AN OPTIMIZED PERSONALITY RECOGNITION MODEL FOR SOCIAL MEDIA USERS

BY

ADEBIYI, BUKOLA TAIBAT (PG 18/0184)

B.Sc (Ilorin)

A Dissertation submitted to the department Computer Science, College of Physical Sciences Federal University of Agriculture Abeokuta in partial fulfillment of the requirement for the award of Masters of Science Degree in Computer Science

June, 2025

ABSTRACT

Convolutional Neural Networks (CNN) have been effectively used in extracting pattern from text and determining its architecture with hyper-parameters for personality recognition from online social network. Achieving a reliable and generic model that function consistently across a wide range of platforms and user populations is difficult. Existing CNN models often lack transparency, hence selecting the most optimal solution is challenging. This study adopted particle swarm optimization (PSO) algorithm to identify optimal weight parameters for CNN, in the context of personality recognition which yields the Particle Swarm Optimization Convolutional Neutral Network (PSO-CNN) technique. Mayer Briggs Type Indicator (MBTI) Personality twitter(X) data were used as the dataset for this study. Natural language process technique was employed to analyze data from twitter(X), which included demographic and activity information of 8,328 users with more than 1.5 million tweets. The acquired data sets underwent various phases of storage, generation, analysis, data collection, pre-processing and data transmission. The dataset was divided into training and testing data. The PSO-CNN was developed and implemented along side CNN algorithm using python to extract features from pre-processed data and conduct personality detection. A set of stylistic features was incorporated at the feature extraction phase, due to the existence of non-verbal indicators of emotion within the domain of social media. The style-based feature included the normalized counts of word length. Each word in a sentence was tokenized using WordPiece method which later became a sequence of words token. The wordpiece tokenization was represented as n-gram document representation which refers to the pattern of extracting needed information from raw text. The evaluation of PSO-CNN and CNN on twitter(X) data was based on metrics such as false positive rate, accuracy, recall, precision, f-measure

and area under the curve. The performance of PSO-CNN showed average accuracy, recall, false positive rate, precision and f-measure of 98.2, 99.8, 0.30, 98.9 and 99.6%, respectively, while the performance of CNN showed average accuracy, recall, false positive rate, precision and f-measure of 76.9, 75.7, 1.3, 79.3 and 80.3%, respectively. The PSO-CNN model exhibited promising levels of robustness and generalization across various social network platforms than CNN model. The study achieved the selection of optimal CNN hyper-parameters such as the number of layers, weight parameters, filter and batch size by utilizing the Particle Swarm Optimization (PSO). The novel technique deployed demonstrated better performance for personality recognition compared to traditional CNN techniques. This study found that particle swarm optimization convolutional neural network approach holds potential applicability in diverse real-world of social network and as such could be employed in useful investigations, prevention, and intent analysis