Detecting Personality Traits from Texts using an Hierarchy of Tree-Transformers and Graph Attention Network with Word Embedding Refinement

Anonymous EACL submission

Abstract

Automatic detection of personality traits from individuals' written texts aids professionals in evaluating mental health and individuals in identifying their strengths and weaknesses, facilitating informed decisions on personal growth, workplace compatibility, and lifestyle choices. Psychologists have discerned a collection of personality traits that can manifest within an individual's character. While BERT-based models have been successful in categorizing writings into specific personality traits, they require significant time and resources for fine-tuning. This research introduces a novel approach that utilizes a hierarchical structure of tree-transformers and a graph attention network (GAT) to classify personality traits derived from written text. It also employs an heterogeneous GAT (H-GAT) to refine Roberta word embeddings. The proposed model demonstrates substantial performance enhancements compared to previous works, as evidenced by superior results on benchmark datasets.

1 Introduction

Personality refers to the enduring traits and patterns of behavior that an individual consistently displays. It encompasses a person's moods, attitudes, and opinions, which are explicitly manifested in their interactions with others. Personality encompasses a wide range of behavioral characteristics, both innate and acquired, that are observable in an individual's social relationships and their interactions with the surrounding environment (Ramezani et al., 2022b). The Big-Five personality traits, also known as OCEAN, are the widely accepted and commonly used model of personality (John et al., 2008). OCEAN represents personality through five dimensions: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (or alternatively, emotional stability) (Kazemeini et al., 2021). Another frequently employed personality model is the Myers-Briggs Type Indicator (MBTI) (Myers et al., 2000), which categorizes individuals into 16 distinct personality types based on four binary categories: Extroversion or Introversion, Sensing or Intuition, Thinking or Feeling, and Judging or Perceiving. These traits significantly influence an individual's future prospects and life outcomes (Roberts et al., 2007).

The utilization of artificial intelligence (AI) has become increasingly significant in assisting psychiatrists and healthcare professionals in addressing the escalating occurrence of mental health issues and disorders (Kazemeini et al., 2021).

According to a 2020 Harris Poll ((Samet, 2020), social media usage has significantly risen among adults in the United States, with approximately 50% reporting increased usage during the pandemic (Kim et al., 2021). The increased use of social media during the pandemic has led to the creation of extensive digital footprints. Researchers, such as Kosinski et al. (2013), have shown that these footprints can provide insights into an individual's personality and emotional traits.

Previous scholarly investigations have delved into the intricate interplay linking personality traits and mental health disorders. Empirical studies have substantiated the pivotal role of neuroticism in the genesis of depression and anxiety disorders (Kendler et al., 1993; Goldberg and Huxley, 1992); have revealed an inverse correlation between resilience and neuroticism, while establishing positive associations with conscientiousness and extraversion; a statistically significant positive correlation has been observed between openness and resilience (Campbell-Sills et al., 2006). Consequently, the automated comprehension of an individual's personality holds considerable potential in enhancing the therapeutic process for mental health concerns, thereby augmenting treatment outcomes and alleviating the strain on mental health services.

The majority of prevailing and contemporary cutting-edge models (Kazameini et al., 2020a;

Mehta et al., 2019) for classifying personality traits predominantly revolves around the utilization of BERT-based architectures. Performance in task-specific contexts can be enhanced through fine-tuning, however, it is worth noting that this entails a substantial demand of computational time and resources. Additionally, certain BERT-based models encounter challenges when handling lengthier texts, as they are constrained by the token intake limitations inherent to the BERT-based architectures.

In this study, we present a novel approach aimed at addressing the aforementioned challenges. We propose three distinct models that incorporate tree-transformers (Ahmed et al., 2019b), a graph attention network (GAT) (Veličković et al., 2017), and an heterogeneous graph attention network (H-GAT) (Wang et al., 2020). Each of these architectures employs an hierarchical structure comprising tree-transformers and GAT layers. The tree-transformers serve as sentence encoders, while the subsequent GAT layer encodes complete statements using the derived sentence vectors. To update the leaf nodes of the tree-transformers and sentence nodes, an H-GAT has been deployed, which leverages the statement embedding. Notably, the three models vary in the specific application of the H-GAT. By fine-tuning the word embeddings, these models effectively serve the purpose of BERT fine-tuning. Their advantage is reducing the need for substantial computational resources to fine-tune the millions of parameters found in the BERT-based models and enabling essentially unlimited input text lengths. In our study, we have conducted an extensive analysis of the performance of the proposed models on well-established personality trait identification datasets. Through rigorous analysis, the findings unequivocally show the superior performance of our proposed model when compared to previously prominent models in the field. We will upload the materials once the paper gets published.

2 Related Work

In recent years, there have been notable contributions in employing various deep learning models for the identification of personality traits. Kalghatgi et al. (2015) employed neural networks, specifically multilayer perceptrons (MLP), along with hand-crafted features to detect personality traits. Similarly, Su et al. (2016) utilized recurrent neural networks (RNN) along with hidden Markov models (HMM) to identify personality traits using Chinese Language Inquiry and Word Count (LIWC) annotations extracted from dialogues. Sun et al. (2018) and Tandera et al. (2017) utilized long-short-term-memory (LSTM) and convolutional neural networks (CNN) to detect personality traits from text data sourced from Facebook posts. Van de Ven et al. (2017) have conducted experiments on 275 LinkedIn profiles and provided evidence that extroversion can be accurately inferred from self-descriptions in user profiles. Lynn et al. (2020) employed message-level attention over Facebook posts to analyze users' personality traits. Gjurković et al. (2020) have introduced their self-created corpus in the context of personality analysis and applied S-BERT (Reimers and Gurevych, 2019) over it. Kazameini et al. (2020a) utilized an ensemble of SVMs with BERT embeddings and achieved superior performance compared to other models for the Big Five trait classification using the Essays corpus (Pennebaker and King, 1999). Mehta et al. (2020) have conducted experiments with various combinations of psycholinguistic features and BERT-based models, analyzing the impact of each feature on trait prediction. Stachl et al. (2021) and Mehta et al. (2019) delve into computational perspectives in their review article, exploring various aspects and considerations within the field of personality trait identification. Ramezani et al. (2022a) have incorporated a knowledge graph with CNN, RNN, LSTM and Bi-LSTM for automatic personality trait classification. In their follow-up work, Ramezani et al. (2022b) have applied attention over the knowledge graph and achieved current state-of-the-art performance.

3 Methodology

The personality trait classification model utilizes an hierarchical framework with a sentence encoder and a statement encoder. The sentence encoder generates vectors for each sentence, while the statement encoder synthesizes these vectors for the entire text. We have experimented with two tree-transformer variants as sentence encoders: constituency (CTT) and dependency (DTT) tree-transformers. A graph attention network (GAT) merges sentence representations and produces a statement representation. The heterogenous GAT (H-GAT) layer refines sentence and word nodes using the statement vector. Three model architectures have been examined, varying the configuration of the H-GAT layer. It enhances sentence and word representation by incorporating information from the statement vector. This section explains the individual components and then describes the comprehensive model.

