Received: 16 October 2024 | Revised: 27 November 2024 | Accepted: 16 December 2024 | Published online: 5 February 2025

RESEARCH ARTICLE

Artificial Intelligence and Applications
2025, Vol. 00(00) 1–10
DOI: 10.47852/bonviewAIA52024561

BON VIEW PUBLISHING

PersonaG: A Quinpartite Graph Convolutional Network for Interpretable Personality Recognition from Text

Mohmad Azhar Teli[1,*] **and Manzoor Ahmad Chachoo**[1]

[1]*Department of Computer Science, University of Kashmir, India*

Abstract: Automatic personality recognition from text has wide-ranging applications in social media analysis, targeted marketing, and personalized user experiences. As such, a lot of researchers have focused on personality recognition in the last two decades. However, existing methods often rely on only shallow semantic features or psycholinguistic features to capture the semantic information in textual data. We propose PersonaG, a novel approach that integrates psycholinguistic categories with WordNet semantics to address these limitations and construct quinpartite graph representations. Our approach combines semantic relationships with the psycholinguistic categories. Classification is performed using a Dynamic Deep Graph Convolutional Network. Our results on the benchmark Essays dataset outperform recent methods, achieving state-of-the-art performance and demonstrating the superiority of our approach. To conclude, the quinpartite graph enables PersonaG to understand the latent personality patterns from text, making it a comprehensive and effective solution for personality recognition.

Keywords: personality computing, APRT, lexical hypothesis, natural language processing, graph convolution network

1. Introduction

Personality is a complex psychological construct encompassing an individual's thoughts, emotions, and behaviors [1]. The study of the personality works to clarify the contrasts in individual conduct. So, the fundamental objective of personality psychology is "to recognize inward properties of the individual from obvious practices and to research the causal connections between them" [2]. The relationship between personality and language has intrigued researchers ever since Theophrastus [3]. The idea gained more impetus with the emergence of the lexical hypothesis, which suggests that significant aspects of personality are encoded in language [4]. With the rise of social media and ever-increasing consumer data availability, it is argued that textual data can be used to train personality computational models to infer traits. These models can be extremely beneficial with widespread applications in social media analytics [5], personalized marketing [6], user experience design [7], and mental health monitoring [8]. As such the last two and a half decades have seen a rising interest in Automatic Personality Recognition from Text (APRT). A wide array of techniques ranging from traditional machine learning methods to sophisticated deep learning and graph-based approaches [9, 10] have been used. Some of these approaches have shown promising results, but there is a need for models that can effectively capture the intricate relationships between personality and natural language. The existing research is filled with benchmark datasets like YouTube blogs [11], MyPersonality [12], Pandora [13], Kaggle [14], Twisty [15], PAN-AP-2015 [16], and Essays [17], we chose Essays for our study as it is the first, yet the most notorious in terms of results [18]. Several psychological scales have been utilized for computational modeling of personality traits, but majority of the researchers follow either MBTI [19] or the BIG 5 [20]. However, we find there is a slight favor for Big 5 scale (Table 1) in [21] and fair amount of underlying correlation between the two scales [22].

Building upon the existing foundation of research, particularly the TrigNet framework [23], we introduce PersonaG. This novel approach extends the Dynamic Deep Graph Convolutional Network (DGCN) [24] with a multi-partite graph structure, dynamic multi-hop (*DmH*) mechanism, and the Learn-to-Connect (L2C) approach. PersonaG combines psycholinguistic categories with WordNet-based semantics to create a comprehensive graph representation. This multi-partite graph is used to capture the relationships among words, sentences, documents, psycholinguistic categories, and semantic associations, producing a highly interpretable model for association between personality and language. This approach offers a more nuanced method for recognizing personality traits from text, addressing a key limitation in previous work, where models often struggled to capture the full spectrum of personality expression in text.

Key contributions of PersonaG include the following:

1) A multi-partite graph structure integrating psycholinguistic and semantic information effectively.
2) Enriched node representation with the integration of WordNet embeddings.
3) A *DmH* mechanism that allows for more flexible information propagation through the graph.
4) Incorporating a L2C approach enables the model to adjust connections between nodes dynamically.

***Corresponding author:** Mohmad Azhar Teli, Department of Computer Science, University of Kashmir, India. Email: mohmadazhar.student@kashmiruniversity.net