Personality prediction via social media user using recurrent neural network

[1]Sakshi Shinde, [2]Sakshi Kumbhar, [3]Atharva Kulkarni, [4]Ajay Jagtap, [5]Prof.pravin Kamble

[1,2,3,4,5]Information Technology BE, Trinity College of Engineering and Research, Pune, India

Abstract— **People differ from one another, which makes personality vital. Research on personality prediction is still ongoing in this area. The use of social data from the media for personality prediction shows promise because it eliminates the need for user-filled surveys, reducing time and boosting trust. The daily usage of social media has surged. The internet is constantly receiving new content and images in large quantities. These days, the work mostly uses decision trees and Naive Bayes machine learning on the popular Facebook dataset.**

Index Terms— **Machine Learning, Facebook, Personality Sdetection, sentiment analysis.**

I. INTRODUCTION

Social interactions, communication, and cooperation among people alter as a result of social networks' growing popularity. On the one hand, users can communicate with friends, leave comments, talk about public issues, and more using the network platform. However, social networks are becoming more and more important in day-to-day life, and they have even had a significant impact on the reconstruction of the network of genuine social relationships.

In the meantime, social computing has grown in importance as a research topic in the fields of computer science and information technology because social network behavior and status are readily observed, collected, and analyzed. Because personality and behavior are closely related, personality prediction has both broad academic applications and business opportunities.

As a result, businesses and academics are now concentrating on using social networking to access users' social information. A person's personality is a highly generalized combination of their unique individual traits. This is why, even in the same setting, different people will behave differently from one another. One of the subfields of

psychology called personality psychology focuses on identifying and analyzing people's intrinsic traits primarily from their outward behavior. Personality traits are typically used in psychology to characterize individuals and provide an explanation for their behavior and preferences.

II. LITERATURE REVIEW

The personality types of the Big Five Factor model can be quantified from user profile images. To measure the effectiveness, proposed two models using convolution Neural Networks to classify each personality of the user. Done performance analysis among two different models for efficiently predict personality traits from profile image[1].

This paper introduces a new prediction using multi model deep learning architecture combined with multiple pre-trained language model such as BERT, RoBERTa, and XLNet as features extraction method on social media data sources. Finally, the system takes the decision based on model averaging to make prediction[2].

Personality prediction has gained lot of focus nowadays. It studies behavior of users and reflects the thinking, feelings etc. Traditional ways to take survey that was time consuming, for large number of users there is need of automatic prediction. Users are dynamic and can have their account on multiple

platforms as they can have multi-context information[3].

In this paper, we have tried to review the work carried out for personality prediction of social media users in the past decade using the information extracted from their digital footprints. Further, we have also discussed different machine and deep algorithms, datasets, personality measures, and applications of automatic personality detection[4].

There is a growing number of research papers related to a user's behaviour in social networks that has recently attracted more attention in the international research community. Personality recognition is studied by two main disciplines: computational linguistics and Social Network Analysis. From the area of computational linguistics[5].

Pennebaker and King 1999 [6] wrote a pioneering work dedicated to personality extraction from text. They examined words in a variety of domains such as diaries, college writing assignments and social psychology manuscripts to study personality related features with linguistic cues. Their results show that agreeable people tend to use more articles while introverts and those low in conscientiousness use more words signalling distinctions. Neurotics use more negative emotion words.

Argamon et al. 2005 [7] classified neuroticism and extraversion using linguistic features such as function words, judgemental and appraisal expressions and modal verbs. Their results revealed that neuroticism is related to the use of functional lexical features, for instance appraisal lexical taxonomy, whereas the results for extraversion were less clear.

Other studies linked neuroticism to irrational beliefs or poor coping efforts on well-being personality [8] oberlander and Nowson 2006 classified the extraversion, stability, agreeableness and conscientiousness of bloggers using the Naive Bayes prediction model as a learning algorithm using different sets of n-grams as features.

Mairesse et al. 2007 [9] examined correlations between the Big 5 personality traits, using LIWC and RMC as feature sets. While LIWC features included word classification such as positive emotions or anger, RMC features included results about word age of acquisition or word imageability.