3.1 Sentence Encoder Module

To effectively analyze an individual's personality traits through textual data, we take into account the syntactical structure of sentences. Motivated by the findings of Tai et al. (2015), we address this requirement by exploring two types of tree-structured transformer models which capture correlations between distant words and the phrasal structures present in sentences. While attention mechanisms (Bahdanau et al., 2015; Vaswani et al., 2017) have made significant strides in addressing the issue of long-distance dependencies, they still fall short when compared to tree-structured models (Ahmed et al., 2019a,b).

To convey comprehensive information about a sentence, two types of tree-based representations are employed: constituency trees, which capture distinct aspects of sentence syntax, and dependency trees, which achieve the relationships between individual words positioned at various locations within the sentence (Ahmed et al., 2019b). Through recursive computations entailing attention across branches, the models analyze each sub-tree and generate a sentence vector representation at the tree's root.

To generate the self-attention over the branches query ($\mathcal{Q}$), key ($\mathcal{K}$), and value ($\mathcal{V}$) matrices are computed (Vaswani et al., 2017) (see Eqs. 1-3).

$$\mathcal{K} = \omega_k \mathcal{M}_k \quad \text{s.t.} \quad \omega_k \in \mathbb{R}^{d\times d} \tag{1}$$

$$\mathcal{V} = \omega_v \mathcal{M}_v \quad \text{s.t.} \quad \omega_v \in \mathbb{R}^{d\times d} \tag{2}$$

$$\mathcal{Q} = \omega_q \mathcal{M}_q \quad \text{s.t.} \quad \omega_q \in \mathbb{R}^{d\times d} \tag{3}$$

For a DTT, the matrix $\mathcal{M}$ is constructed by concatenating the word vectors of all child nodes associated with each parent node. Conversely, in a CTT, the matrix $\mathcal{M}$ is formed by concatenating the word vectors within a constituent. The self-attention matrix (α) is calculated by leveraging the $\mathcal{Q}, \mathcal{V}, \mathcal{K}$ matrices in the following manner:

$$\alpha = \text{softmax}(\frac{\mathcal{Q}\,\mathcal{K}^T}{\sqrt{d_k}})\mathcal{V} \tag{4}$$

Here, the dimension of the key ($\mathcal{K}$) matrix is denoted as d_k.

To carry out multi-branch attention, denoted as $\mathcal{B}_i$, with n branches, n sets of the key ($\mathcal{K}$), query ($\mathcal{Q}$), and value ($\mathcal{V}$) matrices with n corresponding weight matrices (ω_i) are used. Subsequently, a scaled dot product attention is performed on each branch:

$$\mathcal{B}_i = \alpha_{i\in[1,n]}(\mathcal{Q}_i\ \omega_i^{\mathcal{Q}}, \mathcal{K}_i\ \omega_i^{\mathcal{K}}, \mathcal{V}_i\ \omega_i^{\mathcal{V}}) \tag{5}$$

Next, a residual connection is introduced to the tensors obtained from the multi-branch attention operation followed by a layer-wise batch normalization layer to normalize the tensor outputs and finally a scaling factor μ is used:

$$\tilde{\mathcal{B}}_i = \text{LayerNorm}(\mathcal{B}_i\omega_i^b + \mathcal{B}_i) \times \mu_i \tag{6}$$

In the subsequent stage, a position-wise CNN (PCNN) is applied to each $\tilde{\mathcal{B}}_i$. The PCNN layer consists of two convolution operations performed at each position, with a ReLU activation function separating the convolution operations:

$$\texttt{PCNN}(x) = \texttt{Conv}(\textit{ReLU}(\texttt{Conv}(x) + b_1)) + b_2 \tag{7}$$

Then, the attentive representation of the semantic sub-spaces is generated by applying a linear weighted summation over the PCNN layer-derived features (see Eq. 8). Here γ is a trainable hyper-parameter of the model that determines the weights assigned to each semantic sub-space.

$$\texttt{BranchAttn} = \sum_{i=1}^{n} \gamma_i \texttt{PCNN}(\tilde{\mathcal{B}}_i) \tag{8}$$

Finally, a residual connection is established between the output of the `BranchAttn` layer and the subsequent step. Then, a non-linear activation function (tanh) is applied to the resulting tensor. The parent node representation is then computed by performing an element-wise summation (`EwS`) which combines the representations of the child nodes:

$$\texttt{ParentNode} = \texttt{EWS}(\tanh((\chi_{att} + \chi)\omega + b)) \tag{9}$$

Here, the input features to the attention calculation module are denoted as χ, while the output features are represented as χ_{att}.

To incorporate both the word-level dependencies and the underlying phrasal information present in the sentences, mean pooling is utilized over the sentence vectors obtained from the DTT and CTT. This is elaborated in Section 3.4.

3.2 Statement Encoder Module

Over the sentence representations generated from the sentence encoding module (see Section 3.1), the Graph Attention Network (GAT) (Veličković et al., 2017) is employed to generate the vector representation of the statement. The graph $\mathcal{G} = \{V, E\}$ is designed in such a way that there is an edge between the statement node $\mathcal{D}$ and all n sentence nodes $(S_1, S_2, ..., S_n)$ in the statement. Thus, $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ where $\mathcal{V} = \{S_1, S_2, ..., S_n, \mathcal{D}\}$ and $\mathcal{E} = \{S_1 \to \mathcal{D}, S_2 \to \mathcal{D}, ..., S_n \to \mathcal{D}\}$. The sentence nodes are initialized with the sentence embeddings that are generated by the sentence encoder module and by applying mean pooling over them, $\mathcal{D}$ is initialized. GAT is applied over these sentence nodes to generate the vector representation for node $\mathcal{D}$ (see Eqs. 10-12).

$$\kappa_{\mathcal{D},S_j} = \mathsf{LeakyReLU}(\omega_a[\omega_q \mathcal{D} || \omega_k S_j]) \tag{10}$$

$$\alpha_{\mathcal{D},S_j} = \frac{\exp(\kappa_{\mathcal{D},S_j})}{\sum_{l \in \mathcal{N}_\mathcal{D}} \exp(\kappa_{\mathcal{D},l})} \tag{11}$$

$$\mathcal{D} = \sigma\Big(\sum_{j \in \mathcal{N}_\mathcal{D}} \alpha_{\mathcal{D},S_j} \omega_v S_j\Big) \tag{12}$$

Here, the concatenation operation is denoted by $||$. ω_a, ω_q, ω_k, and ω_v are the trainable weight matrices. The set of neighbouring nodes for a given node (S_i or $\mathcal{D}$) is represented by $\mathcal{N}_i$. $\alpha_{i,j}$ denotes the attention value between any two nodes in the graph. The GAT layer incorporates multi-head attention. Utilizing $\mathcal{M}$ attention heads, this multi-head attention formulation can be expressed as follows:

$$\mathcal{H}_i = ||_{m=1}^{\mathcal{M}} \sigma\Big(\sum_{j \in \mathcal{N}_\mathcal{D}} \alpha^m_{\mathcal{D},S_j} \omega^m S_j\Big) \tag{13}$$

This final hidden representation $\mathcal{H}_i$ is used as the statement representation vector ($\mathcal{D} = \mathcal{H}_i$).

