unstable vs. calm	Awful, lazy, worse, irony, depressing

In the present-day context, there has been a dramatic increase in transformers-based LLM-related investigations (Kumar and Renuka, 2023). While text analytics has gained considerable attention in research, it can be refined further with fine-tuning trials using LLMs demonstrating state-of-the-art performance (Kjell et al., 2023). This refinement enhances real-world applications, particularly downstream NLP tasks such as text classification and sentiment analysis due to the capability of LLMs to focus on the context of words (Lewis et al., 2020; Kjell et al., 2023). Transformers (Vaswani et al., 2017) outperform previous techniques in text analysis due to the self-attention mechanism for capturing long-distance dependencies; parallel training for faster computation; and versatile capabilities facilitated by their unique architecture. Various studies have experimented with diverse models, such as BERT (Devlin et al., 2019; Zhao and Wong, 2023), RoBERTa (Putra and Setiawan, 2022), ULMFiT (El-Demerdash et al., 2021), and DistilBERT (Sanh et al., 2019), to harness the full potential of transformers (Rajapaksha et al., 2021).

Despite the sheer rise in LLMs, their deployment poses a few major challenges. One of the primary obstacles is the ever-increasing consumption of scarce resources, as these models are large and data-intensive (Schick and Schütze, 2020; Fu et al., 2023; Hsieh et al., 2023). These sizeable models extensively consume computational and memory resources (Hsieh et al., 2023). Thereby, contributing to energy inefficiency and a large carbon footprint through carbon dioxide equivalent emissions (CO₂e) into the environment (Schick and Schütze, 2020; Patterson et al., 2021). Comparative estimates of CO₂e exhibit a significant difference between the T5 model (11 billion parameters) and the GPT-3 (175 billion parameters), with 46.7 and 552.1 metric tons of CO₂e, respectively (Patterson et al., 2021). Additionally, LLMs are also associated with the substantial water footprint crisis, an often-overlooked environmental threat (Li et al., 2023; Rillig et al., 2023). Conversely, smaller language models offer an opportunity to mitigate the risks accompanying LLMs. These models can be fine-tuned to be computationally efficient, match, or sometimes even surpass the accuracy of larger models (Kazameini et al., 2020).

Previous studies have predominantly emphasized increasing LLM size as a means to increase accuracy. Nevertheless, ongoing research is centered on attaining higher accuracy with smaller

models. This study focuses on fine-tuning a small language model (Albert-Base 2 with 11.8 million parameters) in comparison with a larger one (Roberta-Base with 125 million parameters) for APP. By comparing these two language models for multi-output personality prediction, this research makes the following contributions to the existing literature on APP. First, the comparative analysis provides insights into the optimal selection of a language model offering a comparative level of error reduction for APP. Second, the results offer perspectives on minimizing computational resource constraints encompassing time consumption, heat emission, and computational power usage. Third, it examines the viability of predicting personality using a continuous scale for each of the five traits of the Big Five Model, through multi-output regression. As recommended in several researches (Feizi-Derakhshi et al., 2022; Johnson and Murty, 2023).

2 Background and literature

“The web sees everything and forgets nothing” (Golder and Macy, 2014). The digital footprint has enabled the field of computational social science to extend its mining into human behavior on a massive scale. This big data analysis has led to the fine-grained investigation of critical phenomena such as social network analysis (Letzring and Human, 2014), public opinion (Christian et al., 2021; Berggren et al., 2024), and social influence on political mobilization during electoral events (Cui and Qi, 2017; Tandra et al., 2017; Tadesse et al., 2018). Moreover, time spent on social media can help decipher the emotional states of smartphone users’ indicating boredom and loneliness (Kazameini et al., 2020; Theil et al., 2023). In turn, such emotional states and tones can help recognize user demographic traits (Volkova and Bachrach, 2016). While social influence can be assessed from the friends on a social media network, the use of language can be indicative of user intention leading to the detection of depression and suicidal tendencies (Jukić et al., 2022; Matz et al., 2023; Peters and Matz, 2023). Social media data, when utilized responsibly, carries immense potential to significantly benefit the community through data-driven decision-making.

Pivoting to the business sphere, social media has acted as a catalyst for fostering analytical insights across diverse business functions. Consumer sentiments and attitudes toward culturally diverse brands can be inferred from the data available on social networks and blogs (Alexander et al., 2020). Such data can offer valuable insights into user interests across a diverse spectrum of health, religion, movies, music, and arts (Lewenberg et al., 2015). Emotional drivers and user influence on social media can be assessed and integrated into information systems for enhanced business decision-making (Chung and Zeng, 2020). Data harnessed from social media discourse can also help us infer demographic information such as income, gender, opinions, sentiments, and personality traits (Wynekoop and Walz, 2000; Volkova et al., 2016; Hinds and Joinson, 2018; Tang et al., 2018; Tomat et al., 2021; Kleć et al., 2023; Theil et al., 2023). Such pertinent insights from digital traces (as exhibited in Figure 1) can be extremely beneficial when implemented in targeted advertising, marketing, and customer relationships (Kosinski et al., 2014; Matz et al., 2017). While digital

footprints can be utilized in these diverse applications, this paper primarily focuses on their use in APP.

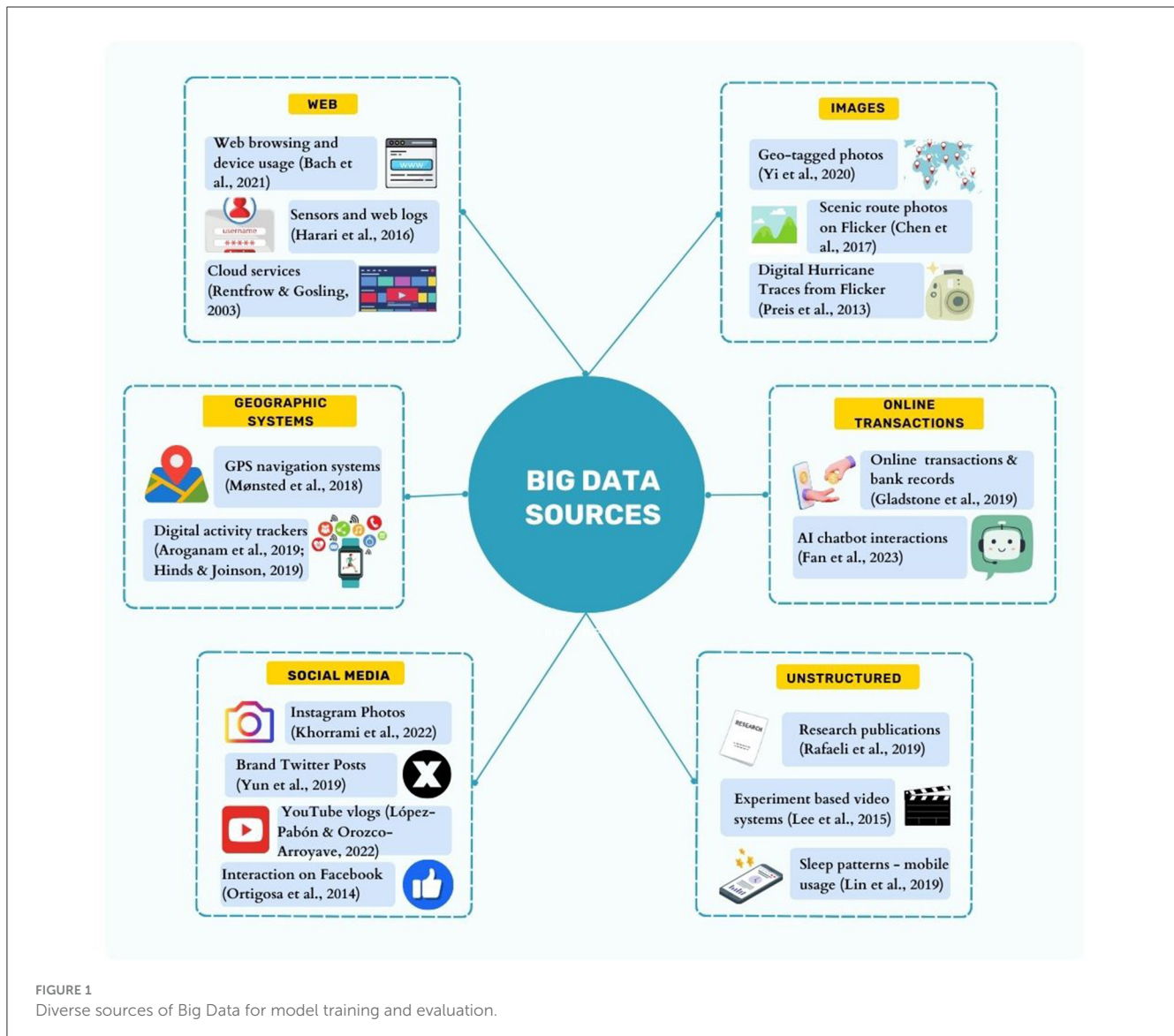
2.1 Automated personality prediction

Personality has been a subject of extensive research, which has significantly contributed to our understanding of human behavior and actions (Putra and Setiawan, 2022). Funder (2012) describes personality as patterns of emotions, thoughts, and behavior consistent across situations and over time. Personality distinguishes individuals and forms clusters of individuals to reveal consistent behavioral patterns (Kulkarni et al., 2018). It refers to the characteristic amalgamations of emotional reactions developed from biological makeup and circumstantial factors, resulting in consistent differences (Karanatsiou et al., 2022).

Numerous studies have investigated the diverse aspects of human personality. Among various others such as the Personality Enneagrams (Sutton et al., 2013) and Myers Briggs Type Indicator (MBTI) (Gjurković and Šnajder, 2018), the Big Five model is appreciated as a widely used framework for psychological assessment (Zimmermann et al., 2019; Gjurkovic et al., 2021). It provides scores for five OCEAN (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism) traits that rate the respondents on a continuum (McCrae and John, 1992). The Big Five personality model is deemed one of the most credible and widely known scales for gauging personality against five traits of natural language for fine-grained analysis (Goldberg, 1982; Phan and Rauthmann, 2021). In the new technologically advanced global economy, the upsurge in digital data has posed immense opportunities for data utilization. Digital footprint entails several means of assessing personality. Personality traits can be expressed through likes, comments, musical preferences (Nave et al., 2018), pictures (Segalin et al., 2017), product selection (Hirsh et al., 2012), text, and language (Tutaysalgiir et al., 2019; Mehta et al., 2020a). Together, these studies indicate the pertinent role of digital data in personality detection.

2.2 Personality and prediction models

Early research on APP focused on traditional statistical methods. Tausczik and Pennebaker (2010) used Linguistic Inquiry and Word Count (LIWC) to perform a psychometric analysis of words. Psycholinguistic feature extraction formed the foundation for Decision Trees, SVM, Naïve Bayes, and several other traditional machine-learning classifiers for personality identification analyzing text (Quercia et al., 2011; Alam et al., 2013; Markovikj et al., 2013; Mohammad and Kiritchenko, 2013). Later, ensemble models, basic neural networks, and text-mining experiments elevated the sphere of APP (Peng et al., 2015; Cui and Qi, 2017; Tadesse et al., 2018; Yang and Huang, 2019). Taken together, these studies highlight an important theme of personality identification, opening doors to the continuous evolution of methodology. In essence, the literature largely focuses on personality detection employing classification methods, with limited attention given to personality prediction as a multi-output regression problem.