IV. PROPOSED SYSTEM

We extract social data and questionnaire, and focus on how to use the user text information to predict their personality characteristics. We use the correlation analysis and principal component analysis to select the user information, and then use the machine learning classifiers, the prediction model and the multitasking model to predict and analyse the results.
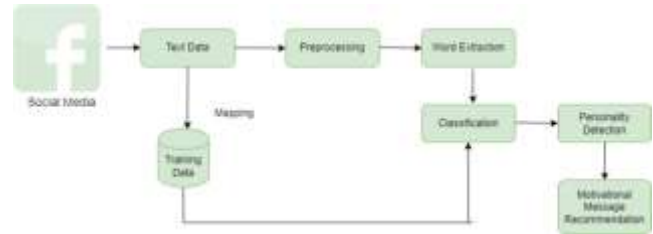

Fig. System Architecture
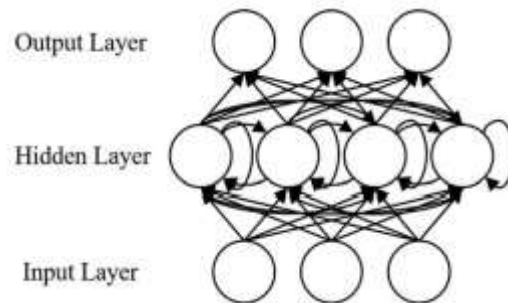
1. RNN (Recurrent neural network)

Fig. Recurrent neural network

As shown in Fig. , for a RNN, let our input x be a sequence whose length is T, x = {x1 , x2 ,...,xT }, and each item xt is a feature vector. At time step t, given the previous hidden layer state ht−1 , the current hidden layer state ht and the output layer state yt can be calculated by,

$$h_t = \sigma_h(w_h x_t + U_h h_{t-1} + b_h)$$

$$y_t = \sigma_y(w_y h_t + b_y)$$

where Wh and Wy denote the input-to-hidden and hidden-to-output weight matrices, respectively, Uh is the matrix of the recurrent weights between the hidden layer and itself at two adjacent time steps, bh and by are the biases, and σh and σy denote the activation functions.

At each time step, the input is propagated in a standard feedforward fashion, and then, a learning rule is applied. The back connections lead to the result that the context units always maintain a copy of the previous values of the hidden units (sincethey propagate over the connections before the learning rule is applied). Thus, the network can maintain a state, allowing it to perform such tasks as sequence prediction that are beyond the power of standard multilayer perception.

Formula for calculating current state:

$$h_t = \int (h_{t-1}, x_t)$$

Where,

h_t-> Current state

h_{t-1} -> Previous state

x_t-> Input state

Formula for applying Activation function:

$$h_t = \text{activation}(W_{hh}h_{t-1} + w_{xh}x_t)$$

Where,

W_{hh}->Weight at recurrent neuron
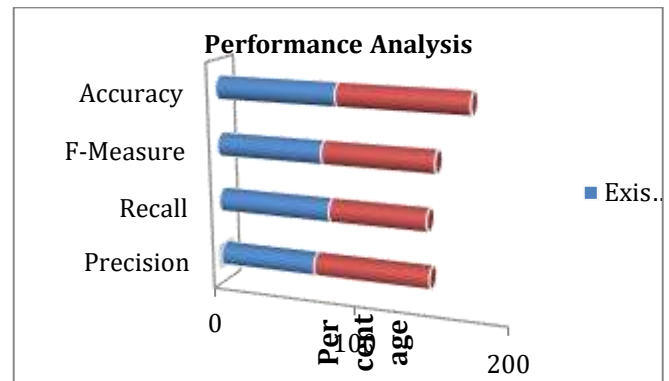
w_{xh}->Weight at input neuron

Formula for calculating output:

$$y_t = w_{hy}h_t$$

y_t->Output

w_{hy}->Weight at output layer

V. RESULTS AND DISSCUSSION

Performance Analysis

	Existing System	Proposed System
Precision	67.78	78.70
Recall	79844	65.64
F-Measure	73.11	74.31
Accuracy	83.29	87.26

V. CONCLUSION

We present an overview of research directions in personality psychology and social networks in this paper. By examining the connection between users' personalities and their actions on social networks, the study looks at the literature on the application of social media frameworks as behavioral feature studies. We performed a comparative analysis of the top behavioral indicators for Facebook usage of the same set of features to capture the ways in which users interact, communicate, and form connections with one another in order to forecast a user's personality.

