International Journal of Latest Trends in Engineering and Technology
Vol.(23)Issue(1), pp.007-011
DOI: http://dx.doi.org/10.21172/1.231.02
e-ISSN:2278-621X

REVIEW ON PERSONALITY PREDICTION USING NATURAL LANGUAGE PROCESSING (NLP)

Himanshi Joshi[1], Himanshu Agrawal[1], Amrit Kaur[1]

***Abstract-** Personality prediction through sentiment analysis is an evolving research area that aims to deduce an individual's personality traits based on their emotional tone in text. It is substantive as a rightly fixed feature of an individual which mentions individual's priorities. Sentiment analysis is the process which helps in identifying people's personality. In this paper, we will review the various personality prediction techniques.*

***Keywords** –Personality, Sentiment analysis, Review, NLP, Machine learning.*

I. INTRODUCTION

Personality prediction [1] using sentiment analysis involves analyzing emotional tone and sentiment in text to infer personality traits. It leverages natural language processing (NLP) techniques to analyze textual data, providing insights into personal characteristics based on models like the Big Five personality traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism.

Numerous forms of interactions, including social relationships, music choice prediction, and the relationship between job tasks and personality, have demonstrated the relevance of personality.

The information obtained from social media users is utilised to depict user behaviour in light of various real-world scenarios. It is possible to comprehend the mood by using machine learning algorithms.

Sentiment analysis, also known as opinion mining, is a field within natural language processing (NLP) that focuses on identifying and categorizing opinions expressed in a piece of text. The primary goal is to determine the writer's attitude towards a particular topic, product, or service, typically categorized as positive, negative, or neutral. This technology has a wide range of applications, including social media monitoring, customer service, market research, and more [2].

1.1 Natural Language Processing

[1]*Department of Computer Science Engineering, Punjabi University, Patiala, Punjab, India*

The goal of the natural language processing group is to create effective algorithms that can process texts and provide their content to computer programmes. Natural language processing (NLP) is the technique by which a computer programme can comprehend spoken language from humans. An element of artificial intelligence (AI) is NLP [3].The NLP group's mission is to create software that can analyse, comprehend, and produce natural languages used by humans. Eventually, you will be able to speak to your computer as if it were a human being.

The creation of NLP applications is difficult since, historically, computers have required people to "speak" to them using a precise, clear-cut, highly structured programming language or, occasionally, a small number of well-pronounce voice commands. However, human speech is not always exact; it is frequently ambiguous, and the linguistic structure can vary depending on a wide range of intricate factors, such as social context, regional dialects, and slang. Over 50 years have passed since the development of computers, when NLP research first began [21]. The methods used in NLP nowadays are based on machine learning, a subset of artificial intelligence that looks for patterns in data and exploits them to enhance a program's comprehension.

1.2 Personality Types and Characteristics

There is a significant difference between personality types and their characteristics, which can be negative, positive, very negative, very positive, or neutral in general. On the other hand, if a person's personality is being identified from unknown individuals using social media platforms like Facebook and Twitter, it becomes critical to identify the personality. For this reason, we need a word and sentence data set with its description and context where these are used and when.

Openness, conscientiousness, extraversion, introversion, extroversion, agreeableness, sensitivity, and neuroticism are some of the other so-called big-5 personality traits [4] that set an individual apart from others. A personality can be obtained by following a specific and suitable process, as personalities have many types, qualities, and traits.

1.3 Five factor model of personality

We now briefly [5] describe the five personality traits:

1.3.1 Openness to experience

It evaluates a person's creativity, inquisitiveness, desire for novel experiences, and interest in ideas, culture, and the arts. It has to do with political liberalism, tolerance, and emotional sensitivity. Individuals with high Openness tend to be very appreciative of art, adventure, and novel or unconventional concepts. Individuals with low Openness are typically more authoritarian, traditional, and lacking in creativity. They are typically more conservative and conventional, and they have a tendency to resist change for its own sake.