Literature has extensively explored the construct of personality prediction from social media texts. Deep learning methods are expected to yield superior results owing to their automatic feature extraction capabilities, unique structures, noteworthy performance, and relatively low computational cost (Xue et al., 2018; Deilami et al., 2022). It is worth mentioning that, in recent years, notable deep neural networks exemplifying personality prediction have emerged. The neural networks employed in APP encompass embeddings, Convolutional Neural Networks, Recurrent Neural Networks, and Long Short-Term Memory Networks with various experiments in combinations (Majumder et al., 2017; Tandra et al., 2017; Yu and Markov, 2017; Sun et al., 2018; Xue et al., 2018). Following the escalating momentum in deep learning, Mehta et al. (2020b) reviewed the advances in APP based on deep learning, This review highlighted that the quality of training data is a significant determinant of the strength of ML models.

What follows is a comprehensive account of the latest developments in the field with pre-trained LLMs built on transformers' attention mechanism. Transformers have significant advantages over previous methods of text analysis such as RNNs

and CNNs. Their self-attention mechanism captures long-range dependencies and contextual information in the text (Vaswani et al., 2017), while the parallelized architecture leads to faster training and inference times (Fan et al., 2023b). Moreover, the encoder-decoder framework, bidirectional context, and multihead attention enable transformers to effectively handle diverse NLP tasks (Rothman, 2021). LLM inference requires lower programming and computational resources after being fine-tuned (Church et al., 2021). Since the base models are generic, they are required to be fine-tuned for a specific downstream task at some point. Fine tuning implies the capability to change the layers expressing more control over the output (Church et al., 2021). Prior fine-tuning using various training procedures can help reduce the burden on the computational resources at the inference stage which consumes more resources as compared to the training process (Schick and Schütze, 2020; Strubell et al., 2020). Despite the significance of fine-tuning LLMs, a few researchers such as Tomat et al. (2021) and Wynekoop and Walz (2000) have started exploring the viability of zero-shot learning without explicit training or fine-tuning it for the specific task. However, research on zero-shot learning is still

in the nascent stages, as far as its reported response accuracies are concerned.

LLMs exhibit strong transferability facilitating emotion detection outperforming state-of-the-art models (Peng et al., 2024). Empirical investigations on Bard, GPT, and TAS have claimed that LLMs when trained on sufficient data can achieve or even surpass the benchmark of human emotional intelligence (Patel and Fan, 2023). Such progressions have facilitated much complex Emotional Support Conversations which can prominently elevate customer service chats, counseling, psychotherapy, and mental health support (Zhang et al., 2024). GPT 3.5 has been fine-tuned to detect the target demographic aligning itself when used in public-dealing tools showing varying accuracies (Sicilia et al., 2024). Fine-tuning BERT-like LMs on text authored by balanced demographic cohorts can mitigate the biases that usually arise (Garimella et al., 2022). Alpaca has been successfully fine-tuned to predict survey responses escalating opinion mining and trend analysis of diverse social issues (Kim et al., 2023). The proliferation of LLMs has facilitated a wide array of applications in diverse fields. Among these, a large body of research emphasized the automated prediction of personality through LLMs. Using transfer learning, various investigations have been conducted employing attentive networks such as BERT, RoBERTa, and XLNet with varying combinations, evaluated on text datasets including FriendsPersona, Reddit comment, and Facebook (Jiang et al., 2020; Kazameini et al., 2020; Christian et al., 2021; Gjurkovic et al., 2021).

One of the drawbacks observed in previous studies was the concentration on the classification of personality traits into labels. However, this does not accurately reflect the continuous nature of these traits in humans (Mehta et al., 2020a; Feizi-Derakhshi et al., 2022; Johnson and Murty, 2023). By contrast, predicting personality using regression analysis provides a more realistic representation of an individual's traits. Only a few studies have employed regression as a technique to assess Big Five personality traits (Xue et al., 2018; López-Pabón and Orozco-Arroyave, 2022). Following the proposed methods in the aforementioned studies, our research is focused on multi-output regression as a method to assess the Big Five personality traits from the text analysis of Reddit comments.

Additionally, despite all the developments made by APP using various models as shown in Table 2, there are a few challenges in the continuation of research on LLMs and their application in the industry settings. López-Pabón and Orozco-Arroyave (2022) assert that LLMs tend to be computationally expensive, water-intensive, costly, and time-consuming. They also suggest that further research is needed to compare large and small models. A few studies have advocated the importance of smaller and computationally efficient models for real-time inference. By specializing in specific tasks using model specialization, smaller models derived from larger ones can achieve comparable performance, effectively leveraging the capabilities of large models (Araci, 2019; Yang et al., 2020; Fu et al., 2023). In essence, smaller models have been proposed to achieve similar levels of accuracy, sometimes outperforming larger models (Hsieh et al., 2023). They also require less training data, are more computationally efficient, produce less carbon footprint, and are ultimately cost-effective (Schick and Schütze, 2020; Fu et al., 2023). Therefore, the research is shifting toward the viability of smaller language models with comparable performance. In this

study, we are aiming to compare a smaller language model with a larger one, in terms of training loss and resource consumption.

Several recent studies have focused on APP from text and other sources such as images, voice, and video (Kazameini et al., 2020; Moreno et al., 2021). However, it is still at a developing stage in the field of business and artificial intelligence. Previous studies have predominantly emphasized increasing LLM size as a means to increase accuracy. The majority of which have employed classification techniques for personality predictions. Contrary to these researches, we propose a small pre-trained language model ALBERT in comparison with a larger pre-trained language model RoBERTa for personality prediction from the text of Reddit comments. Unlike the classification approach, we also propose multi-output regression to produce continuous prediction scores for five personality traits simultaneously. We evaluate the models based on loss function particularly, error reduction. Additionally, we assess the training process differences given computational GPU resource consumption, commitment, and heat emission. We demonstrate that small language models exhibit performance comparable to ten times large language models in APP owing to the similarity in training data quality.

3 Materials and methods

We aim to compare the performance of a large and a small language model by examining the training loss, time-to-train of the pre-trained models, and GPU computation resources consumed while keeping the training parameters constant to train both models.¹ Figure 2 illustrates the flow chart of our study detailing the flow of data preprocessing, the architectural differences between RoBERTa and ALBERT, and the customization of the model at the last layer.

3.1 Dataset

The dataset used for training and fine-tuning this model is titled Personality And Demographics Of Reddit Authors (PANDORA). Gjurković and Šnajder (2018) presented this large-scale dataset collected from the social media platform Reddit.com. Despite being a popular discussion website, it is often overlooked for personality prediction tasks. It encompasses a wide array of topics while preserving user anonymity (Jukić et al., 2022). It started with MBTI categories for 9,000 Reddit users. Later, Gjurkovic et al. (2021) added the Enneagrams and the Big Five scores which resulted in a comprehensive collection of comments posted on reddit.com by 10,288 users. This dataset has been used to support various research focused on predicting personality from text. Li et al. (2021) used it for personality prediction through multi-task learning whereas Jukić et al. (2022) utilized PANDORA to explore the relationship between evaluative language and personality traits. Moreover, Radisavljević et al. (2023) attempted to create a similarity

¹ The model training is tracked using the platform *Weights and Biases*. This model workflow can be accessed at <https://wandb.ai/nlp-thesis/huggingface/workspace?workspace=user-fatima-habib>.

TABLE 2 Models used in previous research and their evaluation metrics.

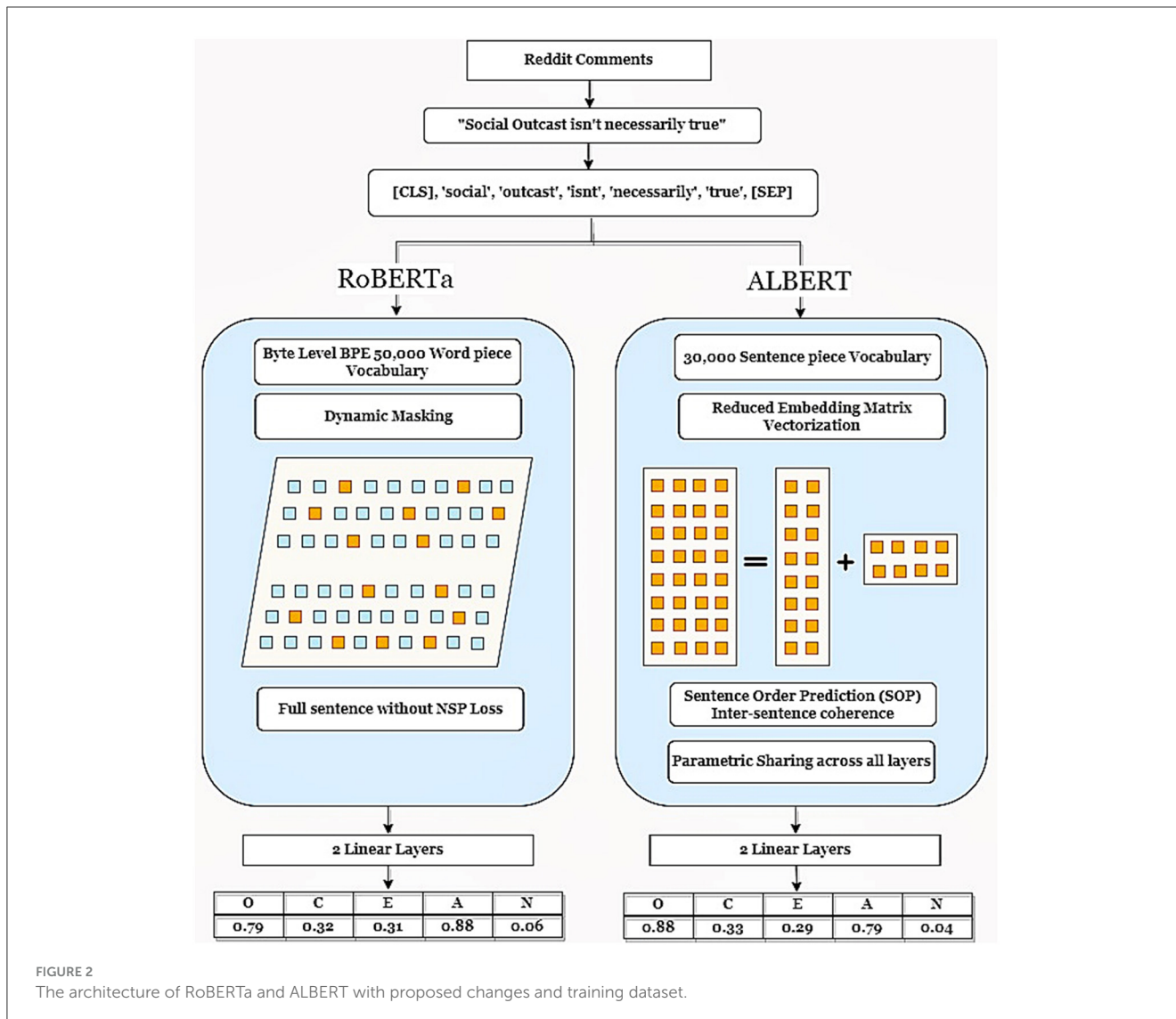
Research	Model	C	R	Evaluation metrics
Statistical and machine learning models				
Quercia et al. (2011)	M5 algorithm		✓	RMSE
Markovikj et al. (2013)	Simple minimal optimization (SMO) and boost algorithms with POS Tag, AFINN and H4LVD parameters	✓		TP rate, FP rate, precision, recall, ROC Area
Mohammad and Kiritchenko (2013)	SVM	✓		Recall, accuracy and F1-score
Alam et al. (2013)	SMO (sequential minimal optimization for SVM), Bayesian Logistic Regression (BLR) and Multinomial Naïve Bayes (MNB) sparse modeling	✓		Macro-averaged precision, recall and F1; weighted average accuracy (WA) and un-weighted average accuracy (UA)
Peng et al. (2015)	SVM	✓		Precision and recall
Cui and Qi (2017)	Comparison of SVM, NB classifier and a basic NN with ReLU activation function	✓		Accuracy
Tadesse et al. (2018)	XGBoost classifier	✓		Accuracy
Tutaysalgir et al. (2019)	Clustering algorithms	✓		Silhouette coefficient scores
Yang and Huang (2019)	M5 Regression Tree and SVM		✓	MSE
Deep learning models				
Majumder et al. (2017)	Word2vec embeddings with CNN layer	✓		Accuracy
Tandera et al. (2017)	MLP (multi-layer perceptron), LSTM, GRU (gated recurrent unit), and CNN 1D (1-dimensional)	✓		Accuracy
Xue et al. (2018)	Hierarchical DNN (AttRCNN) CNN-based inception structure		✓	MAE
Yu and Markov (2017)	Multi-model (FC, CNN with RNN)	✓		F1 and accuracy
Sun et al. (2018)	Multi-model CLSTM (CNN and BiLSTM)	✓		Precision
Deilami et al. (2022)	CNN with AdaBoost	✓		Accuracy and MAE
Quwaider et al. (2023)	Artificial NN	✓		Accuracy
Pre-trained language models				
Jiang et al. (2020)	RoBERTa	✓		Accuracy
Keh and Cheng (2019)	Fine-tuned BERT	✓		Accuracy
Kazameini et al. (2020)	BERT with Bagged-SVM classifier	✓		Accuracy
Gjurkovic et al. (2021)	BERT with NN		✓	r between Big 5 and MBTI results
López-Pabón and Orozco-Arroyave (2022)	Word2Vec, GloVe, and BERT (base and large)	✓	✓	Regression (R square, MAE and RMSE) and classification (F1, accuracy and AUC)
Mehta et al. (2020a)	BERT (base and large)	✓		Accuracy
Ramezani et al. (2022)	KGrAt-Net (knowledge graph attention network text classifier)	✓		Recall, precision, accuracy and F1-score
Theil et al. (2023)	BERT-base and RoBERTa-base		✓	MAE
Johnson and Murty (2023)	Enhanced knowledge graph with BERT	✓		Recall, precision, accuracy and F1-score
Matz et al. (2023)	GPT-3		✓	R
Peters and Matz (2023)	GPT-3.5 and GPT 4 with zero-shot learning		✓	R