VII. REFERENCES

[1] T.S. Kanchana and B.S.E.Zoraida A Deep Learning Model for Predicting User Personality Using Social Media Profile Images IJCSNS International Journal of Computer Science and Network Security, VOL.22 No.11, November 2022

[2] Hans Christian1 , Derwin Suhartono2* , Andry Chowanda2 and Kamal Z. Zamli Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging 2021

[3] Hetal Vora, Mamta Bhamare, Dr. K. Ashok Kumar, Personality Prediction from Social Media Text: An Overview International Journal of Engineering Research & Technology (IJERT) 2020.

[4] Eishita Sharmaa , Rhea Mahajanb , Remia Mahajanc , Vibhakar Mansotrad Automated Personality Prediction of Social Media Users: A Decade Review 2020

[5] M. S. H. Mukta, M. E. Ali, and J. Mahmud, "User generated vs. Supported contents: Which one can better predict basic human values?" in Proc. Int. Conf. Social Inform. Cham, Switzerland: Springer, 2016, pp. 454–470.

[6] P. Howlader, K. K. Pal, A. Cuzzocrea, and S. D. M. Kumar, "Predicting Facebook-users' personality based on status and linguistic features via flexible regression analysis techniques," in Proc. 33rd Annu. ACM Symp. Appl. Comput. (SAC), Pau, France, Apr. 2018, pp. 339–345.

[7] KAUSHAL AND M. PATWARDHAN, "EMERGING TRENDS IN PERSONALITY IDENTIFICATION USING ONLINE SOCIAL NETWORKS—A LITERATURE SURVEY," ACM TRANS. KNOWL. DISCOV. DATAT, VOL. 12, NO. 2, PP. 1–30, JAN. 2018.

[8] W. Youyou, M. Kosinski, and D. Stillwell, "Computer-based personality judgments are more accurate than those made by humans," Proc. Nat. Acad. Sci. USA, vol. 112, no. 4, pp. 1036–1040, Jan. 2015.

[9] G. Farnadi, G. Sitaraman, S. Sushmita, F. Celli, M. Kosinski, D. Stillwell, S. Davalos, M.-F. Moens, and M. De Cock, "Computational personality recognition in social media," User Model. User-Adapted Interact., vol. 26, nos. 2–3, pp. 109–142, 2016.

[10] S. Mallik, T. Bhadra, and U. Maulik, "Identifying epigenetic biomarkers using maximal relevance and minimal redundancy based feature selection for multi-omics data," IEEE Trans. Nanobiosci., vol. 16, no. 1, pp. 3–10, Jan. 2017, doi: 10.1109/tnb.2017.2650217.

[11] T. Tandera, D. Suhartono, R. Wongso, Y. L. Prasetio et al., "Personality prediction system from facebook users," Procedia Computer Science, vol. 116, pp. 604–611, 2017.

[12] ONG, A. D. RAHMANTO, D. SUHARTONO, A. E. NUGROHO, E. W. ANDANGSARI, M. N. SUPRAYOGI ET AL., "PERSONALITY PREDICTION

BASED ON TWITTER INFORMATION IN BAHASAINDONESIA," IN COMPUTER SCIENCE AND INFORMATION SYSTEMS (FEDCSIS), 2017 FEDERATED CONFERENCE ON. IEEE, 2017, PP. 367–372.

[13] F. Iatan, "Predicting human personality from social media using a f uzzy neural network," in Issues in the Use of Neural Networks in Information Retrieval. Springer, 2017, pp. 81–105

[14] T. Bhadra, S. Mallik, and S. Bandyopadhyay, "Identification of multiview gene modules using mutual information-based hypograph mining," IEEE Trans. Syst., Man, Cybern., Syst., vol. 49, no. 6, pp. 1119–1130, Jun. 2019, doi: 10.1109/tsmc.2017.2726553.