1.3.2 Conscientiousness

It measures the inclination towards a planned lifestyle as opposed to an unplanned one. Individuals with high conscientiousness are more likely to be trustworthy, consistent, and well-organized. They set long-term objectives, plan ahead, and strive for success. People with low conscientiousness tend to be more laid back, impulsive, and imaginative. They are typically less constrained by guidelines and plans and more accepting. Extroversion is a scale that quantifies an individual's propensity to express good emotions, seek out happy experiences in the outside world, and interact with people. It is common for extroverts to be more gregarious, gregarious, and socially engaged. They typically enjoy being the centre of attention, are gregarious and talkative, and find it easier to make new friends.

In addition to being more at ease in their own company, introverts might be reclusive and prefer surroundings with less external stimuli.

1.3.3 Agreeableness

It evaluates how much attention a person devotes to preserving wholesome social relationships. Those with high agreeableness scores are typically amiable and caring, but they could find it difficult to deliver a harsh reality. They may find it difficult to defend their own opinions since they are more inclined to act cooperatively, trust others, and adjust to the requirements of others.

1.3.4 Neuroticism

The tendency to suffer mood swings and unpleasant feelings including guilt, wrath, anxiety, and despair is sometimes referred to as emotional instability. Extremely neurotic people can be emotionally guarded, whereas those with lower neuroticism tend to be calmer and more self-assured. Highly neurotic persons are more prone to feel stressed and anxious. A task in natural language processing called personality classification involves determining the opinions expressed about a certain topic in a text.

II.LITERATURE REVIEW

A evaluation method using personality visual attributes in the user's social features that are collected from social media was provided by Nie et al. [6]. The assessment of personality visual qualities can be divided into three main challenges: (1) feature selection (2) feature fusion and (3) feature absence. These tasks offer a fresh method for assessing the personality gap between evocative social media pictures.

A personality character analysis on the impact of online social relationships was proposed by Huang et al. [7]. Five personality theories are used in this study to measure the data collected on personality outcomes. Furthermore, a literature review examining machine learning methods for assessing personality in social networks was provided by Bleidorn and Hopwood [8]. A few personality traits, including data extraction, data gathering, and data prediction, are examined in this review.

A consistent categorization of a user's Facebook personality traits has been proposed by Lo Coco et al. [9]. This classification assesses the laws of association that have been studied between Facebook usage profiles, relational traits, and personality traits in online social interactions.

In this research, Souri et al. [10] put out the concept that individuals with similar personalities should exhibit mutual behavioural patterns while collaborating through social networks. They gathered data about users' personality features and Facebook profiles in order to analyse user behaviour with the aim of personality recognition. As a result, they developed an application using Facebook's API. The study comprises one hundred Facebook users who volunteered to participate. In May 2012, they gave the subjects a month to complete the NEO personality assessment. A link asking participants to allow the application access to their profiles was included at the end of the inquiry. Classifiers were trained using various data mining approaches based on all the gathered data in order to identify user personality based only on their profile and without completing any questionnaires. When comparing the performance of the classifiers, the boosting-decision tree model that they suggested had an accuracy of 82.2%, which was higher than that of earlier studies that used the variables in their profiles across five categories to predict personality.

Applications are increasingly using personality factors to give users a personalised experience, as highlighted by Ferwerda et al. [11]. Social media footprints proved to be a dependable tool for personality acquisition. But up until now, social media trail analysis has concentrated on what has been revealed: the contents of objects that have been revealed. When there is no content, these strategies are unable to develop personality (non-disclosure). This study examined whether or not disclosure occurred rather than the information that was communicated. They took out forty elements from various Facebook profile pages that allow users to reveal or withhold information. In addition to asking people to fill out a personality test, they requested respondents to an online survey to indicate the degree to which they reveal the things. They discovered that, based only on whether individuals share certain parts of their profiles, it is possible to predict the personalities of one hundred participants. This makes personality acquisition possible in the absence of content.