R, regression; C, classification.

connection among multiple personality models including the Big Five, the Enneagrams, and MBTI.

We selected the Big Five model for our study and extracted author profiles of ~1,608 users, with a total of 27,859 comments from the dataset. The final dataset can be accessed at Habib (2024). Table 3 displays a statistical summary of the original PANDORA

Big Five subset while Table 4 presents the statistical account of each of the five traits. Before the training, the dataset was pre-processed to eliminate any noise that could distort the analysis (López-Pabón and Orozco-Arroyave, 2022). Reddit data, which were already anonymous, were further pseudonymized to protect the privacy of the authors, as suggested by Volodina et al. (2020).



3.2 Data pre-processing

The Natural Language Toolkit (NLTK) library of Python (Wang and Hu, 2021) was used to preprocess the text in the dataset.² The text was standardized by converting it into lowercase and removing any hyperlinks and URLs, punctuation, new lines, and special characters. Additionally, the text was tokenized to represent the input text as a sequence of word tokens. Lemmatization was avoided to preserve the linguistic context (Ramezani et al., 2022). Removing stop words has been shown to have no significant impact on the performance of LLMs (Qiao et al., 2019). Therefore, stop words have been retained to maintain the contextual integrity of natural language patterns, essential to LLMs functionality. Afterward, sentences with fewer than five words were filtered out, and non-English comments were removed through NLP's

² The source code for preprocessing text data and model training is available at: <https://github.com/Fatima0923/NLP>.

TABLE 3 PANDORA statistics summary (Big Five model).

Property	Count
Number of comments	27,859
Average word count	44.88
Maximum words	3,430
Minimum words	1
Average sentences	2.5
Maximum sentences	71
Minimum sentences	1

language detection process (Rajanak et al., 2023), using the LangDetect package in Python. The final dataset contained only English sentences to maintain uniformity for better fine-tuning (see Table 5).

TABLE 4 Statistics of personality scores for five traits.

Property	Agreeableness	Openness	Conscientiousness	Extraversion	Neuroticism
Count	27,859.00	27,859.00	27,859.00	27,859.00	27,859.00
Mean	37.213450	67.675365	30.006605	35.726139	49.522722
Std	29.833761	22.208940	27.382423	30.034822	29.106680
Min	0.00	9.00	1.00	0.00	0.00
25%	9.00	50.00	9.00	7.00	20.00
50%	38.00	74.00	20.00	31.00	50.00
75%	57.00	85.00	43.00	59.00	72.00
Max	99.00	98.00	98.00	99.00	99.00

TABLE 5 Pre-processed text from the training dataset and corresponding Big Five Scores.

	Text	AGR	OPN	CON	EXT	NEU
1	His name was kim kimble originally wow thats some messed up parents	9	61	13	4	72
2	Theyre better than the normal posts on ryugioh id rather have them then the same topic posted multiple times in the week after the banlist	50	85	50	85	50
3	How the fuck does this even happen hi youre cute you too ive had a crush on you for a while um i uh inserts finger in butthole	15	85	15	85	15
4	It probably does ive learned a lot about myself by browsing this subreddit over the months	71	53	17	3	31
5	Yea those are the same sound to me still	64	44	33	8	88
6	Long term shifting is the cart titans gimmick though the fact that she can do it doesnt mean eren can	50	85	50	85	50
7	Texas is molly weasley i love it	79	84	86	53	1
8	Yeah those are good points my experiences with recruiting is all with really open ended type work that	85	95	15	50	15

AGR, agreeableness; OPN, openness; CON, conscientiousness; EXT, extraversion; NEU, neuroticism.

3.3 Model comparison and customization

This paper proposes a comparison of two pre-trained language models, RoBERTa and ALBERT. Each model was customized using two additional linear layers. Liu et al. (2019) introduced RoBERTa using transformers as the underlying mechanism (Kumar and Renuka, 2023). Liu et al. (2019) claim that RoBERTa has been trained on a large English corpus of more than 50,000 byte-level Byte-Pair encoding tokenized vocabulary the masking patterns were dynamically altered, adding to the robustness of the results by eliminating duplicate data during training. RoBERTa's focus is on understanding language. Hence, RoBERTa is deemed to be one of the top-performing models for predicting personality traits (Theil et al., 2023). On the other hand, ALBERT was selected to compare the results of RoBERTa with those of a more modestly sized model. As explained by Lan et al. (2019), ALBERT shares the same architecture as BERT, analogous to the training and fine-tuning processes. It uses matrix-factorized embeddings with sentence-order predictions to better comprehend sentence connections. It carries smaller embedding sizes and also shares parameters across all layers, requiring less memory to store the parameter weights. The cross-layer parameter sharing helps the model to converge faster and enhance parameter efficiency (Plummer et al., 2020). The ALBERT model was selected

owing to its smaller size and efficient handling of contextualized text representations.

To accomplish the multi-output regression proposed in this study, both RoBERTa and ALBERT models were customized to handle regression tasks with multiple outputs. ALBERT and RoBERTa models from the Hugging Face Transformers library were leveraged as the core component of the custom models. Pretrained on a large corpus, these models are capable of contextual understanding of language (Devlin et al., 2019). Linear layers were added to each model, followed by an activation function. Initially, the model processes an input text sequence and generates a contextualized representation labeled "the_last_hidden_state." The mean of this output is computed into "a pooled output" which represents the summary of the input text. This pooled output is then passed through the linear layers added, to produce the final regression output. The hidden and output layers of these models carry the free parameters which can be altered by adding new trainable layers and an output layer. Such customization can be viable for multi-output regression utilizing the transfer learning technique (Emami and Martinez-Munoz, 2024). The first linear layer mapped the hidden size of contextualized word embeddings into a 128-dimensional vector. This transformation helped in reducing the dimensionality and focusing on the most significant features. Afterward, a Rectified Linear Unit (ReLU) function is


```

AlbertForPreTraining(
  (albert): AlbertModel(
    (embeddings): AlbertEmbeddings(
      (word_embeddings): Embedding(30000, 128, padding_idx=0)
      (position_embeddings): Embedding(512, 128)
      (token_type_embeddings): Embedding(2, 128)
      (LayerNorm): LayerNorm((128,), eps=1e-12, elementwise_affine=True)
      (dropout): Dropout(p=0, inplace=False)
    )
    (encoder): AlbertTransformer(
      (embedding_hidden_mapping_in): Linear(in_features=128, out_features=768, bias=True)
      (albert_layer_groups): ModuleList(
        (0): AlbertLayerGroup(
          (albert_layers): ModuleList(
            (0): AlbertLayer(
              (full_layer_layer_norm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
              (attention): AlbertAttention(
                (query): Linear(in_features=768, out_features=768, bias=True)
                (key): Linear(in_features=768, out_features=768, bias=True)
                (value): Linear(in_features=768, out_features=768, bias=True)
                (attention_dropout): Dropout(p=0, inplace=False)
                (output_dropout): Dropout(p=0, inplace=False)
                (dense): Linear(in_features=768, out_features=768, bias=True)
                (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
              )
              (ffn): Linear(in_features=768, out_features=3072, bias=True)
              (ffn_output): Linear(in_features=3072, out_features=768, bias=True)
              (activation): NewGELUActivation()
              (dropout): Dropout(p=0, inplace=False)
            )
          )
        )
      )
      (pooler): Linear(in_features=768, out_features=768, bias=True)
      (pooler_activation): Tanh()
    )
    (predictions): AlbertMLMHead(
      (LayerNorm): LayerNorm((128,), eps=1e-12, elementwise_affine=True)
      (dense): Linear(in_features=768, out_features=128, bias=True)
      (decoder): Linear(in_features=128, out_features=30000, bias=True)
      (activation): NewGELUActivation()
    )
    (sop_classifier): AlbertSOPHead(
      (dropout): Dropout(p=0.1, inplace=False)
      (classifier): Linear(in_features=768, out_features=5, bias=True)
    )
  )
)

```

FIGURE 3
Structure of ALBERT model customized for regression with additional layers.