3.3 Refinement module

The refinement module employs the heterogeneous graph attention network (H-GAT) to update the word and sentence embeddings based on the statement embeddings generated from the statement update module. This refinement module is inspired by the H-GAT (Wang et al., 2020). Originally designed to enhance cross-sentence relationships and to generate more informative sentence representations for extractive summarization tasks, we have adapted this approach to improve the quality of statement representations for our task. In our methodology, the H-GAT module is utilized at each iteration, following the completion of forward passes of the sentence encoder and statement encoder modules. By incorporating the statement-to-sentence, sentence-to-word, and statement-to-word update steps and subsequent forward passes of the sentence and statement encoder modules, this module enriches the statement vectors, leading to an enhancement in the overall quality of the statement representations for the personality trait detection task. This section outlines the general concept of the refinement module. The varying placements of this module are elaborated in Section 3.4.

Considering a statement has n sentences, the **statement-to-sentence update module** works on a graph $G = \{V, E\}$ where the set of vertices $V = \{S_1, S_2, ..., \mathcal{D}\}$ and set of edges $E = \{S_1 \to \mathcal{D}, S_2 \to \mathcal{D}, ..., S_n \to \mathcal{D}\}$ (similar to the graph $\mathcal{G}$ in the statement encoding module). After constructing the graph G, the feature values of the nodes are modified using a Graph Attention Network (GAT) (Veličković et al., 2017). Let $h_i \in \mathbb{R}^{d_h}$ represent the hidden states of the statement and sentence nodes, where $i \in \{1, 2, ..., (n+1)\}$, and d_h denotes the dimension of the hidden states. The GAT layer, which operates on this graph, can be formulated as follows:

$$\kappa_{i,j} = \mathsf{LeakyReLU}(\omega_a[\omega_q h_i; \omega_k h_j]) \tag{14}$$

$$\alpha_{i,j} = \frac{\exp(\kappa_{i,j})}{\sum_{l \in \mathcal{N}_i} \exp(\kappa_{i,l})} \tag{15}$$

$$\mathcal{Z}_i = \sigma\Big(\sum_{j \in \mathcal{N}_i} \alpha_{i,j} \omega_v h_j\Big) \tag{16}$$

where ω_a, ω_q, ω_k, and ω_v are the weight matrices in the GAT layer are updated during the backpropagation based on the gradients. The set of neighbouring nodes for a given node i is represented by $\mathcal{N}_i$, and the attention score between hidden states h_i and h_j is denoted as $\alpha_{i,j}$.

To enhance the expressiveness of the GAT layer, it can be extended to incorporate multi-head attention with $\mathcal{M}$ heads. This extension allows the model to capture multiple aspects or perspectives of the relationship between nodes. The formulation of the GAT layer with multi-head attention can be expressed as follows:

$$\mathcal{Z}^i = ||_{m=1}^{\mathcal{M}} \sigma\Big(\sum_{j \in \mathcal{N}_i} \alpha^m_{i,j} \omega^m h_i\Big) \tag{17}$$

Finally, a residual connection is established in the model. This connection allows the final hidden state representation h_i to incorporate the information u_i from the residual connection. The updated hidden state representation is formulated as $h_i = u_i + h_i$. This addition operation ensures that the information from previous layers is preserved and combined with the current representation, helping to alleviate the issue of vanishing gradients.

At each iteration, the sentence nodes undergo updates using the GAT layer and a position-wise feed-forward network (FFN) layer. Following the approach introduced by Wang et al. (2020), the updates are performed considering the information from the statement node. The updates can be described by the following equations:

$$\mathcal{Z}^{t+1}_{\mathcal{D}\to S} = GAT(\mathcal{H}^t_S, \mathcal{H}^t_{\mathcal{D}}, \mathcal{H}^t_{\mathcal{D}}) \quad (18)$$

$$\mathcal{H}^{t+1}_S = FFN(\mathcal{Z}^{t+1}_{\mathcal{D}\to S} + \mathcal{H}^t_{\mathcal{D}}) \quad (19)$$

In Eq. 18, at the first iteration ($t = 0$), $\mathcal{H}^0_S$ corresponds to the initial set of sentence nodes, which are obtained from the sentence encoder module. On the other hand, $\mathcal{H}^0_{\mathcal{D}}$ represents the statement representation derived from the statement encoder module. Within the GAT layer, $\mathcal{H}^t_S$ serves as the query matrix, while $\mathcal{H}^t_{\mathcal{D}}$ is utilized as both the value and key matrices. This configuration is inspired by the approach proposed by Vaswani et al. (2017), aiming to capture the attention-based relationships between the sentence nodes and the statement representation.

Both the sentence-to-word and statement-to-word update steps are designed following the same principle of the statement-to-sentence update step. The **sentence-to-word** update step tries to refine the word embeddings based on the sentence embedding so that the word vectors can preserve the essence of the sentence. For any sentence S containing p words, the word nodes are updated by a GAT layer as follows:

$$\mathcal{Z}^{t+1}_{S\to w} = GAT(\mathcal{H}^t_w, \mathcal{H}^t_S, \mathcal{H}^t_S) \quad (20)$$

$$\mathcal{H}^{t+1}_w = FFN(\mathcal{Z}^{t+1}_{S\to w} + \mathcal{H}^t_S) \quad (21)$$

where, at the first epoch ($t = 0$), $\mathcal{H}^0_w$ represents the initial set of word nodes. These word nodes correspond to the RoBERTa-based embeddings (Liu et al., 2019) for the words present in the sentence. $\mathcal{H}^t_S$ depicts the updated sentence representations obtained from the statement-to-sentence update step.

The statement-to-word update step applies GAT to produce refined word embeddings with the knowledge of the statement embedding. For the statement $\mathcal{D}$, this step is defined as:

$$\mathcal{Z}^{t+1}_{\mathcal{D}\to w} = GAT(\mathcal{H}^t_w, \mathcal{H}^t_{\mathcal{D}}, \mathcal{H}^t_{\mathcal{D}}) \quad (22)$$

$$\mathcal{H}^{t+1}_w = FFN(\mathcal{Z}^{t+1}_{\mathcal{D}\to w} + \mathcal{H}^t_{\mathcal{D}}) \quad (23)$$

where, initially ($t = 0$), $\mathcal{H}^0_w$ is the set of word nodes present in the statement $\mathcal{D}$ and initialized with the RoBERTa word embeddings. $\mathcal{H}^0_{\mathcal{D}}$ is the statement vector generated by the statement encoder.

3.4 Model Architecture

By varying the position and utilization of the refinement module units, we have investigated three architectures for the automatic personality trait detection task. The architectural structures of the proposed models are portrayed in Figure 1. In the context of the personality trait detection task, all of the models require two forward passes separated by a refinement step. The first forward pass is a common step shared by all of the models. During the initial forward pass, RoBERTa word embeddings are used as the initial input to the model. Subsequently, the aforementioned inputs undergo simultaneous processing by both the DTTs and CTTs in the sentence encoder module. This step outputs two sentence representations for each sentence in the statement: $S_{DTT} \in \{S^1_{DTT}, S^2_{DTT}, ..., S^n_{DTT}\}$ and $S_{CTT} \in \{S^1_{CTT}, S^2_{CTT}, ..., S^n_{CTT}\}$, accordingly. Following this stage, a mean-pooling procedure is executed, resulting in the generation of an intermediate sentence representation denoted as S_{avg}. Thus for a statement D containing n sentences, n sentence representations ($S_{avg} \in \{S^1_{avg}, S^2_{avg}, ..., S^n_{avg}\}$) are generated. These sentence representations from the sentence encoder are passed to the statement encoder module. The GAT layer in the statement encoder computes the statement representation $\mathcal{D}$ and the first forward pass ends here.

