Optimizing Automatic Personality Prediction

Sarvesh Vishnu[1]

[1]*Department of Computer Science*

[1]sm8969@srmist.edu.in

SRM Institute of Science and Technology
Chennai, India

***Abstract*— Several Artificial Intelligence solutions such as recommender systems, human computer interface operates on the foundations set by psychology, personality models to influence learning methods and approaches to applications. As systems get more personalized and these applications are used more extensively it becomes crucial for personality recognition procedures to be accurate and comprehensive. Therefore, we attempt to create an optimized automatic personality recognition method that accounts for its generalization, interpretability and accuracy. Hence, addressing the existing challenges of personality recognition — privacy, high accuracy.**

***Keywords* — Personality, Generalization, Interpretability, Accuracy, Psychology**

I. INTRODUCTION

Automatic Personality Recognition is a facet of Personality Computing. Several solutions to real-world problems can benefit greatly by leveraging personality theories and psychological characteristics. These methods can be used across various domains to better personalize services, products and information.

The assessment of user's personality using a questionnaire is not possible in certain circumstances. E.g., Analysing an existing anonymous dataset. Moreover, we must also consider self-bias while assessing through a questionnaire. APR is used to solve this dilemma. APR is the process of mapping the data related to a subject to a personality score that represents the personality type of the subject.

II. AUTOMATIC PERSONALITY RECOGNITION

APR schemes are generally divided into three classes. Text-based APR, Multimedia-based APR and behaviour-based APR. In the context of user personality from online social network data, the source data would be, text format from posts or tweets, voice or video data and behavioural patterns of the user respectively for the cases mentioned above.[1]

A. Text Based APR

Text-based APR is inspired by the fact that some language psychology theories claim that the choice of words can reveal some psychological states such as emotions and personality traits of the subject. One of the most prominent techniques for text-based APR is Linguistic Inquiry and Word Count (LIWC).

The procedure categorizes the analysed text into various psychologically relevant sets known as "buckets" like "function words" (e.g., conjunctions, articles, pronouns), 'social processes' (e.g., mate, talk, friend) and 'affective processes' (e.g., happy, nervous, cried). It is then followed by measuring the frequency of words in each of these buckets and predicts the personality of the subject accordingly. Another famous linguistic database is the Medical Research Council (MRC) psycholinguistics database. LIWC & MRC have been proven to achieve acceptable accuracy to detect the user's personality traits from its text.

B. Multimedia-based APR

Multimedia-based APR detects the user's personality traits by analysing related photos or videos and tries to associate features of these data with the facets of personality traits. For example, user who frequently post phots related to art might achieve a high score of openness trait.

C. Behaviour-based APR

The procedure detects the user's personality trait by analysing behavioural patterns and associate them with relevant dominant traits.

D. Psychological Models

Personality refers to the individual differences in characteristics patterns of thinking, feeling and behaving. This construct is commonly used to predict whether a job candidate will perform well in a specific job role and engage well in a prospective cultural environment. There are a variety of models that can be used to assess personality. The model in focus is the "Big Five" traits, also called as the five-factor model or the OCEAN model. This model provides researchers and practitioners with a well-defined taxonomy for selecting job applicants. The core features of the big five are categorized and applied in different cultural contexts. These factors are openness, conscientiousness, extraversion, agreeableness, and neuroticism.

- Openness: the degree to which an individual is imaginative and creative.
- Conscientiousness: the degree to which an individual is organized, thorough, and thoughtful.
- Extraversion: the extent to which an individual is talkative, energetic and assertive.
- Agreeableness: the degree to which an individual is sympathetic, kind and affectionate.
- Neuroticism: reflects the tensions, moodiness, and anxiety an individual may feel.

III. RELATED WORK

Works in text-based methods have been used in several languages utilizing individual linguistic theories from different languages. Silva *et al.* proposed supervised models for APR of text in Brazilian Portugues from Facebook posts — Han et al. used a word embedding technique and prior-knowledge lexicons to automatically construct Chinese semantic lexis for personality analysis. Other work includes deep learning approaches to personality prediction from short texts using combined models using LSTMs and CNNs, done in the "SENTIPEDE" paper by A. Darliansyah, M. A. Naeem, F.Mirza and R.Pears.

Multimedia-based approaches have leveraged computer vision to predict personalities. Kim *et. al.* proposed computer vision techniques to recognize user's personality traits from social media imaged. The study comprised of an online survey of 179 university students and 25,394 photos in total were analysed from Instagram Social media platforms. Ferwerda et al. conducted work on a similar dataset of 22,398 images from Instagram. These studies effectively conclude that there is correlation between features obtained from the pictures and the relevant personality characteristics. Other work in multimedia-based APR includes deep learning approaches. E.g., Zhu *et al.* introduced an end-to-end weakly supervised CNN composed of a classification, regression network.

There has been lesser work done in Behaviour-based APR in comparison to the other two approaches. Tadesse *et al.* Analysed and compared four machine learning models to investigate the relationship between user book preferences by analysing labels generated by users from the online website Goodreads.com and mapped it to personality scores from Facebook users. Nave *et al*. Investigated the possibility of personality prediction using musical preferences. The study was able to infer that, reactions to unfamiliar musical excerpts predicted individual differences in personality.

It can be observed that research in the field of Automatic Personality Recognition has been varied and spread across different domains and applications. However, one essential question we must answer is the generalization of these methodologies across other data groups, subjects. Several methods have been adopted in the works mentioned above which target various individual facets through approaches such as DNNs and ensemble methods.

A methodology that addresses the essential limitations of the previous works. i.e., Interpretability, Generalization and taken into consideration. Additionally, a model that qualifies relevance, accuracy across non-conventional domains and in cross-domain settings would be an ideal approach.