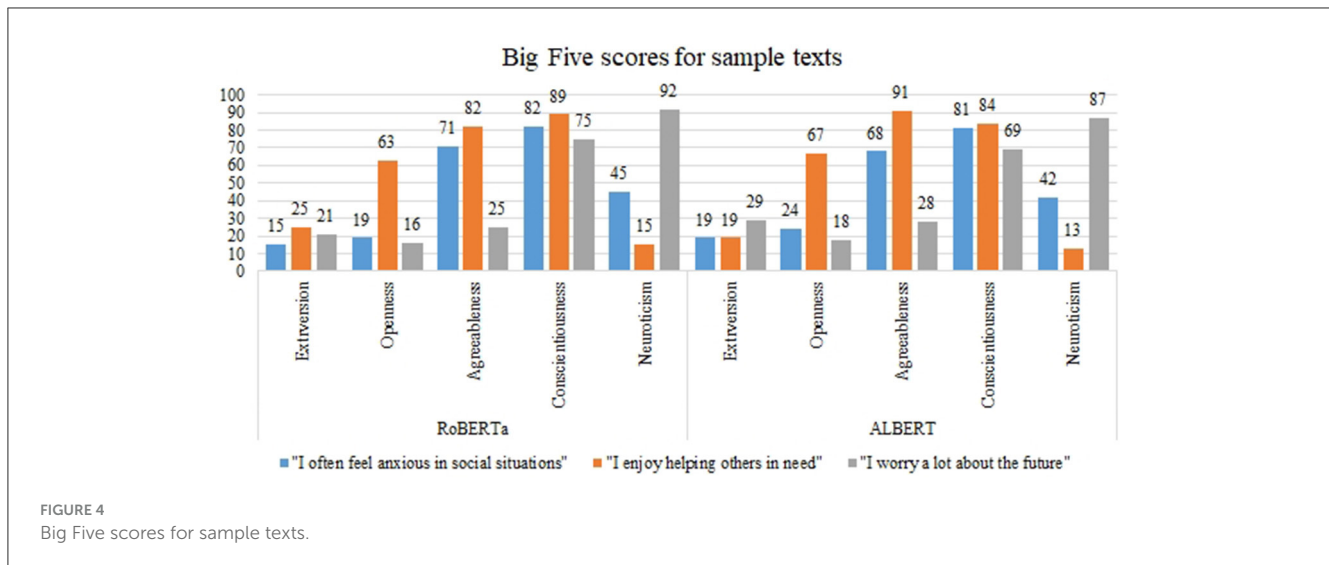
applied for the model to learn more complex language patterns. The second linear layer maintains the dimensionality at 128 units refining the learnings from the previous layer. The second linear layer was followed by a Tanh (Hyperbolic Tangent) which normalizes the representations into values ranging between -1 and 1 . This activation function thus stabilizes the learning process. The final linear layer was pivotal in the regression task as it mapped the 128-dimensional vector to the specific number of regression targets, five personality scores in this case (see [Figure 3](#) for ALBERT model structure customized for regression).

The models were configured with their respective tokenizers and custom regression layers. The tokenizer facilitates the embedding of tokens in a fixed representation in the vector space. These linear layers were instrumental in converting the model outputs into a more suitable form for regression. The architecture

was designed to learn and ultimately extract relevant features from textual data to predict personality traits as continuous values. These transformations were aimed at capturing the intricate patterns in the data.

3.4 Fine tuning

According to [Church et al. \(2021\)](#), pre-trained models are typically trained on unlabeled datasets for general purposes, whereas fine-tuning calls for the training of the base model on particular downstream tasks with labeled data. Fine-tuning enables us to modify only a few layers of the model's neural network for related but different specialized tasks ([Vrbanić and](#)



Podgorelec, 2020). In this research design, we used the PyTorch-based versions of RoBERTa and ALBERT with RoBERTa and ALBERT tokenizers for fine-tuning. The training configuration parameters included 40 training epochs, 16 batch sizes, and a maximum token length of 512. The evaluation strategy was set to epochs, and the learning rate was fixed at $2e-5$ for both models, with a weight decay of 0.01. These hyperparameters were selected based on a synthesis of previous research and empirical testing. Christian et al. (2021) experimented with $1e-5$ and $3e-5$ learning rates in addition to 16 and 32 batch sizes in various combinations. We initially used a learning rate of $1e-5$ and subsequently increased the value to $2e-5$, as supported by Yang et al. (2021) and El-Demerdash et al. (2022). Furthermore, we adopted a batch size of 16 which was substantiated by literature. Increasing the batch size further impeded the training process, due to the available GPU resources requirements. Additionally, El-Demerdash et al. (2022) also recommended setting the token length to a maximum of 512 tokens. A weight decay of 0.01 is generally recommended in pytorch documentation.³ Regarding the number of epochs, previous research has utilized a wide range, from three epochs (El-Demerdash et al., 2022) to 60 epochs (Deilami et al., 2022) have been employed. We selected 40 training epochs to sufficiently train the models, simultaneously staying within the designated resource limits. The purpose of this fine-tuning was to optimize the resultant performance of the models while making efficient use of the limited computational resources available. Thus, we initiated the process with values derived from previous works and fine-tuned the hyperparameters as the project progressed. These training arguments were implemented by employing a trainer-class API (Trainer API, 2023) for comprehensive feature training.

3.5 Evaluation metrics

The performance of the model was assessed by comparing its predictions with actual values. Evaluation metrics that can

discriminate between the method results were used (Deilami et al., 2022). The compute loss functions in RoBERTa and ALBERT were superseded. Functions from Python's Scikit-learn library provide regression metrics for evaluation, including Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). The MSE loss is commonly employed as a loss function in regression-based tasks. According to Yang and Huang (2019), a smaller MSE determines the effectiveness of the proposed model.

4 Analysis and results

In line with the previous research (Gjurkovic et al., 2021; Yang et al., 2021; Jukić et al., 2022), this study employs the extensive PANDORA dataset, which comprises 27,000 comments from 1,608 authors on the Reddit platform for APP. The dataset was leveraged to train the LLMs, including RoBERTa and ALBERT, with RoBERTa having ~ 10 times the number of parameters of ALBERT. Tesla P100-PCIE-16GB GPU was used to execute the two model trainings. The predictions entailed a continuous number for each of the Big Five traits, on a scale of 0–100. Furthermore, multi-output regression has been used as a mechanism for the simultaneous prediction of all five traits, as proposed in many studies (Xue et al., 2018; López-Pabón and Orozco-Arroyave, 2022). The execution of text pre-processing, and its input into ALBERT and RoBERTa with the subsequent output regression scores, are illustrated in Figure 2.

When evaluating sample texts, both models produced remarkably similar scores on Big Five traits. The predictions made by RoBERTa and ALBERT were consistently close, demonstrating agreement in their assessments. For instance, the input text "I prefer spending time alone with books" yielded comparable predictions with ALBERT scores of "38, 82, 63, 78, 27" and RoBERTa scores of "40, 81, 61, 75, 29" indicative of extroversion, openness, agreeableness, conscientiousness, and neuroticism respectively. These customized models produce an array of scores for each text input which can be used to comprehend its personality. Figure 4 shows a comparison of the scores predicted

³ <https://pytorch.org/docs/stable/optim.html#torch.optim.AdamW>

TABLE 6 Comparison of evaluation metrics between RoBERTa and ALBERT.

Epoch	RoBERTa					ALBERT				
	Training loss	Validation loss	MSE	RMSE	MAE	Training loss	Validation loss	MSE	RMSE	MAE
1	2,508.5298	2,465.8377	2,465.8364	48.9216	39.77474	2,499.2832	2,456.2976	2,456.3022	48.8059	39.6621
5	1,837.4262	1,811.1443	1,811.1420	41.9133	33.25334	1,831.1945	1,804.6882	1,804.6829	41.8195	33.1853
10	1,328.2593	1,305.9076	1,305.9078	35.7156	29.17931	1,325.3960	1,302.8401	1,302.8418	35.6608	29.1646
15	1,045.4495	1,041.1628	1,041.1605	32.0867	27.14586	1,045.0584	1,040.5450	1,040.5442	32.0710	27.1474
20	914.5754	915.7201	915.7194	30.2150	25.95275	915.2727	916.2628	916.2623	30.2219	25.9626
25	848.7351	856.8837	856.8833	29.2569	25.27504	849.6119	857.6313	857.6346	29.2697	25.2848
30	818.1030	828.8650	828.8630	28.7618	24.91060	818.8715	829.5237	829.5251	28.7741	24.9214
35	799.1436	816.3414	816.3398	28.5274	24.71394	799.7687	816.9009	816.9002	28.5383	24.7246
40	800.2508	812.7495	812.7492	28.4584	24.64961	800.8514	813.2720	813.2730	28.4687	24.6602

TABLE 7 Comparison of training hyperparameters in RoBERTa and ALBERT.

Model	RoBERTa	ALBERT
Train output		
Global step	40,440	40,440
Training loss	1,141.85	1,140.59
Metrics		
Train runtime	35,939.5414 (9:54:11)	37,620.0265 (10:25:41)
Train samples per second	17.996	17.192
Train steps per second	1.125	1.075
Total flos	0.0	0.0
Train loss	1,141.85	1,140.59
Epoch	40.0	40.0

by ALBERT and RoBERTa for the given sample texts across all the traits of the Big Five Model, showcasing very close results between the two models. Additionally, Table 6 demonstrates the training loss and reduction in MSE, RMSE, and MAE over 40 epochs while training RoBERTa and ALBERT on the PANDORA dataset. The aforementioned hyperparameters, including the number of epochs, evaluation strategy, learning rate, and batch size of the input, were kept constant for training both models. Despite the differences in their sizes, both models seem to produce similar results in terms of training loss and reduction in MSE, RMSE, and MAE.