Bachrach et al. [12] shown the relationship between Facebook users' behaviour and their personality as determined by the conventional Five Factor Model. The personality profiles and Facebook profile information of 180,000 people make up their dataset. Researchers looked at relationships between Facebook profiles, including buddy network size and density, number of uploaded images, number of events attended, number of group memberships, and number of times user has been tagged in photos, and personality. Their findings indicate strong correlations between different Facebook profile characteristics and personality variables. They demonstrated how multivariate regression makes it possible to estimate a user's personality attributes based on their Facebook profile.

These forecasts are most accurate for Extraversion and Neuroticism, least accurate for Agreeableness, and in the centre for Openness and Conscientiousness.

According to Xu [13] et al., a user's ongoing social media content, such as postings on Facebook Timeline, is seen by others as a "exhibition" about them. Therefore, maintaining this exhibition of content for impression management requires deliberate and laborious work. We created a prototype called Personality Insight to help with impression management and to increase awareness of previous topics. Using computational psycholinguistic analysis, the system lets users see how their previous text posts might give away ideas about their personalities. Users can then edit their posts in response to these visualisations.

In order to assess the design, we conducted a user study. Overall, users discovered that the tool increased their awareness of the concept of impression management and the ways in which their personality may be expressed through past content. However, they also felt that the tool needed to be improved in order to provide concrete recommendations for content modification and to encourage careful consideration of impression management as one of many values people have about their digital past.

Individual personality characteristics impact users' online activities just as much as they do offline, according to Kosinski [14] et al. Based on a sample of more than a third of a million users, this study looks at the relationship between a user's online behavior—measured by the standard Five Factor Model personality questionnaire—and their personality as shown by the websites they choose and the elements of their Facebook profiles. The findings indicate a psychologically significant relationship between consumers' choices for websites, Facebook profiles, and personalities. They demonstrated the differences in personality amongst website audiences, the connections between personality and Facebook profile elements, and the ability to forecast a person's personality based on Facebook profile features.

Cai and colleagues [15] introduced a method for content detection. The components of the suggested method pertain to two distinct opinion types: positive and negative opinions. This technique's primary shortcoming was found to be that the classifier it was employed with did not close the "insight of what drives these sentiments" gap.

Raina [16] presented an opinion mining engine that performed the SA using common sense, which was knowledge retrieved from semantic and concept nets. He deployed the classifier, which has produced results up to 71% accurate with 91% for impartial view, and employed a large dataset to extract the information from news articles for evaluating the data mining engine.

Neri and colleagues [17] conducted a sentiment analysis. The applied dataset consisted of one thousand real-time Facebook posts. The posts were pertinent to Rai's (the Italian Public Broadcasting Service) newscasts and their tone.

The automated sentiment classifier for emotions was proposed by Lima et al. [18]. The Naïve Bayes algorithm was utilised for analysis, and it was applied to the tweets that contained the words. This technique had a flaw in that it classified two elements as either positive or negative while ignoring the neutral condition.

Wang et al.'s [19] main areas of interest were the quantitative computation of sentiment words and the polarity analysis of words that had never been used. The demonstrated experiments were adaptable and successful.

Through their research, Zhu et al. [20] discussed the most current advancements in sentiment analysis. Based on a survey they conducted in three main areas—sentiment analysis, feature extraction, and framework—they have made observations.

III. CONCLUSION

Personality analysis is the process which helps in identifying people's personality. The personality of the people can be expressed in positive or negative ways. Mostly, parts of speech are used as feature to extract the personality of a person. Personality analysis is an evolving field with a variety of use applications. From this paper, we concluded that the integration of sentiment analysis into personality prediction offers a nuanced understanding of individual traits through the emotional tone of language.