The major difference between the investigated models is the utilization and design of the refinement step. For the first model (see Figure 1(a)), the refinement module uses only the statement-to-word update step. In the second investigated model (see Figure 1(b)), the refinement module uses the statement-to-sentence and the sentence-to-word update steps. The statement-to-sentence step, at first, updates the averaged sentence representations (S_{avg}). These updated sentence representations are then used by the sentence-to-word update module to update the word embeddings. The last model (see Figure 1(c)) also uses the statement-

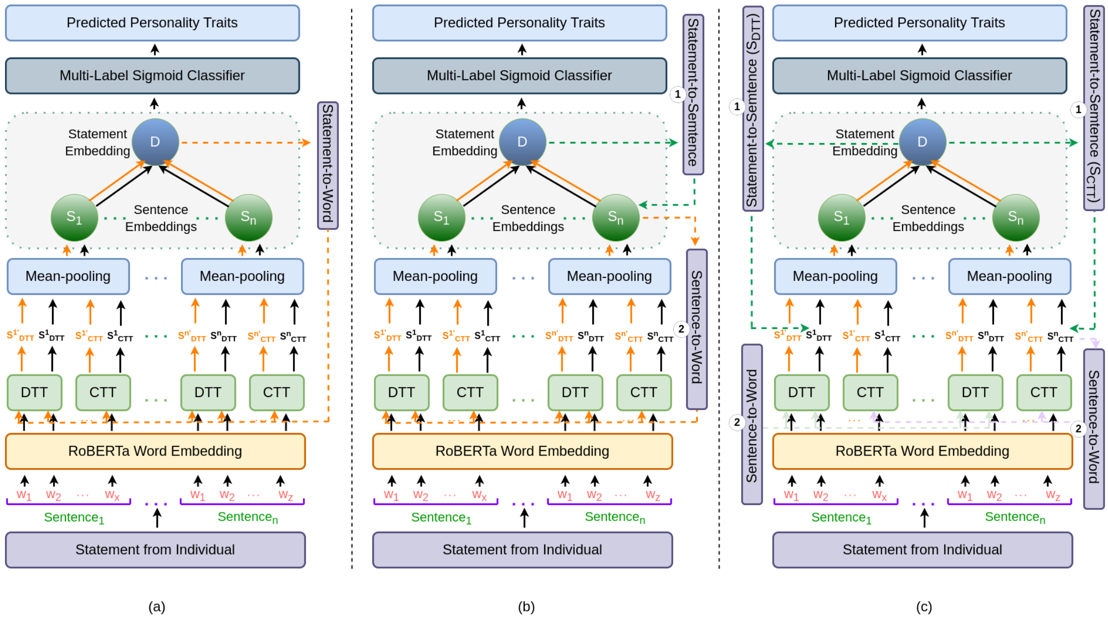

Figure 1: Structure of the investigated systems for identifying personality traits. (a) the word embeddings are updated using statement vector, (b) the statement vector updates the S_{avg} and subsequently updates the word embeddings, and (c) the statement vector updates the S_{DTT} and S_{CTT}. These sentence vector updates the word embeddings separately. All the unbroken orange straight lines indicate the second forward pass with the updated word vectors. For (b) and (c) the refinement steps labeled with numbers indicate the order of occurrence.

to-sentence and the sentence-to-word update steps. But here, the statement-to-sentence update module updates the S_{DTT} (S'_{DTT}) and S_{CTT} (S'_{CTT}). Then, the sentence-to-word refinement step is utilized twice: once to update the word embeddings based on the updated S_{DTT}, and another time based on the updated S_{CTT}.

After the refinement module is employed, the second forward pass is initiated. For the first two models, with the updated word embeddings, the forward pass is the same as the first forward pass. But for the third model, the sentence encoder module works with two different word embeddings. The CTT intakes the word embeddings updated by the S'_{CTT}, and the DTT is fed with word embeddings updated by the S'_{DTT} as inputs. The following steps are similar to the other two models. This second forward pass generates a refined statement vector ($\mathcal{D}'$). Subsequently, $\mathcal{D}'$ is fed into a dense layer, followed by a *sigmoid* classifier that assigns a probability score to each individual personality trait. For model training, we have employed the binary cross-entropy loss function to evaluate and calculate the overall loss of the model.

4 Experimental Setup

Here, we give an assessment of our model's efficacy in discerning personality traits, employing accuracy and F-1 score as the evaluation metrics. Individuals can exhibit multiple traits concurrently, given that these characteristics are not inherently exclusive. Consequently, we have framed the identification of personality traits as a multi-label classification problem, gauging the model's performance against each distinct class label. Furthermore, this section provides a synopsis of the benchmark datasets used in our experiments.

We have experimented on two benchmark corpora: (i) Essays (Pennebaker and King, 1999), (ii) Kaggle MBTI (Jolly, 2017). The "Essays" dataset encompasses a collection of 2468 compositions penned by students, meticulously annotated with binary labels pertaining to five distinct personality traits: Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N). They were annotated by analyzing a standardized self-report questionnaire for each student. The "Kaggle MBTI" dataset comprises a substantial collection of 8675 records, with each entry containing the 50 most recent contributions made by individuals on the PersonalityCafe website. Each entry is associated with a binary MBTI personality type. This corpus encompasses four binary class labels, namely: (i) Extroversion or Introversion (I/E), (ii) Sensing or Intuition (S/I), (iii) Thinking or Feeling (T/F), and (iv) Judging or Perceiving (J/P). The data pre-processing step follows the approach used in (to preserve anonymity we don't cite the work. Upon acceptance we will add the citation.). The statistics of the corpora are presented in Appendix A.3. For both corpora, the same train, test, and validation splits are used as in (Mehta et al., 2020).

The model employs an initial learning rate of 0.1, which is subsequently reduced by 80% in each iteration if the validation accuracy declines compared to the previous iteration. The batch size is 10. For the tree-transformers, the same hyper-parameter settings are used as in Ahmed et al. (2019b). The statement encoding unit utilizes a GAT (Graph Attention Network) with six attention heads. The model's parameters are trained using the "Adagrad" optimizer (Lydia and Francis, 2019).

The output representations for the sentence encoders (DTT and CTT), the statement encoder, and the model itself, are 768-dimensional vectors. The model employs two forward passes to generate the statement vector. During the first forward pass, RoBERTa word embeddings are utilized. In the second pass, the updated word representations obtained from the "refinement module" are employed, as described in Section 3.4. The performance evaluation of our models has been conducted using 10-fold cross-validation. To facilitate this cross-validation process, we have utilized the StratifiedKFold function from the scikit-learn package. All experiments have been conducted in an Ubuntu 22.04 LTE environment, leveraging a 48GB NVIDIA RTX A6000 GPU. For parsing the sentences and generating the tree representations, we have used the Stanford Core-NLP parser.