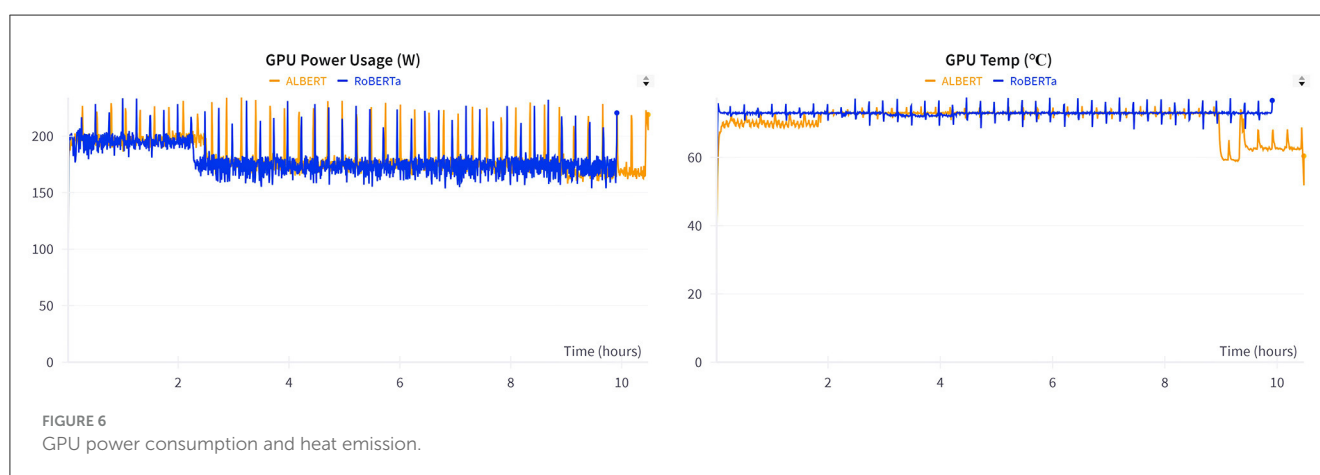
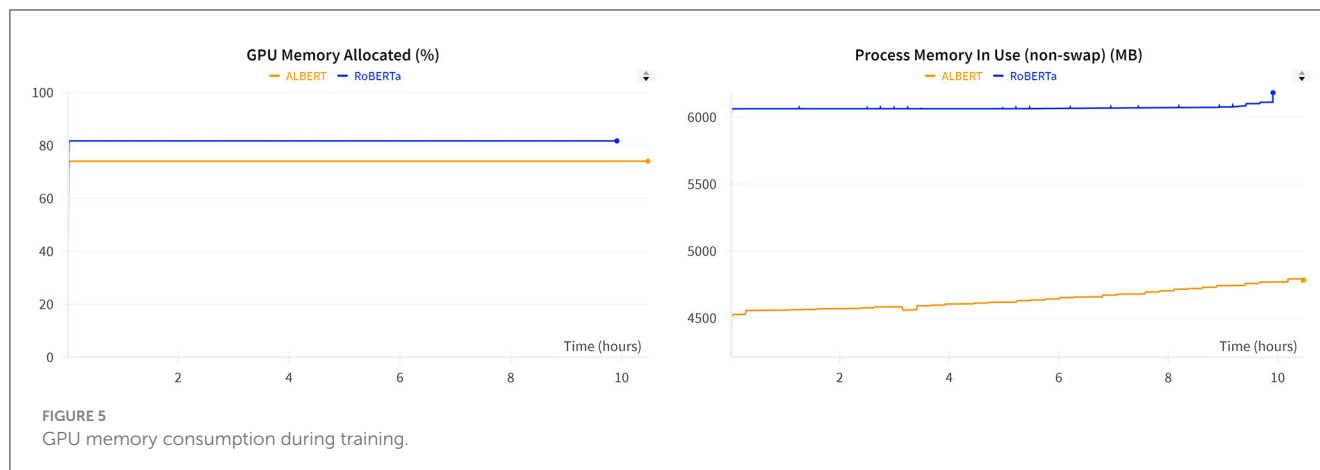
Table 7 presents a comparison of the overall performance of the models across 40 epochs. During training, ALBERT consumed 30 min more than RoBERTa because of the slower training steps in ALBERT. However, no significant difference was observed in the overall training loss between the two models. The training loss comparison shows a negligible variance of 0.11%, which is statistically insignificant. Additionally, the metrics in Figure 5 illustrate the GPU memory allocation and resource consumption by RoBERTa and ALBERT when trained

separately. ALBERT was found to have a 6% lower allocation of GPU memory, with an average difference of 1,500 MBs overall. Furthermore, Figure 6 illustrates that ALBERT results in a relatively lower emission of heat in GPUs, with differences ranging from 4°C at the beginning, consistent 1°C in the middle, to a significant 14% at the end of the training procedures. Moreover, ALBERT also exhibits relatively lower power consumption in Watts.

Taken together, it is interesting to note the striking similarity between the results produced by training both language models. Despite the difference in architecture and number of parameters, training on the same data has led to very close training results. In summary, this study indicates that the size of the language model does not have a discernible difference in learning and the consequent predictions of the model.

5 Discussion

This research uncovered two salient aspects of comparison drawn between large and small language models, owing to their parameters. Building upon the notion of textual analysis and its potential to predict the personality of individuals, our study focused on two primary aspects of automated text-based predictions. The initial objective of this study was to explore the feasibility of smaller language models in contrast with LLMs. The second point of focus was the use of multi-output regression to show human personality on a continuum. The resulting values ranged from 0 to 100, exhibiting values for all five traits of the Big Five model employing a large-scale personality dataset. Addressing the aforementioned objectives, the current investigation revealed that no noticeable difference could be observed by training language models on the same dataset. This finding is in line with previous data science research which implies that the quality of the training dataset is crucial in determining the performance of various AI-based models (Stuart Geiger et al., 2020; Liu et al., 2023). Specifically, it reinforces the findings by Mehta et al. (2020b), where they point out that deep learning-based personality prediction is also affected by data quality.



5.1 Implications of the study

The contextual alignment of personality detection and computational developments underscore the significance of our study, offering valuable insights building upon the evolving landscape of research in psychology, particularly within the context of APP. First, we shed light on personality predictions using automated methods. Moreover, our research design has been accentuated to incorporate the most recent technological breakthroughs in LLMs, especially transformers proposed by Vaswani et al. (2017). Although pre-trained LLMs have been a less researched methodology, they have become much more desirable areas in research. This desire stems from their sophisticated and resource-intensive computation, with a fraction of the effort and cost invested (Kumar and Renuka, 2023). As mentioned in the literature, transfer learning has revolutionized the realm of NLP (El-Demerdash et al., 2021; Rajapaksha et al., 2021; Yuan et al., 2023) and our empirical investigation reinforces the significance of transfer learning. We found that we can leverage small language models, emphasizing their learning on a specialized task of predicting personality from the textual data, as suggested by recent studies (Araci, 2019; Li et al., 2020; Yang et al., 2020; Kjell et al., 2023). In this study, the execution parameters were kept identical to rule out any other cause of similar results being produced. The findings of this study, presented in Tables 5, 6, show that there is

little to no difference in the performance of the training models of varying sizes. This is consistent with the arguments put forth by Hsieh et al. (2023) and Sanh et al. (2019), and supports the idea of smaller language models proposed by Schick and Schütze (2020), after the continuous up-scaling of the language models beyond resources.

Second, in addition to investigating large and small pre-trained LLMs, this study examined online textual data to predict personality, specifically the Big Five traits (Gjurković and Šnajder, 2018). Our approach aligns with the perception of human personality to be evaluated on a continuum instead of labeled classes (Johnson and Murty, 2023). This model investigation follows the work of López-Pabón and Orozco-Arroyave (2022) and Xue et al. (2018), who used regression techniques for APP. We adapted these studies by customizing the pre-trained language models RoBERTa and ALBERT to produce a multi-output regression. This study confirms regression as a viable statistical technique to predict the values of five personality traits, supporting the proposal of Mehta et al. (2020a).

This study offers various practical implications in diverse contexts. APP can be extremely beneficial for maintaining general wellbeing (Moreno et al., 2021) and detecting suicidal tendencies and mental health risks (Deilami et al., 2022). In addition, this concept is expected to be valuable in social network analysis and deception detection (Xue et al., 2018) and voter

inclination toward elections (Tutaysalgiir et al., 2019). These models can also be deployed to enrich the experience with autocars, robots, voice assistants, and other human-machine interaction agents (Kazameini et al., 2020). Additionally, personality-based psychometric analysis can massively contribute to improvements in crucial business performance indicators such as sales and social media clicks (Matz et al., 2017). Such psychological profiling can influence the behavior of people by personalizing business strategies according to their personalities.

Moreover, our study has utilized a Tesla P100-PCIE-16GB GPU to increase the pace of models' training. The resource consumption statistics of our method validate the feasibility of faster processing units, such as GPUs and TPUs in commercial use. Such advanced processing units together with big data, have enabled companies to adopt state-of-the-art computational methods seamlessly (Lecun et al., 2015). Search engines, recommendation systems, search rankings, fake news identification, and translations are just a few applications already employed in organizations (Pais et al., 2022). Corporations have shown an immensely growing tilt to apply these studies to their business processes.

Since employing models is a challenging task in finding solutions to NLP-related business problems (Paleyes et al., 2022), our research empirically investigated the feasibility of using pre-trained language models to predict the personality traits of individuals from their texts. Furthermore, our research suggests that smaller models can be effectively utilized in diverse business contexts. Additionally, the reduced usage of computational resources lowers the CO₂e emissions, thereby lowering potential climate impacts, hence addressing the concerns put forward in recent research (Henderson et al., 2020; Patterson et al., 2021). A lower heat emission confirms the decreased necessity for water to cool down the systems, thereby lowering the water footprint reinforcing the argument by Li et al. (2023). According to McDonald et al. (2022), model inference requires nearly 80% of the computational demand. Therefore, analyzing and comparing the computational resource consumption of different LLMs at the training stage is essential to minimize energy consumption and carbon footprint during inference. Our study not only validates the use of pre-trained language models to predict personality but also emphasizes the practicality of employing smaller language models in various organizational settings. Our findings support the previous literature which emphasizes prioritizing energy and computational efficiency when selecting models (Strubell et al., 2020; Tamburrini, 2022). This underscores the practical application of the proposed pre-trained smaller language models in contexts where human personality plays a crucial role. Overall, companies can leverage the benefits of such pre-trained models while minimizing their financial and technical computation budgets, aligning with sustainable business practices.

6 Limitations and future implications

We acknowledge the presence of certain limitations in our research. Our research offers initial evidence of the similar performances of a large and a ten times smaller model with other stable parameters. However, because of the unavailability of more powerful GPU resources, we could not include larger models.

Models such as Llama (Touvron et al., 2023), LaMDA (Thoppilan et al., 2022), and GPT-3 (Brown et al., 2020) may provide more insightful results as they are much larger. A comparison with such substantial models would offer a more robust perspective on the study. Second, future research could also use datasets with varying sizes in parameters; and quality such as with biased sampling; subjective labels; imbalanced classes; or limited diversity. This would enable a deeper comparison of the models in analyzing which one performs better if data quality is poor.

From the methodological approach, variations in error margin for different texts were observed. Although the model successfully produced the multioutput ratings of the big five traits, the predictions are expected to improve with longer training employing more efficient computational resources. Furthermore, we also propose to contrast the performance output of the auto-encoder and auto-regressive model architectures (Yang et al., 2019; Zhang et al., 2020). This comparison can indicate an architecture that is more suitable for a specialized task of personality prediction. Another architectural comparison can be of single-label classification, multi-label or multi-class classification as well as single-output regression and multi-output regression. Such multi-level analysis may provide insight into the customization criteria for pre-trained models for optimal performance, regardless of their parameter size. Since our study has major implications regarding the use of computational resources by LLMs and its environmental impact, future studies could extend this line of inquiry by employing techniques for carbon-footprint reduction such as power-capping or energy-aware scheduling (McDonald et al., 2022).

7 Conclusion

The present study integrates the theoretical underpinnings of the Big Five personality model with state-of-the-art technology. This integration is intended to assess the potential of employing pre-trained language models to predict human personality based on their language. This paper commences with a comparative account of ML and deep learning techniques used for similar objectives by previous researchers. Additionally, our paper highlights the advancements in pre-trained models since their emergence in NLP. Furthermore, our analytical outcomes establish a comparable performance yielded by the two models, RoBERTa and ALBERT, despite their different parameter sizes. Our results also provide logical evidence in support of multi-output regression. Moreover, we observe a reduced heat emission as well as lower carbon and water footprint by smaller models. Our novel findings are expected to stimulate more nuanced questions, to be raised in this direction, thereby broadening the scope of research and industrial applications alike.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The social media texts from the Reddit platform are kept anonymous. The research is conducted with utmost respect for individuals' privacy and commitment to transparent and ethical data usage practices throughout the study.

Author contributions

FH: Writing – review & editing, Writing – original draft, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. ZA: Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization. AA: Writing – review & editing, Writing – original draft, Supervision, Project administration. KK: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration. FP: Writing – review & editing, Writing – original draft.

