Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1, No.10 : 2024
ISSN : **1906-9685**

PERSONALITY DETECTION THROUGH TEXT USING DEEP LEARNING

Mr. Veera. Hemanth, MCA Student, Department of Master of Computer Applications, Vignan's Institute of Information Technology (A), Visakhapatnam-530049.
Dr. B. Prasad, Assistant Professor, Department of Information Technology, Vignan's Institute of Information Technology(A),Visakhapatnam-530049.

Abstract
Personality profoundly shapes our lives, influencing decisions, well-being, health, and myriad other preferences. Traditionally described through the lens of the Big Five personality traits—Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N)—assessing personality traits has relied on time-consuming questionnaires. However, with the rise of technology, there's a growing demand for automated personality detection, driven by practical applications. This need is particularly pressing in the context of user-generated content, where individual input statements and personal content serve as valuable indicators of personality. By harnessing the wealth of self-disclosed information and digital footprints, personality traits can be inferred from user-defined data. Machine learning algorithms, especially those utilizing deep learning techniques like Convolutional Neural Networks (CNNs), offer a promising avenue for extracting personality traits from textual data. These algorithms are trained to discern patterns and nuances within user-defined input, enabling accurate personality predictions. By leveraging user-provided data, this approach facilitates a deeper understanding of an individual's personality, paving the way for tailored interventions and personalized services in various domains. This automated approach, rooted in the OCEAN traits, holds promise for personalized services, targeted interventions, and a deeper understanding of human behavior in the digital age, spanning fields from psychology to marketing and beyond.

Keywords
- Big Five Personality Traits
- Machine Learning Algorithms
- Automatic Personality Prediction
- Convolutional Neural Networks (CNN)

Introduction
Personality is a blend of an individual's actions, feelings, inspiration and thoughts. Personality has a great impact on a person's life as it also affects one's choices in life. Personality is an important psychological construct accounting for individual differences in people and aimed at explaining the wide variety of human behaviors in terms of a few, stable and measurable individual characteristics. It not only reflects an individual's consistent patterns of behavior, thought and interpersonal communication. To reliably, validly and efficiently recognize an individual's personality is a worthwhile goal.

Our personality has large impacts on one's life, health, well-being and numerous other preferences. Social media is a place where users express them to the world. Numerous research efforts have been pushed towards the automatic assessment of personality dimensions relying on a set of information gathered from social media platforms such as list of friends, interests of music and movies, endorsements and likes an individual has ever performed. Turning this information into signals and giving them as inputs to supervised learning approaches has resulted in being particularly effective and accurate in computing personality traits and types. Despite the demonstrated accuracy of these approaches, the sheer amount of information needed to put in place such a methodology makes them

unfeasible to be used in a real usage scenario. With the development of online social networks, it is easy to obtain online social texts created by users. The rich digital traces and self-disclosed personal information on social networking platforms render it possible to analyze the user's behaviors and infer their personality traits on the web. Texts are direct way for people to translate their internal thoughts and emotions into a form that others can understand.

This project uses an approach for determining personality traits depending upon the Big Five model which can be considered as multi-label classification problem. This is because an individual can possess many traits and each of these personality traits corresponds to a classifier. The approach segments texts in words, then it learns word vector representations as embeddings that are then used to feed a Convolution Neural Network classifier. We demonstrate the effectiveness of the approach by measuring the accuracy of the learned model.

In the Big Five model, character is divided into five traits:

- Openness(O)
- Conscientiousness(C)
- Extroversion(E)
- Agreeableness(A)
- Neuroticism(N)

Trait	**Description**
Openness	Curious, original, intellectual, creative and open to new ideas.
Conscientiousness	Organized, Systematic, punctual, achievement oriented and dependable.
Extraversion	Outgoing, talkative, confident, bold, sociable and enjoys being in social situations.
Agreeableness	Willing, appealing, engaging, Affable, tolerant, sensitive, trusting, kind and warm.
Neuroticism	Anxious, feelings of guilt, envy, anger, anxiety, irritable, temperamental, and moody.

Table 1.1: Big Five traits

The Personality traits can be obtained by classifying the text using deep learning. Deep learning is a subset of machine learning in artificial intelligence (AI) that has networks capable of learning unsupervised from data that is unstructured or unlabeled. Deep learning is also known as deep neural learning or deep neural network. Deep learning is a particular kind of machine learning that achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones.