5 Analysis of Results

Tables 1 and 2 showcase the performance of the proposed models on the Essays and Kaggle MBTI corpora, respectively. The results clearly demonstrate that our proposed models, with one exception, have outperformed previous models, including the current state-of-the-art (SOTA) (Ramezani et al., 2022a), by a significant margin without using any additional features like the other models do. For the Essays corpus, the model incorporating the statement-to-word update module exhibits slightly lower accuracy and F-1 scores compared to the current SOTA (Ramezani et al., 2022a). However, the second model, which incorporates the statement-to-sentence (S_{avg}) and sentence (S_{avg})-to-word refinements, surpasses the current SOTA by an average margin of 2.6 percentage points (p.p.) in accuracy and 3.7 p.p. in F-1 score. The third model, which incorporates separate statement-to-sentence (S_{DTT} and S_{CTT}) and sentence (S_{DTT} and S_{CTT})-to-word update modules, exhibits an additional 0.6 p.p. and 1.0 p.p. improvement over the second model in accuracy and F-1 score, respectively, on average. Among all the proposed models, the third model yields the best performance. The same trend in improvement is observed for each individual class, as well. Compared to the BERT-based models, the performance gain is substantially higher. The third model has gained 15.0 and 15.9 p.p. accuracy boosts over the best performing BERT-based models: BERT-base + MLP and BERT-large + MLP (Mehta et al., 2020), respectively. A similar performance boost is observed in the experiments with the Kaggle MBTI corpus, as well. In terms of accuracy, our best performing proposed model has shown 9.6 and 8.6 p.p. accuracy gain over the BERT + Bagged SVM (Kazameini et al., 2020b) and BERT-large + MLP, respectively.

A major reason behind such improvement over the BERT-based models is that the BERT-based models work with only the first 512 tokens of the statements due to the token input limitation of BERT. Our model has surpassed that limitation by using the tree-transformer based sentence encoder module. It works with individual sentences from the statement and thus it is not dependent on the statement length. Furthermore, the statement encoder module imposes attention over the sentences which helps the model understand which sentences are important when identifying personality traits.

Another reason for the improved performance is that the models proposed by Kazameini et al. (2020b) and Mehta et al. (2020) use the pre-trained BERT models without any fine-tuning for this task which does not allow the models to inherit task specific knowledge. On the other hand, our proposed models, using the refinement module, update the word embeddings as well, based on the generated statement representation which in the end helps to produce more enriched statement representations. This approach is quite similar to the concept of fine-tuning BERT-based models, but demands less computational resources (122M vs 345M parameters) and one-fourth of the training time compared to the BERT-fine-tuning. This has made our model more suitable to run on computers with less computational resources. The proposed

Table 1: Performance analysis of the proposed models along with the other prominent works over the Essays dataset. The best results are presented in bold texts. Missing values are presented with -.

	F-1 Score						Accuracy					
Model	O	C	E	A	N	Ave.	O	C	E	A	N	Ave.
Previous Works												
BERT + Bagged SVM (Kazameini et al., 2020b)	-	-	-	-	-	-	62.1	57.8	59.3	56.5	59.4	59.0
Psycholinguistic + MLP (Mehta et al., 2020)	-	-	-	-	-	-	60.4	57.3	56.9	57.0	59.8	58.3
BERT-base + MLP (Mehta et al., 2020)	-	-	-	-	-	-	64.6	59.2	60.0	58.8	60.5	60.6
BERT-large + MLP (Mehta et al., 2020)	-	-	-	-	-	-	63.4	58.9	59.2	58.3	58.9	59.7
CNN-AdaBoost-2channel (Mohades Deilami et al., 2022)	-	-	-	-	-	-	61.9	62.1	59.9	60.6	64.9	61.9
KGrAT-Net (Ramezani et al., 2022a)	75.0	76.5	78.1	72.8	69.9	74.4	72.2	73.4	74.2	71.2	71.0	72.4
Proposed Models												
Model-1	74.6	76.2	77.6	72.6	74.1	75.0	71.9	73.2	73.8	71.0	70.6	72.1
Model-2	77.8	78.8	81.1	75.5	77.6	78.1	74.9	75.8	77.2	73.9	73.1	75.0
Model-3	78.5	79.6	81.8	76.1	79.4	79.1	**75.6**	**76.4**	**77.8**	**74.5**	**73.8**	**75.6**

Table 2: Performance analysis of the proposed models along with the other prominent works over the Kaggle MBTI dataset. The best results are presented in bold texts. Missing values are presented with -.

	F-1 Score					Accuracy				
Model	I/E	S/I	T/F	P/J	Ave.	I/E	S/I	T/F	P/J	Ave.
Previous Works										
BERT + Bagged SVM (Kazameini et al., 2020b)	-	-	-	-	-	79.0	86.0	74.2	65.4	76.1
Psycholinguistic + MLP (Mehta et al., 2020)	-	-	-	-	-	77.6	86.3	72.0	61.9	74.5
BERT-base + MLP (Mehta et al., 2020)	-	-	-	-	-	78.3	86.4	74.4	64.4	75.9
BERT-large + MLP (Mehta et al., 2020)	-	-	-	-	-	78.8	86.3	76.1	67.2	77.1
Proposed Models										
Model-1	88.6	95.3	86.4	74.6	86.2	84.3	90.4	83.2	73.0	82.7
Model-2	89.1	97.8	89.1	77.6	88.4	84.7	93.2	85.8	76.1	85.0
Model-3	89.6	98.4	89.8	78.4	89.1	**85.3**	**93.9**	**86.7**	**76.7**	**85.7**

model takes slightly more time compared to the BERT-large model in the testing phase. However, there are ways to parallelize our model to reduce this computational time difference.

Using the tree-transformers allows the model to better capture structural knowledge at the sentence level. Through our experiment, we have found that while dealing with complex sentences the BERT-based models fail to identify all of the personality traits properly as there exist dependencies between different phrases at various distances in the sentence (see Appendix A.2 for an example).

Among the proposed models, we observe that the second and third models perform much better than the first one. The second model directly refines the word representations based on the generated statement vector. It requires a lot of nodes to be refined all together based on the value of only one node (the statement vector) and the refinement ignores the sentence-level information. The third model uses two separate statement-to-sentence and sentence-to-word update modules so the two tree-transformers get different word embeddings and allows the model to have more semantic information during the second pass, helping it to achieve higher performance compared to the second model. To show the importance of the individual components of these methods, an ablation study is shown in Appendix A. The ablation study demonstrates that the refinement module helps the model to achieve superior performance. However, our model makes wrong predictions in some cases. Each class within the Essays corpus is exemplified by a singular case in Appendix A.4.

6 Conclusions and Future Work

This study introduces three innovative architectures that leverage an hierarchical structure of tree-transformers and a graph attention network for the classification of personality traits inferred from written text. The refinement module proposed in this research aids in the precise adjustment of word vectors while preserving enriched semantics and syntactical information. The proposed models have demonstrated a substantial performance improvement compared to previous prominent works. A potential extension of this work could involve the incorporation of a knowledge graph, similar to the approach taken by Ramezani et al. (2022a).

Limitations

In this study, our focus was primarily on the Big Five Model (OCEAN) and Myers-Briggs Type Indicator (MBTI) personality trait classifications. However, it is important to note that there are two other noteworthy personality trait models that warrant attention: Eysenck's Personality Dimensions and the HEXACO Model. These alternative models offer distinct frameworks for understanding and categorizing personality traits. We are not sure how well these models will perform when working with them.