	S.no	Title	Journal Published	Methodology	Year
Text-based APR	1	Learning Personality Traits from Facebook text	IEEE Latin America Transactions	Proposes supervised models for APR of text in Brazilian Portuguese from Facebook Posts	2018
	2	Knowledge of words	Knowledge based systems	Uses word embedding techniques and prior-knowledge lexicons suitable for personality analysis.	2020
	3	Group Level Personality Detection	Springer	Proposes a model of group-level personality detection by learning the influence of text generated networks.	2018
	4	SENTIPEDE: A smart system for Sentiment-based Personality detection from short texts	Journal of Universal Computer Science	Leverages Neural network language model for personality detection from short texts by using a unified model that combines LSTM with a CNN.	2019
	5	Computing personality recognition from Facebook	Review of Hypermedia and Multimedia	Discusses the effectiveness of using psycholinguistic knowledge in APR & performed series of individual experiments of APR from Facebook text.	2020
Multimedia -based APR	6	Using computer vision techniques on Instagram to link users' personalities and genders to the features of their photos	Journal of Information Processing & Management	An online survey of 179 university students was conducted to measure user characteristics and 25,394 photos in total were downloaded and analyzed from the respondents' Instagram accounts.	2018
	7	Inferring Personality Traits from Attentive Regions of User Liked Images Via Weakly Supervised Dual Convolutional Network	Springer	Introduces an end-to-end weakly-supervised dual CNN for personality detection composed of a classification network and a regression network. The classification network detects personality class-specific attentive image regions. While the regression network is used for detecting personality traits.	2020
	8	PersEmoN	IEEE Transactions on Affective Computing	Proposed PersEmoN, an end-to-end trainable and deep Siamese-like-network, PersEmoN is composed of two CNN branches, the first for emotion and the second for personality traits. Both networks share their bottom feature extraction module and are optimized within a multi-task learning framework	2020
Behaviour-based APR	9	Personality Prediction Based on User Behavior on the Facebook Social Media Platform	IEEE Access	Analyzes and compared four machine learning models to investigate the relationship between user behavior on Facebook and big-five personality traits.	2018
	10	Personality Prediction Based on User Behavior on Facebook	IEEE Access	Investigates the relationship between the user book preferences by analyzing labels "tags" generated by users from Goodreads.com and match it with personality scores collected from Facebook users.	2018
	11	Musical Preferences Predict Personality: Evidence from Active Listening and Facebook Likes	Journal of Psychological Science	Investigated the possibility of personality prediction using musical preferences. Their finding using data of active listening and Facebook likes show that reactions to unfamiliar music excerpts predicted individual differences in personality.	2018

IV. Conclusion & Future Work

The paper fundamentally addresses the ggeneralization of the prediction over different domains. The Personality Recognition models and algorithms have been tested over fields such as social media data, movies, music etc. However, it is questionable if the personality prediction is relevant to other domains. i.e., Effectiveness of the prediction in a cross-domain setting. The paper incorporates interpretability in the models to better aid in communicating the working of the model, alleviating bias and reducing privacy related concerns associated with providing personalized predictions based on personality.

As technologies get more personalized, we must work towards improving the accuracy and comprehensive nature of the fundamental operations supporting these advancements. In this context, the automatic personality recognition model. Data sensitivity is an increasing issue and must be addressed further in the data sources used. Decentralizing the data is certainly a viable option. Open research, practical areas include testing out various other environments of applications, eliminating all biases and providing an even more accurate, inclusive system.

V. References

[1] N. Annalyn, M. W. Bos, L. Sigal, and B. Li, "Predicting Personality from Book Preferences with User-Generated Content Labels," IEEE Transactions on Affective Computing

[2] S. Han, H. Huang, and Y. Tang, "Knowledge of words: An interpretable approach for personality recognition from social media," Knowledge Based Systems

[3] B. Silva and I. Paraboni, "Learning Personality Traits from Facebook Text," IEEE Latin America Transactions

[4] L. Li, H. Zhu, S. Zhao, G. Ding, and W. Lin, "Personality-Assisted Multi-Task Learning for Generic and Personalized Image Aesthetics Assessment," IEEE Transactions on Image Processing

[5] X. Sun, B. Liu, Q. Meng, J. Cao, J. Luo, and H. Yin, "Group-level personality detection based on text generated networks," World wide web

[6] A. Darliansyah, M. A. Naeem, F. Mirza, and R. Pears, "SENTIPEDE: A Smart System for Sentiment-based Personality Detection from Short Texts," Journal of Universal Computer Science

[7] A Survey on Personality-Aware Recommendation Systems. Sahraoui Dhelim, Nyothiri Aung, Mohammed Amine Bouras, Huansheng Ning and Erik Cambria | Artificial Intelligence Review

[8] W. R. dos Santos, R. M. S. Ramos, and I. Paraboni, "Computational personality recognition from Facebook text: psycholinguistic features, words and facets," New Review of Hypermedia and Multimedia

[9] L. Zhang, S. Peng, and S. Winkler, "PersEmoN: A Deep Network for Joint Analysis of Apparent Personality, Emotion and Their Relationship," IEEE Transactions on Affective Computing

[10] Kim and J. H. Kim, "Using computer vision techniques on Instagram to link users' personalities and genders to the features of their photos: An exploratory study," Information Processing & ManagementJ.

[11] C. Segalin, A. Perina, M. Cristani, and A. Vinciarelli, "The Pictures We Like Are Our Image: Continuous Mapping of Favorite Pictures into Self-Assessed and Attributed Personality Traits," IEEE Transactions