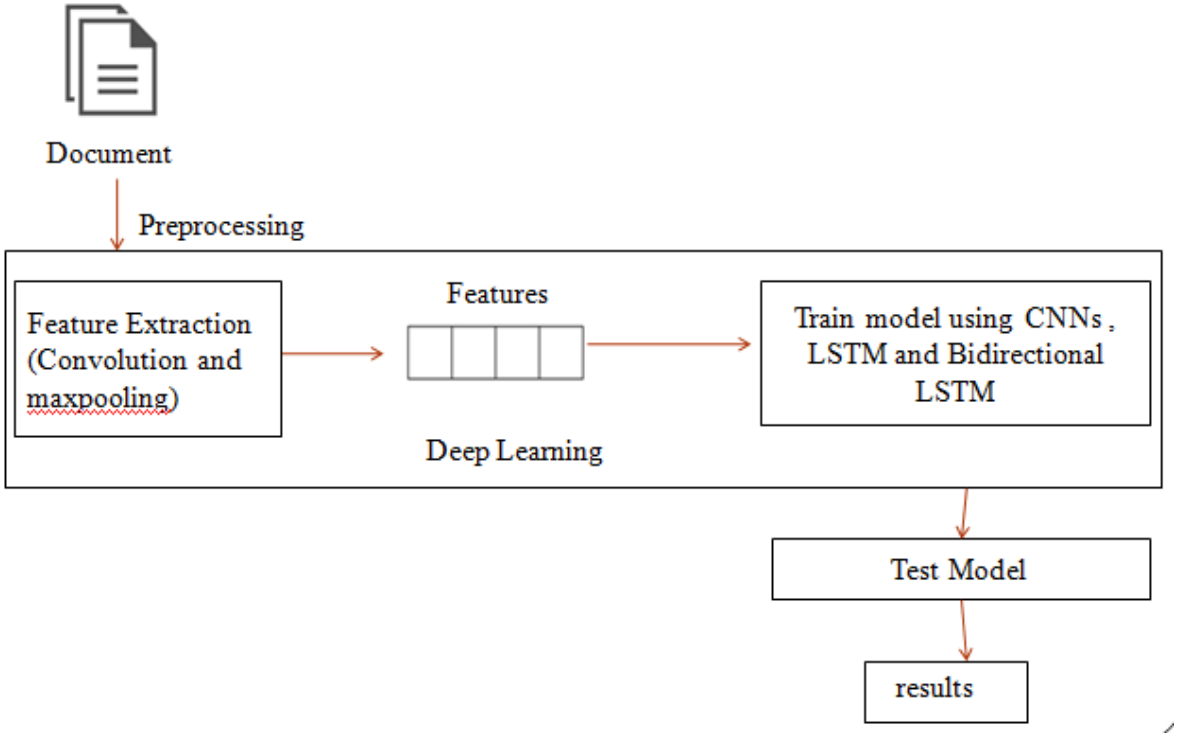
In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-

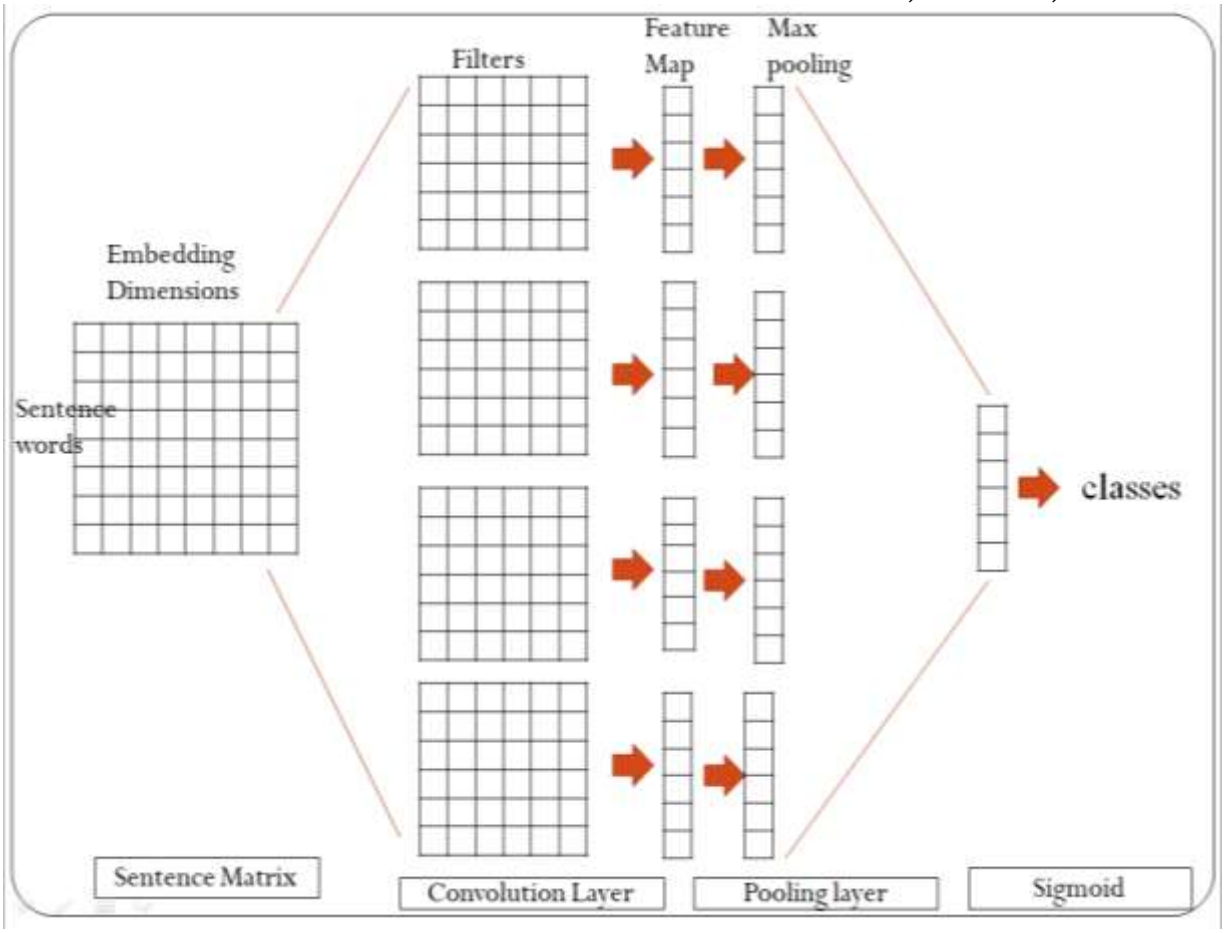
level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers.