Moreover, while the proposed models have indeed exhibited a substantial performance improvement, it is important to acknowledge that there is a trade-off in terms of computational time. The utilization of two forward passes plus parsing required for the tree-structured transformers in the model leads to an increase in time required for the generation of results compared to the other models. This computational overhead should be taken into consideration when considering the deployment and scalability of the proposed models in practical applications. However, with some parallelization in the model implementation, the computational time it requires can be reduced.

Ethical Disclaimer

This work is not intended to be used in clinical practice or by individuals to determine mental health conditions.

References

Mahtab Ahmed, Muhammad Rifayat Samee, and Robert E. Mercer. 2019a. Improving tree-LSTM with tree attention. In *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, pages 247–254.

Mahtab Ahmed, Muhammad Rifayat Samee, and Robert E Mercer. 2019b. You only need attention to traverse trees. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 316–322.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations*.

Laura Campbell-Sills, Sharon L Cohan, and Murray B Stein. 2006. Relationship of resilience to personality, coping, and psychiatric symptoms in young adults. *Behaviour Research and Therapy*, 44(4):585–599.

Matej Gjurković, Mladen Karan, Iva Vukojević, Mihaela Bošnjak, and Jan Šnajder. 2020. Pandora talks: Personality and demographics on reddit. *arXiv preprint arXiv:2004.04460*.

David P Goldberg and Peter Huxley. 1992. *Common Mental Disorders: A Bio-social Model.* Tavistock/Routledge.

Oliver P. John, Laura P. Naumann, and Cristopher J. Soto. 2008. Paradigm shift to the integrative Big Five Trait taxonomy. *Handbook of Personality: Theory and Research*, pages 114–158.

Mitchell Jolly. 2017. (mbti) myers-briggs personality type dataset.

Mayuri Pundlik Kalghatgi, Manjula Ramannavar, and Nandini S Sidnal. 2015. A neural network approach to personality prediction based on the big-five model. *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, 2(8):56–63.

Amirmohammad Kazameini, Samin Fatehi, Yash Mehta, Sauleh Eetemadi, and Erik Cambria. 2020a. Personality trait detection using bagged SVM over BERT word embedding ensembles. *arXiv preprint arXiv:2010.01309*.

Amirmohammad Kazameini, Samin Fatehi, Yash Mehta, Sauleh Eetemadi, and Erik Cambria. 2020b. Personality trait detection using bagged SVM over BERT word embedding ensembles. In *Proceedings of the The Fourth Widening Natural Language Processing Workshop*.

Amirmohammad Kazemeini, Sudipta Singha Roy, Robert E Mercer, and Erik Cambria. 2021. Interpretable representation learning for personality detection. In *2021 International Conference on Data Mining Workshops (ICDMW)*, pages 158–165.

Kenneth S Kendler, Ronald C Kessler, Michael C Neale, Andrew C Heath, and Lindon J Eaves. 1993. The prediction of major depression in women: toward an integrated etiologic model. *American journal of psychiatry*, 150:1139–1139.

Jihyun Kim, Kelly Merrill Jr, Chad Collins, and Hocheol Yang. 2021. Social tv viewing during the covid-19 lockdown: The mediating role of social presence. *Technology in Society*, 67:101733.

Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences*, 110(15):5802–5805.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*.

Agnes Lydia and Sagayaraj Francis. 2019. Adagrad—an optimizer for stochastic gradient descent. *Int. J. Inf. Comput. Sci*, 6(5):566–568.

Veronica Lynn, Niranjan Balasubramanian, and H Andrew Schwartz. 2020. Hierarchical modeling for user personality prediction: The role of message-level attention. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5306–5316.

Yash Mehta, Samin Fatehi, Amirmohammad Kazameini, Clemens Stachl, Erik Cambria, and Sauleh Eetemadi. 2020. Bottom-up and top-down: Predicting personality with psycholinguistic and language model features. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 1184–1189.

Yash Mehta, Navonil Majumder, Alexander Gelbukh, and Erik Cambria. 2019. Recent trends in deep learning based personality detection. *Artificial Intelligence Review*, pages 1–27.

Fatemeh Mohades Deilami, Hossein Sadr, and Morteza Tarkhan. 2022. Contextualized multidimensional personality recognition using combination of deep neural network and ensemble learning. *Neural Processing Letters*, 54(5):3811–3828.

I.B. Myers, L.K. Kirby, and K.D. Myers. 2000. *Introduction to Type: A Guide to Understanding Your Results on the Myers-Briggs Type Indicator*. Introduction to type series. Oxford Psychologists Press.

James W Pennebaker and Laura A King. 1999. Linguistic styles: language use as an individual difference. *Journal of personality and social psychology*, 77(6):1296.

Majid Ramezani, Mohammad-Reza Feizi-Derakhshi, and Mohammad-Ali Balafar. 2022a. Knowledge graph-enabled text-based automatic personality prediction. *Computational Intelligence and Neuroscience*, 2022.

Majid Ramezani, Mohammad-Reza Feizi-Derakhshi, and Mohammad-Ali Balafar. 2022b. Text-based automatic personality prediction using kgrat-net: a knowledge graph attention network classifier. *Scientific Reports*, 12(1):21453.

Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence embeddings using siamese BERT-networks. *arXiv preprint arXiv:1908.10084*.

Brent W Roberts, Nathan R Kuncel, Rebecca Shiner, Avshalom Caspi, and Lewis R Goldberg. 2007. The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4):313–45.

Alexandra Samet. 2020. 2020 us social media usage: How the coronavirus is changing consumer behavior.

Clemens Stachl, Ryan L Boyd, Kai T Horstmann, Poruz Khambatta, Sandra Matz, and Gabriella M Harari. 2021. Computational personality assessment-an overview and perspective. *PsyArXiv*.

Ming-Hsiang Su, Chung-Hsien Wu, and Yu-Ting Zheng. 2016. Exploiting turn-taking temporal evolution for personality trait perception in dyadic conversations. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(4):733–744.

Xiangguo Sun, Bo Liu, Jiuxin Cao, Junzhou Luo, and Xiaojun Shen. 2018. Who am I? Personality detection based on deep learning for texts. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–6.

Kai Sheng Tai, Richard Socher, and Christopher D. Manning. 2015. Improved semantic representations from tree-structured long short-term memory networks. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1556–1566.

Tommy Tandera, Derwin Suhartono, Rini Wongso, Yen Lina Prasetio, et al. 2017. Personality prediction system from facebook users. *Procedia computer science*, 116:604–611.

Niels Van de Ven, Aniek Bogaert, Alec Serlie, Mark J Brandt, and Jaap JA Denissen. 2017. Personality perception based on linkedin profiles. *Journal of Managerial Psychology*, 32(6):418–429.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*, 30.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903*.

Danqing Wang, Pengfei Liu, Yining Zheng, Xipeng Qiu, and Xuanjing Huang. 2020. Heterogeneous graph neural networks for extractive document summarization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6209–6219.

A Appendix

A.1 Ablation Study

Table 3: Ablation Study on the Essays dataset. Here, CTT + GAT is the model where sentences are encoded with only the constituency tree-transformer (CTT) and only the graph attention network (GAT) generates the statement encoding. No refinement module is used. DTT + GAT uses the dependency tree-transformer (DTT) as the sentence encoder and GAT as the statement encoder without any refinement module. DTT + CTT + GAT takes the point-wise average of the sentence representations generated from the DTT and CTT, and the GAT layer computes the statement vector. No refinement module is used here, as well. All the performances are accuracy (in %).