REFERENCES

[1] Souri, A., Hosseinpour, S. & Rahmani, A. M. (2018). Personality classification based on profiles of social networks' users and the five-factor model of personality. Human-centric Computing and Information Sciences, Vol. 8, No. 24, https://doi.org/10.1186/s13673-018-0147-4

[2] Kumar, A., & Sebastian, T. M. (2012). Sentiment analysis: A perspective on its past, present, and future. International Journal of Intelligent Systems and Applications, 4(10), 1-14. https://doi.org/10.5815/ijisa.2012.10.01

[3] Garg, P., Singh, R. & Goel, M. B. (2020). A Novel Approach to Implement Natural Language Interface for Database in Punjabi Language. Journal of Xidian University, Vol. 14, No. 4, pp. 1230-1238.

[4] Azucar, Danny, Marengo D. & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. Personality and Individual Differences, Vol. 124, pp. 150-159.

[5] Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D. & Graepel, T. (2013). Manifestations of user personality in website choice and behaviour on online social networks.

[6] Nie J, Wei Z, Li Z, Yan Y, Huang L. (2018). Understanding personality of portrait by social embedding visual features. Multimedia Tools Applications. https://doi.org/10.1007/s11042-017-5577-x

[7] Huang H-C, Cheng TCE, Huang W-F, Teng C-I. (2018). Who are likely to build strong online social networks? The perspectives of relational cohesion theory and personality theory. Computers in Human Behaviour. Vol. 82, pp. 111–123. https://doi.org/10.1016/j.chb.2018.01.004

[8] Bleidorn W, Hopwood CJ (2018). Using machine learning to advance personality assessment and theory. Personality and Social Psychology Review. https://doi.org/10.1177/108886831-8772990

[9] Lo C. G., Maiorana A, Mirisola A, Salerno L, Boca S, Profita G. (2018). Empirically-derived subgroups of Facebook users and their association with personality characteristics: a Latent Class Analysis. Computers in Human Behaviour. Vol. 86, pp.190–198. https://doi.org/10.1016/j.chb.2018.04.044

[10] Shein, P., K. P., & Nyunt, T. T. (2010). Sentiment Classification based on Ontology and SVM Classifier. Second International Conference on Communication Software and Networks.

[11] Ferwerda, B., Schedl, M. & Tkalcic, M. (2016). Personality Traits and the Relationship with (Non-) Disclosure Behavior on Facebook.

[12] Bachrach Y., Kosinski, M., Graepel, T., Kohli, P. & well, D. S. (2012). Personality and Patterns of Facebook Usage.

[13] Xu, B., Gou, L., Xu, A., Mahmud, J. & Cosley, D. Raising Awareness of Conveyed Personality in Social Media Traces.

[14] Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D. & Graepel, T. (2013). Manifestations of user personality in website choice and behaviour on online social networks.

[15] Cai, K., Spangler, S., Chen, Y., Zhang, L. (2008). Leveraging Sentiment Analysis for Topic Detection. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, pp. 265-271.

[16] Raina, P. (2013). Sentiment analysis in News articles using Senti-computing. 1EEE 131h International conference on Data Mining Workshops, pp.959-962.

[17] Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., Tomas (2012). Sentiment Analysis on social media. IEEE/ACM International conference on Advances in Social Networks analysis and mining, pp. 919-926.

[18] Ana C. E.S. Lima, Leandro N.de Castro. (2012). Automatic Sentiment Analysis of twitter Messages. Publisher IEEE, pp. 52-57.

[19] Wang, M., Shi., H. (2010). Research on Sentiment Analysis Technology and Polarity Computation of Sentiment words. Publisher IEEE, pp. 331-334.

[20] Zhu N., Zou, P., Li W., Cheng M. (2012). Sentiment Analysis: A Literature Review. In proceedings of the 2012 IEEE ISMOT, pp. 572-576.

[21] Garg, P., Marwaha, A., Goel, M. B.(2020). Identification and Classification of Reduplication Words in Punjabi Language.International Journal of Scientific & Technology Research Vol. 9, No.6, pp. 532-537.