System Architecture

Text Classification starts with an input sentence broken up into words or word embeddings i.e. low-dimensional representations generated by models like word2vec.

Words are broken up into features and are fed into a convolutional layer. The results of the convolution are "pooled" or aggregated to a representative number. This number is fed to a fully connected neural structure, which makes a classification decision based on the weights assigned to each feature within the text.





Literature Survey

The literature review examines the effort to predict people's personalities exclusively based on textual information, with a particular emphasis on the Big Five personality traits collectively referred to as OCEAN (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism). This method focuses on using text-based content analysis—such as social media posts or written essays—to identify characteristics that point to these personality traits. Through the use of machine learning algorithms, researchers hope to find patterns and connections between personality traits and linguistic cues. This presents a fresh approach to textual analysis as a means of understanding psychology and human behavior.

Research has shown that OCEAN personality traits can be predicted from text data with encouraging results. Linguistic features and personality traits have been inferred using methods such as Naive Bayes, Support Vector Machines, and Convolutional Neural Networks. Despite obstacles such as restricted data sizes and the requirement for enhanced model precision, these studies underscore the possibilities of text-oriented analysis in personality assessment. The research findings may have implications for interface design and personalized recommendations, as they offer a better understanding of how people's personalities come through in their writing.

Predicting personality from text data has significant practical implications. These kinds of insights can help with targeted marketing campaigns, customized content delivery, and personalized interactions. Furthermore, knowing how language patterns and personality traits relate to one another can provide important insights into personal preferences and differences. To improve prediction models, investigate novel machine learning paradigms, and resolve issues with data scalability and quality, more research is necessary. In the end, the goal of personality prediction from text data has potential to improve a variety of fields, including psychology and marketing, and to deepen our understanding of human behavior.

Conclusion

The aim of Automatic personality detection is to predict the personality traits of a person more accurately. There are many methods which can be used to obtain personality traits but most of them include questionnaire. Machine learning approaches can also be used to improve the accuracy of prediction. To extract relevant features convolutions are applied and model was constructed using convolution neural networks. The correct prediction of personality traits is of utmost importance, the goal is to use the model that gives highest accuracy of prediction.

References

[1] Yilin Li, Tingshao Zhu, Ang Li,Fan Zhang, and Xinguo Xu. "Web behavior and personality:A review."In Web Society (SWS), 2011 3rdSymposium on, pages81–87.IEEE, 2011.

[2] Golbeck, J., Robles, C., Edmondson, M., & Turner, K. "Predicting Personality from Twitter." 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing. doi:10.1109/passat/socialcom.2011.33

[3] Bai, S., Hao, B., Li, A., Yuan, S., Gao, R., & Zhu, T. (2013). "Predicting Big Five Personality Traits of Microblog Users". 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). doi:10.1109/wi-iat.2013.70, 2013.

[4] Pratama, B. Y., & Sarno, R. "Personality classification based on Twitter text using Naive Bayes, KNN and SVM". 2015 International Conference on Data and Software Engineering (ICoDSE). doi:10.1109/icodse.2015.7436992, 2015.

[5] Cambria, E. "Affective Computing and Sentiment Analysis". IEEE Intelligent Systems, 31(2), pages 102–107. doi:10.1109/mis.2016.31, 2016.

[6]K. C. Pramodh and Y. Vijayalata, "Automatic personality recognition of authors using big five factor model," 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, 2016, pp. 32-37

[7] Soujanya Poria, Erik Cambria, Devamanyu Hazarika, and Prateek Vij. "A deeper look into sarcastic tweets using deep convolutional neural networks". In Proceedings of COLING 2016, the 26th International Conference on Com-putational Linguistics: Technical Papers, pages 1601–1612. The COLING 2016 Organizing Committee.

[8] Yu, J., & Markov, K. "Deep learning based personality recognition from Facebook status updates." 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST). doi:10.1109/icawst.2017.8256484, 2017.

[9] D. Sewwandi, K. Perera, S. Sandaruwan, O. Lakchani, A. Nugaliyadde and S. Thelijjagoda, "Linguistic features based personality recognition using social media data," *2017 6th National Conference on Technology and Management (NCTM)*, Malabe, 2017, pp. 63-68.