Model	O	C	E	A	N	Average
CTT + GAT	69.2	68.8	65.9	65.3	68.3	67.5
DTT + GAT	70.1	69.2	66.5	64.8	69.0	67.9
DTT + CTT + GAT	70.9	70.3	67.5	65.7	69.7	68.82

Table 4: Ablation Study on the Kaggle MBTI dataset. Here, CTT + GAT is the model where sentences are encoded with only the constituency tree-transformer (CTT) and only the graph attention network (GAT) generates the statement encoding. No refinement module is used. DTT + GAT uses the dependency tree-transformer (DTT) as the sentence encoder and GAT as the statement encoder without any refinement module. DTT + CTT + GAT takes the point-wise average of the sentence representations generated from the DTT and CTT, and the GAT layer computes the statement vector. No refinement module is used here, as well. All the performances are accuracy (in %)

Model	I/E	S/I	T/F	P/J	Average
CTT + GAT	82.0	88.8	79.1	70.2	80.0
DTT + GAT	82.5	89.3	79.9	70.6	80.6
DTT + CTT + GAT	82.5	89.2	80.5	71.0	80.8

Observing the results in Tables 1, & 3, and Tables 2, & 4, we can clearly say that the performance of the individual units are much lower compared to the proposed models. The refinement unit, present in the proposed models, plays the vital role in the performance boost achieved by the three investigated models.

A.2 Example of BERT-based model providing incorrect personality traits

> Through my academic journey, I evolved from a beginner in school to a more experienced student in college, encountering various challenges, making friends, and gaining knowledge and skills, ultimately realizing that my education wasn't just about physical classrooms, but a path to becoming a better learner and thinker, preparing me for the future.

Here, the original labels are "openness" and "conscientiousness". The BERT-based model identifies the "openness" properly, however fails to capture "conscientiousness". Rather it predicts "extroversion". From Table 1 it is observed that the BERT-based models have always shown their best results on "openness" compared to the other classes. The tree-transformer/GAT model trained on the Essays corpus for OCEAN trait classification predicts the personality traits correctly.

A.3 Statistics of the Corpora

A.3.1 Statistics of the Essays Corpus

Table 5: Statistics of the Essays dataset.

	Openness	**Conscientiousness**	**Extraversion**	**Agreeableness**	**Neuroticism**
Positive	1271	1253	1276	1310	1233
Negative	1197	1215	1192	1158	1235

A.3.2 Statistics of the MBTI Corpus

Table 6: Statistics of the MBTI dataset.

Category	Number of Samples
Extroversion	1999
Introversion	6677
Sensing	1197
Intuition	7479
Thinking	3981
Feeling	4695
Judging	3434
Perceiving	5242

A.4 Errors in Prediction Made by the Proposed Model

A.4.1 Class: Openness

> I just got back from your class, so I decided that I should start to type this paper. I am very happy with my classes, even though I feel like they are going to be rather difficult this year, especially my Calculus class. I have a hard time understanding what my professor is saying. I end up have to go home and teach myself most of the information. Well that's enough about school. I just thought about my exgirlfriend. I have very strong emotions about her. I know that she was my first love. But I also am so mad at sometimes. We had talked about me going off to college and we knew that it probably would work about, so we decided that we would date other people. From my experience this really does work out. The first girl that I dated after her was a girl from my waiting job in New Braunfels. I decided that I should tell my exgirlfriend, whose name is Genie, about the girl. This was a very big mistake. Genie came to the restaurant where I worked and caused a big scene. But this isn't the only thing that makes me mad. Things are totally different now that we decided to see other people. We don't get along and we can't talk to each other. I think women need to just make up their mind. They all act like want this perfect gentlemen that does everything for them, but when the actually get that they don't know how to treat it. Usually the go to far and try to take advantage of it and then the guy starts to despise the girl. I don't really wish that things were back the way they were, I just wish that we could still get along. I really miss talking to her. She was a person that I could tell everything to and still feel comfortable about doing so. I am lucky though, because I have a sister that I am very close to. She also goes to UT and she has been a very big help with getting me settled in here in Austin. She only lives a couple of blocks away from me and she is there for me whenever I need anything, as I am for her. This is my freshman year and I am already dreaming that college would be over. It isn't that I don't enjoy Austin or College, it is just that I am tired of school. I wish that there could be a step in your life that you could just skip, but that is impossible. I would love to just be able to be settled in to a good paying job, but since that will never happen I am prepared to work now to enjoy the benefits later.

This particular instance is designated as negative within the category of "Openness." Nonetheless, our model has made an incorrect prediction by classifying it as positive in this specific class.

A.4.2 Extraversion

> Right now, I am sitting here sick to my stomach and the world feels so small. I am waiting for a phone call that is so important, and if I don't get it, I am going to feel like a really big loser. Yes, I did just get all the blessings I could ever ask for, so I am selfish to be wanting more, but its something I really really want. All I want is to make my parents proud and to give my family

something they can brag about. I have spent my whole life wanting to achieve the best, and I get so sick when I let myself down. Rejection sucks. its so hot in here, and as all my friends call because they just got the call," I feel like a loser. I am proud of myself- but rejection is not something I handle well? What if the call does not come– will I cry, will I blame my inabilities on something else, how will I react? The anxiety I feel right now is extreme. On top of all that, I am homesick. I have a great life here in Austin, but since my family is a huge part of my life, I feel kind of left out being so far away. Everything back home seems to go on without me. my roommate here is annoying and the tv here is always on. she follows me around and sometimes I feel used because she really does not know people here. She is not in a sorority and so sometimes I feel as if she is angry at me for that. I am so anxious. my boyfriend is supportive too, but I wonder sometimes if he really has deep feelings for me. Yes, I know about his fear of commitment and all that crap, but we have been together for way too long for me not to feel totally secure with him. Oh, that stupid seventh heaven song. turn off the dang tv. All I want is peace and quiet without all the noise. Oh, and I have to worry about yesterday too. My sorority is awesome, but it makes me really uncomfortable to drink around some of them. Yes, I know. Its silly if we all drink together. But, sometimes I feel as if I have this image that I have to uphold. and that image reflects back onto all aspects of my life. my family, my faith, my school, my friends. How do I act? How do I dress? Who do I associate myself with? All of these things constantly flood my brain, and sometimes all I want to do is get far away from those thoughts. Do people love me for me? Do they love me for who I am here or the grades I make or the house I live in or the money my parents make? How do people view me? And that tv, always on. what I would give for that chatter to stop for 10 minutes. I can't even study with the noise. I am worried about this year. I need a job, I have bills to pay, I am in hard classes. how will I measure up? I love my life, I love my life. but I could seriously do without the stress. I am determined, and I already have accomplished so much this semester, but will it end? I want it to stay this way, but there is so much to lose. I am scared that I will lose it. How do I not lose it? I pray all the time, and I count my blessings. its hot in my apartment and it smells like paint. why did I choose to live in an apartment with a girl I don't like? What possessed me to do this? Did I feel independent and like a big girl? Now I feel young and naive, and way out of my league. oh, the insanity, but good things come to those who wait and I put all my trust into a higher being so things WILL work out.

This statement has been categorized as positive within the "Extraversion" class; nevertheless, our model has erred in its prognostication, misclassifying it as negative.