References

- Alam, F., Stepanov, E. A., and Riccardi, G. (2013). "Personality traits recognition on social network - facebook," in *Proceedings of the International AAAI Conference on Web and Social Media* (Cambridge, MA: AI Access Foundation), 6–9. doi: 10.1609/icwsm.v7i2.14464
- Alexander, L., Mulfinger, E., and Oswald, F. L. (2020). Using big data and machine learning in personality measurement: opportunities and challenges. *Eur. J. Pers.* 34, 632–648. doi: 10.1002/per.2305
- Araci, D. T. (2019). FinBERT: financial sentiment analysis with pre-trained language models. *arXiv [Preprint]*. arXiv:1908.10063. doi: 10.48550/arXiv.1908.10063
- Aroganam, G., Manivannan, N., and Harrison, D. (2019). Review on wearable technology sensors used in consumer sport applications. *Sensors* 19:1983. doi: 10.3390/s19091983
- Bach, R. L., Kern, C., Amaya, A., Keusch, F., Kreuter, F., Hecht, J., et al. (2021). Predicting voting behavior using digital trace data. *Soc. Sci. Comput. Rev.* 39, 862–883. doi: 10.1177/0894439319882896
- Berggren, M., Kaati, L., Pelzer, B., Stiff, H., Lundmark, L., and Akrami, N. (2024). The generalizability of machine learning models of personality across two text domains. *Pers. Individ. Dif.* 217:112465. doi: 10.1016/j.paid.2023.112465
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., et al. (2020). Language models are few-shot learners. *Adv. Neural Inf. Process. Syst.* Available at: <https://arxiv.org/abs/2005.14165v4> (accessed July 31, 2023).
- Chen, C., Chen, X., Wang, L., Ma, X., Wang, Z., Liu, K., et al. (2017). MA-SSR: a memetic algorithm for skyline scenic routes planning leveraging heterogeneous user-generated digital footprints. *IEEE Trans. Veh. Technol.* 66, 5723–5736. doi: 10.1109/TVT.2016.2639550
- Christian, H., Suhartono, D., Chowanda, A., and Zamli, K. Z. (2021). Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging. *J. Big Data* 8, 1–20. doi: 10.1186/s40537-021-00459-1
- Chung, W., and Zeng, D. (2020). Dissecting emotion and user influence in social media communities: an interaction modeling approach. *Inf. Manag.* 57:103108. doi: 10.1016/j.im.2018.09.008
- Church, K. W., Chen, Z., and Ma, Y. (2021). Emerging trends: a gentle introduction to fine-tuning. *Nat. Lang. Eng.* 27, 763–778. doi: 10.1017/S1351324921000322
- Cui, B., and Qi, C. (2017). *Survey analysis of machine learning methods for natural language processing for MBTI personality type prediction*. Stanford University. Available at: <https://cs229.stanford.edu/proj2017/final-reports/5242471.pdf>
- Deilami, F. M., Sadr, H., and Tarkhan, M. (2022). Contextualized multidimensional personality recognition using combination of deep neural network and ensemble

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- learning. *Neural Process. Lett.* 54, 3811–3828. doi: 10.1007/s11063-022-10787-9
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2019). "BERT: pre-training of deep bidirectional transformers for language understanding," in *Proceedings of NAACL-HLT* (Minneapolis, MN), 4171–4186. doi: 10.48550/arXiv.1810.04805
- El-Demerdash, K., El-Khoribi, R. A., Shoman, M. A. I., and Abdou, S. (2021). Psychological human traits detection based on universal language modeling. *Egypt. Informatics J.* 22, 239–244. doi: 10.1016/j.eij.2020.09.001
- El-Demerdash, K., El-Khoribi, R. A., Shoman, M. A. I., and Abdou, S. (2022). Deep learning based fusion strategies for personality prediction. *Egypt. Informatics J.* 23, 47–53. doi: 10.1016/j.eij.2021.05.004
- Emami, S., and Martinez-Munoz, G. (2024). Deep learning for multi-output regression using gradient boosting. *IEEE Access* 12, 17760–17772. doi: 10.1109/ACCESS.2024.3359115
- Fan, J., Sun, T., Liu, J., Zhao, T., Zhang, B., Chen, Z., et al. (2023a). How well can an AI chatbot infer personality? Examining psychometric properties of machine-inferred personality scores. *J. Appl. Psychol.* 108, 1277–1299. doi: 10.1037/apl0001082
- Fan, L., Li, L., Ma, Z., Lee, S., Yu, H., and Hemphill, L. (2023b). A bibliometric review of large language models research from 2017 to 2023. *arXiv [Preprint]*. arXiv:2304.02020. Available at: <https://arxiv.org/abs/2304.02020v1> (Accessed July 31, 2023).
- Feizi-Derakhshi, A. R., Feizi-Derakhshi, M. R., Ramezani, M., Nikzad-Khaskhaki, N., Asgari-Chenaghlu, M., Akan, T., et al. (2022). Text-based automatic personality prediction: a bibliographic review. *J. Comput. Soc. Sci.* 5, 1555–1593. doi: 10.1007/s42001-022-00178-4
- Fu, Y., Peng, H., Ou, L., Sabharwal, A., and Khot, T. (2023). "Specializing smaller language models towards multi-step reasoning" in *Proceedings of the 40th International Conference on Machine Learning*, 10421–10430. arXiv:1908.10063. Available at: <https://proceedings.mlr.press/v202/fu23d.html> (accessed October 31, 2023).
- Funder, D. C. (2012). Accurate personality judgment. *Curr. Dir. Psychol. Sci.* 21, 177–182. doi: 10.1177/0963721412445309
- Garimella, A., Research, A., Amarnath, A., and Mihalcea, R. (2022). "Demographic-aware language model fine-tuning as a bias mitigation technique," in *Association for Computational Linguistics (Short Papers)*, 311–319. Available at: <https://aclanthology.org/2022.acl-short.38> (accessed April 22, 2024).
- Gjurkovic, M., Karan, M., Vukojevic, I., Bosnjak, M., and Snajder, J. (2021). "PANDORA talks : personality and demographics on Reddit," in *Soc. 2021 - 9th Int. Work. Nat. Lang. Process. Soc. Media, Proc. Work* (New Orleans, LA), 138–152. doi: 10.18653/v1/2021.socialnlp-1.12

- Gjurković, M., and Šnajder, J. (2018). "Reddit: a gold mine for personality prediction." in *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media* (New Orleans, LA: Association for Computational Linguistics), 87–97. doi: 10.18653/v1/W18-1112
- Gladstone, J. J., Matz, S. C., and Lemaire, A. (2019). Can psychological traits be inferred from spending? Evidence from transaction data. *Psychol. Sci.* 30, 1087–1096. doi: 10.1177/0956797619849435
- Goldberg, L. R. (1982). From ace to zombie: some explorations in the language of personality. *Adv. Personal. Assess.* 1, 203–234.
- Golder, S. A., and Macy, M. W. (2014). Digital footprints: opportunities and challenges for online social research. *Annu. Rev. Sociol.* 40, 129–152. doi: 10.1146/annurev-soc-071913-043145
- Habib, F. (2024). "Automated Personality Prediction", *Mendeley Data*, V1. doi: 10.17632/3sndbd4p84.1v
- Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., and Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: opportunities, practical considerations, and challenges. *Perspect. Psychol. Sci.* 11, 838–854. doi: 10.1177/1745691616650285
- Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., and Pineau, J. (2020). Towards the systematic reporting of the energy and carbon footprints of machine learning. *J. Mach. Learn. Res.* 21, 1–43. doi: 10.48550/arXiv.2002.05651
- Hinds, J., and Joinson, A. (2019). Human and computer personality prediction from digital footprints. *Curr. Dir. Psychol. Sci.* 28, 204–211. doi: 10.1177/0963721419827849
- Hinds, J., and Joinson, A. N. (2018). What demographic attributes do our digital footprints reveal? A systematic review. *PLoS ONE* 13:e0207112. doi: 10.1371/journal.pone.0207112
- Hirsh, J. B., Kang, S. K., and Bodenhausen, G. V. (2012). Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. *Psychol. Sci.* 23, 578–581. doi: 10.1177/0956797611436349
- Hsieh, C.-Y., Li, C.-L., Yeh, C.-K., Nakhost, H., Fujii, Y., Ratner, A., et al. (2023). "Distilling step-by-step! Outperforming larger language models with less training data and smaller model sizes," in *Findings of the Association for Computational Linguistics: ACL 2023* (Toronto, ON: Association for Computational Linguistics), 8003–8017. doi: 10.18653/v1/2023.findings-acl.507
- Ihsan, Z., and Furnham, A. (2018). The new technologies in personality assessment: a review. *Consult. Psychol. J.* 70, 147–166. doi: 10.1037/cpb0000106
- Jiang, H., Zhang, X., and Choi, J. D. (2020). "Automatic text-based personality recognition on monologues and multiparty dialogues using attentive networks and contextual embeddings," in *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence* (New York, NY), 13821–13822. doi: 10.1609/aaai.v34i10.7182
- Johnson, S. J., and Murty, M. R. (2023). An aspect-aware enhanced psycholinguistic knowledge graph-based personality detection using deep learning. *SN Comput. Sci.* 4:293. doi: 10.1007/s42979-023-01670-y
- Jukić, J., Vukojević, I., and Šnajder, J. (2022). "You are what you talk about: inducing evaluative topics for personality analysis," in *Findings of the Association for Computational Linguistics: EMNLP 2022* (Abu Dhabi: ACL), 3986–3999. doi: 10.18653/v1/2022.findings-emnlp.294
- Karanatsiou, D., Sermpezis, P., Gruda, D., Kafetsios, K., Dimitriadis, I., and Vakali, A. (2022). "My Tweets bring all the traits to the yard: predicting personality and relational traits in online social networks," in *ACM Transactions on the Web* (New York, NY: Association for Computing Machinery). doi: 10.1145/3523749
- Kazameini, A., Fatehi, S., Mehta, Y., Eetemadi, S., and Cambria, E. (2020). Personality trait detection using bagged SVM over BERT word embedding ensembles. *arXiv [Preprint]*. arXiv:2010.01309. doi: 10.48550/arXiv.2010.01309
- Keh, S. S., and Cheng, I. (2019). Myers-Briggs personality classification and personality-specific language generation using pre-trained language models. *arXiv [Preprint]*. arXiv:1907.06333. doi: 10.48550/arXiv.1907.06333
- Khorrami, M., Khorrami, M., and Farhangi, F. (2022). Evaluation of tree-based ensemble algorithms for predicting the big five personality traits based on social media photos: evidence from an Iranian sample. *Pers. Individ. Dif.* 188:111479. doi: 10.1016/j.paid.2021.111479
- Kim, J., Lee, B., Bearman, P., Baldassarri, D., Bach, J., Bonikowski, B., et al. (2023). *AI-Augmented Surveys: Leveraging Large Language Models and Surveys for Opinion Prediction*. Available at: <https://arxiv.org/abs/2305.09620v3> (accessed April 22, 2024).
- Kjell, O., Kjell, K., and Schwartz, H. A. (2023). Beyond rating scales: with care for validation large language models are poised to change psychological assessment. *arXiv [Preprint]*. doi: 10.31234/osf.io/yfd8g
- Kleč, M., Wiczorkowska, A., Szklanny, K., and Strus, W. (2023). Beyond the Big Five personality traits for music recommendation systems. *Eurasip J. Audio Speech Music Process.* 2023, 1–17. doi: 10.1186/s13636-022-00269-0
- Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., and Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. *Mach. Learn.* 95, 357–380. doi: 10.1007/s10994-013-5415-y
- Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proc. Natl. Acad. Sci. USA.* 110, 5802–5805. doi: 10.1073/pnas.1218772110
- Kulkarni, V., Kern, M. L., Stillwell, D., Kosinski, M., Matz, S., Ungar, L., et al. (2018). Latent human traits in the language of social media: an open-vocabulary approach. *PLoS ONE* 13:e0201703. doi: 10.1371/journal.pone.0201703
- Kumar, L. A., and Renuka, D. K. (2023). "Transformers: State-of-the-Art natural language processing," in *Deep Learning Approach for Natural Language Processing, Speech, and Computer Vision* (Boca Raton, FL), 49–75. doi: 10.1201/9781003348689-3
- Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., and Soricut, R. (2019). "ALBERT: a lite BERT for self-supervised learning of language Representations," in *8th International Conference on Learning Representations, ICLR 2020* (International Conference on Learning Representations, ICLR). Available at: <https://arxiv.org/abs/1909.11942v6> (accessed November 13, 2023).
- Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539
- Lee, M. H., Cha, S., and Nam, T. J. (2015). Impact of digital traces on the appreciation of movie contents. *Digit. Creat.* 26, 287–303. doi: 10.1080/14626268.2015.1087410
- Letzring, T. D., and Human, L. J. (2014). An examination of information quality as a moderator of accurate personality judgment. *J. Pers.* 82, 440–451. doi: 10.1111/jopy.12075
- Lewenberg, Y., Bachrach, Y., and Volkova, S. (2015). "Using emotions to predict user interest areas in online social networks," in *Proc. 2015 IEEE Int. Conf. Data Sci. Adv. Anal. DSAA 2015* (Paris: IEEE). doi: 10.1109/DSAA.2015.7344887
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Adv. Neural Inf. Process. Syst.* 33, 9459–9474. doi: 10.48550/arXiv.2005.11401
- Li, P., Yang, J., Islam, M. A., and Ren, S. (2023). Making AI less "thirsty": uncovering and addressing the secret water footprint of AI models. *arXiv [Preprint]*. Available at: <https://arxiv.org/abs/2304.03271v3> (accessed January 18, 2024).
- Li, Y., Kazameini, A., Mehta, Y., and Cambria, E. (2021). Multitask learning for emotion and personality detection. *arXiv [Preprint]*. Available at: <https://arxiv.org/abs/2101.02346v1> (accessed January 29, 2024).
- Li, Y., Rao, S., Solares, J. R. A., Hassaine, A., Ramakrishnan, R., Canoy, D., et al. (2020). BEHRT: transformer for electronic health records. *Sci. Rep.* 10, 1–12. doi: 10.1038/s41598-020-62922-y
- Lin, Y. H., Wong, B. Y., Pan, Y. C., Chiu, Y. C., and Lee, Y. H. (2019). Validation of the mobile app-recorded circadian rhythm by a digital footprint. *JMIR mHealth uHealth* 7:e13421. doi: 10.2196/13421
- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., et al. (2023). Summary of ChatGPT-related research and perspective towards the future of large language models. *Meta-Radiology* 1:100017. doi: 10.1016/j.metrad.2023.100017
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., et al. (2019). RoBERTa: a robustly optimized BERT pretraining approach. *arXiv [Preprint]*. arXiv:1907.11692. doi: 10.48550/arXiv.1907.11692
- López-Pabón, F. O., and Orozco-Arroyave, J. R. (2022). Automatic personality evaluation from transliterations of YouTube vlogs using classical and state-of-the-art word embeddings. *Ing. Investig.* 42:e93803. doi: 10.15446/ing.investig.93803
- Majumder, N., Poria, S., Gelbukh, A., and Cambria, E. (2017). Deep learning-based document modeling for personality detection from text. *IEEE Intell. Syst.* 32, 74–79. doi: 10.1109/MIS.2017.23
- Markovikj, D., Gievska, S., Kosinski, M., and Stillwell, D. (2013). "Mining facebook data for predictive personality modeling," in *Proceedings of the International AAAI Conference on Web and Social Media* (Cambridge, MA), 23–26. doi: 10.1609/icwsm.v7i2.14466
- Matz, S., Kosinski, M., Nave, G., and Stillwell, D. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proc. Natl. Acad. Sci. USA.* 12714–12719. doi: 10.1073/pnas.1710966114
- Matz, S., Teeny, J., Vaid, S. S., Harari, G. M., and Cerf, M. (2023). The potential of generative AI for personalized persuasion at scale. *PsyArXiv*. doi: 10.31234/osf.io/rn97c
- McCrae, R. R., and John, O. P. (1992). An introduction to the five-factor model and its applications. *J. Pers.* 60, 175–215. doi: 10.1111/j.1467-6494.1992.tb00970.x
- McDonald, J., Li, B., Frey, N., Tiwari, D., Gadepally, V., and Samsi, S. (2022). "Great power, great responsibility: recommendations for reducing energy for training language models," in *Find. Assoc. Comput. Linguist. NAACL 2022* (Seattle, WA: ACL), 1962–1970. doi: 10.18653/v1/2022.findings-naacl.151
- Mehta, Y., Fatehi, S., Kazameini, A., Stachl, C., Cambria, E., and Eetemadi, S. (2020a). "Bottom-up and top-down: predicting personality with psycholinguistic and language model features," in *Proceedings - IEEE International Conference on Data Mining, ICDM* (Sorrento: IEEE), 1184–1189. doi: 10.1109/ICDM50108.2020.00146
- Mehta, Y., Gelbukh, A., Cambria, E., and Majumder, N. (2020b). Recent trends in deep learning based personality detection. *Artif. Intell. Rev.* 53, 2313–2339. doi: 10.1007/s10462-019-09770-z

- Mohammad, S. M., and Kiritchenko, S. (2013). "Using nuances of emotion to identify personality," in *Proceedings of the International AAAI Conference on Web and Social Media* (Cambridge, MA), 27–30. doi: 10.1609/icwsm.v7i2.14468
- Mønsted, B., Mollgaard, A., and Mathiesen, J. (2018). Phone-based metric as a predictor for basic personality traits. *J. Res. Pers.* 74, 16–22. doi: 10.1016/j.jrp.2017.12.004
- Moreno, J. D., Martínez-Huertas, J., Olmos, R., Jorge-Botana, G., and Botella, J. (2021). Can personality traits be measured analyzing written language? A meta-analytic study on computational methods. *Pers. Individ. Dif.* 177:110818. doi: 10.1016/j.paid.2021.110818
- Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D., and Rentfrow, J. (2018). Musical preferences predict personality: evidence from active listening and facebook likes. *Psychol. Sci.* 29, 1145–1158. doi: 10.1177/0956797618761659
- Ortigosa, A., Carro, R. M., and Quiroga, J. I. (2014). Predicting user personality by mining social interactions in facebook. *J. Comput. Syst. Sci.* 80, 57–71. doi: 10.1016/j.jcss.2013.03.008
- Pais, S., Cordeiro, J., and Jamil, M. L. (2022). NLP-based platform as a service: a brief review. *J. Big Data* 9, 1–26. doi: 10.1186/s40537-022-00603-5
- Paley, A., Urma, R. G., and Lawrence, N. D. (2022). Challenges in deploying machine learning: a survey of case studies. *ACM Comput. Surv.* 55, 1–29. doi: 10.1145/3533378
- Patel, S. C., and Fan, J. (2023). Identification and description of emotions by current large language models. *bioRxiv*. doi: 10.1101/2023.07.17.549421
- Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., et al. (2021). *Carbon Emissions and Large Neural Network Training*. Available at: <https://arxiv.org/abs/2104.10350v3> (accessed February 13, 2024).
- Peng, K. H., Liou, L. H., Chang, C. S., and Lee, D. S. (2015). "Predicting personality traits of Chinese users based on Facebook wall posts," in *2015 24th Wireless and Optical Communication Conference, WOCC (Taipei: IEEE)*, 9–14. doi: 10.1109/WOCC.2015.7346106
- Peng, L., Zhang, Z., Pang, T., Han, J., Zhao, H., Chen, H., et al. (2024). "Customising general large language models for specialised emotion recognition tasks," in *ICASSP 2024 - 2024 IEEE Int. Conf. Acoust. Speech Signal Process.* (Seoul: IEEE), 11326–11330. doi: 10.1109/ICASSP48485.2024.10447044
- Peters, H., and Matz, S. (2023). *Large Language Models Can Infer Psychological Dispositions of Social Media Users*. Available at: <https://arxiv.org/abs/2309.08631v1> (accessed October 29, 2023).
- Phan, L. V., and Rauthmann, J. F. (2021). Personality computing: new frontiers in personality assessment. *Soc. Personal. Psychol. Compass* 15:e12624. doi: 10.1111/spc3.12624
- Plummer, B. A., Dryden, N., Zürich, E., Frost, J., Hoefler, T., and Saenko, K. (2020). Shapshifter networks: cross-layer parameter sharing for scalable and effective deep learning. *arXiv [Preprint]*. Available at: <http://arxiv.org/abs/2006.10598>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol.* 88, 879–903. doi: 10.1037/0021-9010.88.5.879
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychol. Bull.* 135, 322–338. doi: 10.1037/a0014996
- Preis, T., Moat, H. S., Bishop, S. R., Treleaven, P., and Stanley, H. E. (2013). Quantifying the digital traces of hurricane sandy on flickr. *Sci. Rep.* 3, 1–3. doi: 10.1038/srep03141
- Putra, R. P., and Setiawan, E. B. (2022). "RoBERTa as semantic approach for Big Five personality prediction using artificial neural network on Twitter," in *2022 International Conference on Advanced Creative Networks and Intelligent Systems (ICACNIS)* (Bandung: IEEE), 1–6. doi: 10.1109/ICACNIS57039.2022.10054870
- Qiao, Y., Xiong, C., Liu, Z., and Liu, Z. (2019). Understanding the behaviors of BERT in ranking. *arXiv [Preprint]*. arXiv:1904.07531. doi: 10.48550/arXiv:1904.07531
- Quercia, D., Kosinski, M., Stillwell, D., and Crowcroft, J. (2011). "Our twitter profiles, our selves: predicting personality with twitter," in *Proceedings - 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom 2011* (Boston, MA: IEEE), 180–185. doi: 10.1109/PASSAT/SocialCom.2011.26
- Quwaider, M., Alabed, A., and Duwairi, R. (2023). Shooter video games for personality prediction using five factor model traits and machine learning. *Simul. Model. Pract. Theory* 122:102665. doi: 10.1016/j.simpat.2022.102665
- Radisavljević, D., Rzepka, R., and Araki, K. (2023). Personality types and traits—examining and leveraging the relationship between different personality models for mutual prediction. *Appl. Sci.* 13:4506. doi: 10.3390/app13074506
- Rafaeli, A., Ashtar, S., and Altman, D. (2019). Digital traces: new data, resources, and tools for psychological-science research. *Curr. Dir. Psychol. Sci.* 28, 560–566. doi: 10.1177/0963721419861410
- Rajanak, Y., Patil, R., and Singh, Y. P. (2023). "Language detection using natural language processing" in *2023 9th International Conference on Advanced Computing and Communication Systems, ICACCS 2023* (Coimbatore: IEEE), 673–678. doi: 10.1109/ICACCS57279.2023.10112773
- Rajapaksha, P., Farahbakhsh, R., and Crespi, N. (2021). BERT, XLNet or RoBERTa: the best transfer learning model to detect clickbaits. *IEEE Access* 9, 154704–154716. doi: 10.1109/ACCESS.2021.3128742
- Ramezani, M., Feizi-Derakhshi, M. R., and Balafar, M. A. (2022). Text-based automatic personality prediction using KGraT-Net: a knowledge graph attention network classifier. *Sci. Rep.* 12:21453. doi: 10.1038/s41598-022-25955-z
- Rentfrow, P. J., and Gosling, S. D. (2003). The do re mi's of everyday life: the structure and personality correlates of music preferences. *J. Pers. Soc. Psychol.* 84, 1236–1256. doi: 10.1037/0022-3514.84.6.1236
- Rillig, M. C., Ågerstrand, M., Bi, M., Gould, K. A., and Sauerland, U. (2023). Risks and benefits of large language models for the environment. *Environ. Sci. Technol.* 57, 3464–3466. doi: 10.1021/acs.est.3c01106
- Rothman, D. (2021). *Transformers for Natural Language Processing Build innovative deep neural network architectures for NLP & Python, PyTorch, TensorFlow, BERT, RoBERTa & more*. Available at: <https://www.packtpub.com/product/transformers-for-natural-language-processing/9781800565791> (accessed April 26, 2024).
- Sanh, V., Debut, L., Chaumond, J., and Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv [Preprint]*. arXiv:1910.01108. doi: 10.48550/arxiv.1910.01108
- Schick, T., and Schütze, H. (2020). "It's not just size that matters: small language models are also few-shot learners," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (ACL)*, 2339–2352. doi: 10.18653/v1/2021.naacl-main.185
- Segalin, C., Perina, A., Cristani, M., and Vinciarelli, A. (2017). The pictures we like are our image: continuous mapping of favorite pictures into self-assessed and attributed personality traits. *IEEE Trans. Affect. Comput.* 8, 268–285. doi: 10.1109/TAFFC.2016.2516994
- Sicilia, A., Gates, J. C., and Alikhani, M. (2024). HumBEL: a human-in-the-loop approach for evaluating demographic factors of language models in human-machine conversations. *Assoc. Comput. Linguist.* 1, 1127–1143. doi: 10.48550/arXiv.2305.14195
- Strubell, E., Ganesh, A., and McCallum, A. (2020). Energy and policy considerations for modern deep learning research. *Proc. AAAI Conf. Artif. Intell.* 34, 13693–13696. doi: 10.1609/aaai.v34i09.7123
- Stuart Geiger, R., Yu, K., Yang, Y., Dai, M., Qiu, J., Tang, R., et al. (2020). "Garbage in, garbage out? Do machine learning application papers in social computing report where human-labeled training data comes from?" in *FAT* 2020 - Proc. 2020 Conf. Fairness, Accountability, Transpar.* (New York, NY: ACM), 325–336. doi: 10.1145/3351095.3372862
- Sun, X., Liu, B., Cao, J., Luo, J., and Shen, X. (2018). "Who am I? Personality detection based on deep learning for texts," in *IEEE International Conference on Communications* (Kansas City, MO: IEEE), 1–6. doi: 10.1109/ICC.2018.8422105
- Sutton, A., Allinson, C., and Williams, H. (2013). Personality type and work-related outcomes: an exploratory application of the Enneagram model. *Eur. Manag. J.* 31, 234–249. doi: 10.1016/j.emj.2012.12.004
- Tadesse, M. M., Lin, H., Xu, B., and Yang, L. (2018). Personality predictions based on user behavior on the Facebook social media platform. *IEEE Access* 6, 61959–61969. doi: 10.1109/ACCESS.2018.2876502
- Tamburrini, G. (2022). The AI carbon footprint and responsibilities of AI scientists. *Philosophies* 7:4. doi: 10.3390/philosophies7010004
- Tandera, T., Hendro, Suhartono, D., Wongso, R., and Prasetyo, Y. L. (2017). Personality prediction system from Facebook users. *Procedia Comput. Sci.* 116, 604–611. doi: 10.1016/j.procs.2017.10.016
- Tang, Z., Chen, L., and L. Gillenson, M. (2018). How to keep brand fan page followers? The lens of person-environment fit theory. *Inf. Technol. People* 31, 927–947. doi: 10.1108/ITP-04-2016-0076
- Tausczik, Y., and Pennebaker, J. (2010). The psychological meaning of words: liwc and computerized text analysis methods. *J. Lang. Soc. Psychol.* 29, 24–543. doi: 10.1177/0261927X09351676
- Theil, K., Hovy, D., and Stuckenschmidt, H. (2023). Top-down influence? Predicting CEO personality and risk impact from speech transcripts. *Proc. Int. AAAI Conf. Web Soc. Media* 17, 832–841. doi: 10.1609/icwsm.v17i1.22192
- Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., et al. (2022). LaMDA: language models for dialog applications. *arXiv [Preprint]*. Available at: <https://arxiv.org/abs/2201.08239v3> (Accessed December 4, 2023).
- Tomat, L., Trkman, P., and Manfreda, A. (2021). Personality in information systems professions: identifying archetypal professions with suitable traits and candidates' ability to fake-good these traits. *Inf. Technol. People* 35, 52–73. doi: 10.1108/ITP-03-2021-0212
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., et al. (2023). LLaMA: open and efficient foundation language models. *arXiv [Preprint]*. Available at: <https://arxiv.org/abs/2302.13971v1> (accessed December 4, 2023).
- Trainer API (2023). Available at: https://huggingface.co/docs/transformers/main_classes/trainer (accessed August 2, 2023).