A.4.3 Conscientiousness

ever since my boyfriend got this new job as a community assistant in an apartment complex, it doesn't seem like he has any time left over to spend with me. also, since he is a higher rank in rotc, he is even busier. so i question. what's going to happen to us? i ask him over and over again and he just gets upset. what am i supposed to think? every time this happens, we end up in an argument and threaten to break up which really hurts. i mean, he can't play with my emotions like that. it's not fair that he can have me waiting for him and giving up all my other plans in the hope that maybe this time, he'll come see me or make plans with never happens. it's not fair how he can just have me on the side when it's convenient to him. why is is that he seems like a totally different person now. not the same from the guy that i met more than a year ago. how can someone just change overnight? i am upset that when he does come and see me, it's is timed cause he says he's trying to squeeze me into his busy schedule. it make me feel like i am in prison and getting visitation rights or something. relationships shouldn't be like that. it was never like that in the begining. but he says he's a different person now. he just called right now and hung up on me because i told him i couldn't talk cause i was doing this thing for the psychology class. he's mad. but what am i supposed to do? after all, the reason i am here, is to go to school and learn and stuff. if he expects me to understand everything he does why can't he understand that i need to do this thing. i feel like i'm gaining a little bit of weight and that bothers me a lot. yet, i'm too stubborn to get into a diet and too lazy to go excercise at the gym. i am sooooo stresed out. not just from the crap i have to put up with my boyfriend but also because of school work and the crap i have to put up to with work. work does not seem fun anymore. it was in the begining when i first started working there for more than a year ago. maybe because it was my very first job and i was getting paid more that i thought i would be. or maybe it was cause i'm new in town and was meeting lots of people then who are my age. but now, it seems like work is just a drag. maybe i'm jealous cause my boyfriend has this wonderful job or may be it's cause a lot of the people and managers that i started working with left to another state or for another occupation and just wanted to get away. i need the money that is why i am still working there. i applied at the hospital a couple of weeks ago but they haven't called me back or anything. then last week, i decided i wanted to volunteer at the children's hospital and when i called to inquire about it to see what i got to do, they told me that they were good. they were good? how can that be. they're a hospital. i thought they always needed help. and i was going to do some services for free. it's not like i was going to ask pay or anything. it was going to be free. my boyfriend's roomate's mom works there and the roomate had told me that he was going to ask his mom to give me a job and he did and she said that all i needed was to give her the hours that i can work. i mean, i can do that but it would be really awkward in my position because the mom is my boyfriends ex mom. i just didn't want to be in that position you know? and i really need to start working in the nursing field and get out of being a cashier at heb because that's my major, nursing. that's another thing i was worried about. what if i don't get accepted to nursing school next semester? then what am i going to do? maybe i can switch to pharmacy just like what my friend did. but i don't think it will be any easier or anything.

This statement is annotated as non-conscientiousness in the corpus. However, our model has predicted the personality trait of the author of this statement as having conscientiousness.

A.4.4 Agreeableness

I thought I would because I've visited with my friends so many times before, but now that I'm actually here it's finally true. I'm away from my parents, it's so great. I live with three great girls in my suite and we're so popular here. I've always been a socially outgoing person, but now I feel like it's going to work. there are always large numbers of people in our living room, bringing in food or beer to contribute to our refrigerator; everyone munches from it. and it's OK. the RA told us about this girl in another room who got so upset because her roommate ate her store bought cookies without asking; she called her mom and was so upset. I'm so glad its not like that here. we all contribute and all consume. But it's not like there's always noise and party's here. only when we all decide. if one person wants to read or study or sleep, we're really considerate. I hope that lasts, I'm pretty sure it will. At our building there are many foreign exchange students which is always a plus because, come on, who minds a foreign accent every once in a while. this guy from Belgium and this one from England are always watching TV in our room, which is another amusing thing: we don't have cable, or an antenna, or a VCR, so we only get FOX Channel 7. We sit around and watch whatever's on. in one way it's good because we don't have arguments over which channel to watch. maybe simplicity is the root of compromise. We had a floor meeting the other night here and they discussed some issues that had come up. it was so funny because almost all of them referred to our room's shananagans. This one guy came here from where he lives in a house to use the laundry (he's one of our friends- our referring to my roommate and I we've been friends since 2nd grad, long time, huh?) anyway, he dropped like half a box of laundry detergent on the stairwell and no one noticed for a week. the RA got mad and cleaned it up herself, but it was amusing because he doesn't even live here. another thing was the "stolen furniture" incident. we are given this loveseat-type couch in our suite's living room that can maybe seat 3 people if you're lucky. and in the lobby of the 3rd floor in front of the elevator there are 2 large couches that just block the pathway, no one ever sits in them, and they could probably seat 5 or 6. so when no one was around, my suitemate and I and 3 other people that happened to b in our room at the time helped us move our dinky little couch into the lobby which is down the hall and around a corner. we hauled the large couch down the way and we had to tilt it sideways and temporarily knock off some of the ceiling tiles just to make it in the doorway without banging down the door across from us. now we have a nice couch that is well used and the RA's are threatening to do a room check to find it. why? its going to more use. It's all kind of a double standard anyway. The head RA is always in our room hanging out and drinking our beer. he has a crush on me so he always brings us stuff and won't mention the couch to the others and lets us into the cafeteria at night. it's pretty funny, one night the night guard knocked on our door because someone had made a noise complaint. we opened the door and the guard stood in the threshold and the head RA stood behind the door quietly while we got reprimanded. it probably wouldn't have been in his best interests to b seen in there. He's only 20, but the building is changing management, so right now he's the head guy. its odd. I'm 18. finally. I could be in a management position at the pool I lifeguard at in the summers, next summer. it seems odd that I'm really an adult. when you're a kid u never think that you're ever going to get to the point where you decide when to come home and when to do this and what to do in this situation, type thing. its like the transition from high school to college really is that much of a change in that you're independent. it feels so good to finally b independent, financially, physically, emotionally. its wonderful responsibility. I am responsible for watching my budget, if I don't, no one will bail me out (well that's probably not true but you know). I guess I'm trying out freedom on borrowed wings, I can always have that security blanket if I want, but I don't want. I want to be independent. I am right now, I hope to stay that way.

This person's personality trait shouldn't be agreeable. However, our model has misclassified it.

A.4.5 Neuroticism

Every day is a rollercoaster of emotions for me. I wake up in the morning with a knot in my stomach, fearing what challenges the day might bring. Even the simplest decisions can send me into a spiral of doubt and anxiety. It's as if a never-ending storm of worry and fear rages inside me. Social interactions are a minefield; I'm constantly second-guessing what I say and how others perceive me. I replay conversations in my head, dissecting every word for hidden meanings or signs of disapproval. Criticism, no matter how constructive, feels like a personal attack, and it takes me days to recover from it. I often find myself unable to let go of past mistakes, no matter how trivial. My mind races with 'what ifs' and 'should haves.' It's a daily struggle to keep my anxiety in check and maintain a semblance of normalcy, but most days, it feels like a battle that I'm losing.

This extended paragraph provides a more detailed and vivid description of a person's experience characterized by high levels of anxiety, constant self-doubt, and sensitivity to social interactions and criticism, all of which are indicative of the neuroticism personality trait. However, our model misclassifies it.