- Tutaysalgir, E., Karagoz, P., and Toroslu, I. H. (2019). "Clustering based personality prediction on Turkish tweets," in *Proc. 2019 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2019* (New York, NY: ACM), 825–828. doi: 10.1145/3341161.3343513
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. Neural Inf. Process. Syst.* 30, 5999–6009. doi: 10.18653/arXiv.1706.03762
- Volkova, S., and Bachrach, Y. (2016). "Inferring perceived demographics from user emotional tone and user-environment emotional contrast," in *54th Annu. Meet. Assoc. Comput. Linguist. ACL 2016 - Long Pap.* 3 (Berlin: ACL), 1567–1578. doi: 10.18653/v1/P16-1148
- Volkova, S., Bachrach, Y., and Van Durme, B. (2016). "Mining user interests to predict perceived psycho-demographic traits on twitter," in *Proc. - 2016 IEEE 2nd Int. Conf. Big Data Comput. Serv. Appl. BigDataService 2016* (Oxford: IEEE), 36–43. doi: 10.1109/BigDataService.2016.28
- Volodina, E., Mohammed, Y. A., Matsson, A., Derbring, S., and Megyesi, B. (2020). "Towards privacy by design in learner corpora research: a case of on-the-fly pseudonymization of Swedish learner essays," in *Proceedings of the 28th International Conference on Computational Linguistics* (Barcelona: ACL), 357–369. doi: 10.18653/v1/2020.coling-main.32
- Vrbancič, G., and Podgorelec, V. (2020). Transfer learning with adaptive fine-tuning. *IEEE Access* 8, 196197–196211. doi: 10.1109/ACCESS.2020.3034343
- Wang, M., and Hu, F. (2021). The application of nltk library for python natural language processing in corpus research. *Theory Pract. Lang. Stud.* 11, 1041–1049. doi: 10.17507/tpls.1109.09
- Wynekoop, J. L., and Walz, D. B. (2000). Investigating traits of top performing software developers. *Inf. Technol. People* 13, 186–195. doi: 10.1108/09593840010377626
- Xue, D., Wu, L., Hong, Z., Guo, S., Gao, L., Wu, Z., et al. (2018). Deep learning-based personality recognition from text posts of online social networks. *Appl. Intell.* 48, 4232–4246. doi: 10.1007/s10489-018-1212-4
- Yang, F., Quan, X., Yang, Y., and Yu, J. (2021). Multi-document transformer for personality detection. *Proc. AAAI Conf. Artif. Intell.* 35, 14221–14229. doi: 10.1609/aaai.v35i16.17673
- Yang, H. C., and Huang, Z. R. (2019). Mining personality traits from social messages for game recommender systems. *Knowledge-Based Syst.* 165, 157–168. doi: 10.1016/j.knsys.2018.11.025
- Yang, Y., Christopher, M., Uy, S., and Huang, A. (2020). FinBERT: a pretrained language model for financial communications. *arXiv [Preprint]*. Available at: <https://arxiv.org/abs/2006.08097v2> (accessed November 1, 2023).
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., and Le, Q. V. (2019). XLNet: generalized autoregressive pretraining for language understanding. *Adv. Neural Inf. Process. Syst.* 165:32.
- Yarkoni, T. (2010). Personality in 100,000 words: a large-scale analysis of personality and word use among bloggers. *J. Res. Pers.* 44, 363. doi: 10.1016/j.jrp.2010.04.001
- Yi, J., Du, Y., Liang, F., Tu, W., Qi, W., and Ge, Y. (2020). Mapping human's digital footprints on the Tibetan Plateau from multi-source geospatial big data. *Sci. Total Environ.* 711:134540. doi: 10.1016/j.scitotenv.2019.134540
- Youyou, W., Kosinski, M., and Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proc. Natl. Acad. Sci. USA.* 112, 1036–1040. doi: 10.1073/pnas.1418680112
- Yu, J., and Markov, K. (2017). Deep learning based personality recognition from Facebook status updates. in *2017 IEEE 8th international conference on awareness science and technology (iCAST)* (Taichung: IEEE), 383–387. doi: 10.1109/ICAWST.2017.8256484
- Yuan, L., Wang, T., Ferraro, G., Suominen, H., and Rizoio, M. A. (2023). Transfer learning for hate speech detection in social media. *J. Comput. Soc. Sci.* 6, 1081–1101. doi: 10.1007/s42001-023-00224-9
- Yun, J. T., Pamuksuz, U., and Duff, B. R. L. (2019). Are we who we follow? Computationally analyzing human personality and brand following on Twitter. *Int. J. Advert.* 38, 776–795. doi: 10.1080/02650487.2019.1575106
- Zhang, G., Liu, Y., and Jin, X. (2020). A survey of autoencoder-based recommender systems. *Front. Comput. Sci.* 14, 430–450. doi: 10.1007/s11704-018-8052-6
- Zhang, H., Chen, Y., Wang, M., and Feng, S. (2024). FEEL: a framework for evaluating emotional support capability with large language models. *arXiv [Preprint]*. Available at: <https://arxiv.org/abs/2403.15699v1> (accessed April 22, 2024).
- Zhao, X., and Wong, C. W. (2023). Automated measures of sentiment via transformer- and lexicon-based sentiment analysis (TLSA). *J. Comput. Soc. Sci.* 7, 1–26. doi: 10.36227/techrxiv.21781109.v2
- Zhong, X. F., Guo, S. Z., Gao, L., Shan, H., and Xue, D. (2018). "A general personality prediction framework based on Facebook profiles," in *Proceedings of the 2018 10th International Conference on Machine Learning and Computing* (New York, NY: Association for Computing Machinery), 269–275. doi: 10.1145/3195106.3195124
- Zimmermann, J., Woods, W. C., Ritter, S., Happel, M., Masuhr, O., Jaeger, U., et al. (2019). Integrating structure and dynamics in personality assessment: first steps toward the development and validation of a personality dynamics diary. *Psychol. Assess.* 31, 516–531. doi: 10.1037/pas0000625