IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 9, Issue.4, April 2021 E-ISSN: 2347-2693

Predicting Personality from Micro-Blogs using Supervised Machine Learning Models

P. Chandra Shaker Reddy[1*], Yadala Sucharitha[2], G. Surya Narayana[3]

[1]Department of CSE, CMR College of Engineering & Technology, Hyderabad, Telangana, India
[2]Department of CSE, CMR Institute of Technology, Hyderabad, Telangana, India
[3]Department of CSE, Vardhaman College of Engineering, Hyderabad, Telangana, India

[*]*Corresponding Author: chandu.pundru@gmail.com, Tel.: +91-9703763343*

DOI: https://doi.org/10.26438/ijcse/v9i4.714 | Available online at: www.ijcseonline.org

Received: 03/Apr/2021, Accepted: 10/Apr/2021, Published: 30/Apr/2021

***Abstract*—** Social media is a place where users present themselves to the world, revealing personal details and insights into their lives. We are beginning to understand how some of this information can be utilized to improve the users' experiences with interfaces and with one another. In this paper, we are interested in the personality of users. Personality has been shown to be relevant to many types of interactions; it has been shown to be useful in predicting job satisfaction, professional and romantic relationship success, and even preference for different interfaces. Until now, to accurately gauge users' personalities, they needed to take a personality test. This made it impractical to use personality analysis in many social media domains. In this paper, we present a method by which a user's personality can be accurately predicted through the publicly available information on their Twitter profile. We will describe the type of data collected, our methods of analysis, and the results of predicting personality traits through machine learning. We then discuss the implications this has for social media design, interface design, and broader domains.

***Keywords*—** personality, user profiles, personalization, cross domains and Twitter

I. INTRODUCTION

Social media on the web has become drastically finished the most recent decade. Social systems have turned out to be generally utilized and prevalent mediums for information dissemination and for social collaboration's. Client activities on social networking give important understanding into singular behavior, encounters, opinions and interests. It is very comparable how a man connects socially with the human instinct and behavior [1]. Personality is the most many-sided human trait and it additionally depicts the uniqueness of a man. Personality is one of the basic angles, by which we can comprehend behavioral identities. It's a long haul objective for clinicians to comprehend personality of a human and its effect on a person behavior. This way of Behaving includes collaboration among a man's hidden personality attributes [2].

The circumstance, that a man gets himself or herself in, assumes a noteworthy part on his or her response. Be that as it may, in a large portion of the cases, individuals react as for their hidden personality qualities. It conceivable to get to and dissect a lot of examples with a specific end to goal consequently recognize personality sorts of author and foresee potential responses and behaviors. People tend to comprehend others' behavior on the premise of the perception of their ordinary behavior. Tremendous number of analysts around the word has been pulled in to deal with this exploration area from various fields particularly scientists in computational phonetics, brain science, counterfeit consciousness, common dialect preparing, human-machine connection, behavioral investigation, and machine learning [3].

Utilization of social networking s has been expanded exponentially as of late. A study of social networking s assessed roughly 115 million individuals over every one of the sites on the web in January 2005, and soon after five years Twitter alone has surpassed 500 million individuals [4]. A standout amongst the most all inclusive on-line conditions, Twitter, is turning into an undeniably every day action of people groups' far and wide. Twitter profiles turned into an essential wellspring of information used to shape impressions about others. During the time spent making users uncover social networking profiles, a considerable measure about themselves both in what they offer and how they say it. Through self-depiction, announcements, photographs, quite a bit and interests, a man's personality turns out using their profile. For instance, individuals look at other individuals' profiles of Twitter when attempting to choose whether to begin dating them, and they are additionally utilized while evaluating work applicants [5]. A man's personality can be defined by their behavior and character in various situations. A person's choice can be influenced by personality in various things such as websites, novels, songs and movies. In Addition, it also affects the interaction with others.

Employee recruitment and other counseling sessions will be done based on personality assessment. A person should know about their personality by conducting different personality tests [6].

A Personality test is like observation of psychologists or self-descriptive report. These are costly traditional methods and less practical. Recently, the personality test is conducting using a website like online questionnaire and this is quite practical. In this type of test the user has to answer various questions. A recent study saying that from the text written by the user, we can obtain the personality traits [7]. The choice of most frequently used words can describe the personality of that particular person. Social media is a platform where represent users to the world by them. Personality classifies a human that affects preferences of an individual and interaction. To find out their personality, people are required to take a personality test. Social media became a platform for people to express their feelings to the world, based on the posts in social media made by users can be used to analyze their behavior [8]. This experiment uses text written by Twitter users to predict personality of the particular user. The used are English and Indonesian. Relations can be presented in among user preferences in different entertainment domains and personality types. The entertainment domains namely movies, music, TV shows and books. For analyzing we have taken a total of 53,224 Twitter user profiles. These profiles were composed of both personality scores from the Five Factor model (openness, conscientiousness, extraversion, agreeableness, neuroticism) and explicit interests in each of the above domains [9]. From our analysis, we extract different personality stereotypes, similarity index of the personality with the personality of different domain and association rules for some of the domains.

Personality is most important and complex attribute among other four attributes like behavioral, emotional, temperamental and mental [10]. This personality describes a unique individual. We know that words used by a person define the information about the speaker and their semantic content. This type of data contains cues of speaker's personality traits. The Big Five can be used to assess the personality and shown in Figure 1.

- Introversion vs. Extraversion (reserved, aloof, shy vs. playful, sociable, assertive)
- Disagreeable vs. Agreeableness (faultfinding vs. cooperative)
- Neuroticism vs. Emotional stability (insecure vs. unemotional)
- Unconscientiously vs. Conscientiousness (careless vs. self-disciplined)
- Openness to experience (unimaginative)

Figure 1: Personality Traits

From the previous analysis 14 out of 60 items can be replaced in the NEO Five Factor Inventory. For the new analysis in high school (N=1960) and adult (N=1493) samples have been taken from the remaining set of revised NEO Personality Inventory items. The results show the reliability of personality traits. These accurate results are appropriate for the age 16 and up in responders [11]. Although, continued use of the present tool is also fair and sensible for many applications. The NEO Five Factor Inventory was not considered to supply definitive amount of the 5 personality factors. Alternatively, it was designed for investigative research as a brief tool which is reasonable for personality factors [12]. This is used in most applications. Rest of the article is organized as follows: Related works are discussed in section 2, the proposed methodology is explained in section 3, section 4 presented a experimental results and finally section 5 presents a conclusion.

II. RELATED WORK

In author proposed another approach for personality discovery which depends on joining the opinion, full of feeling and presence of mind learning from the utilizing assets viz. Sentic-Net, Concept-Net, Emo-Sentic-Net and Emo-Sentic-Space [13]. In their approach, they consolidated good judgment information based highlights with psycho-phonetic highlights and recurrence based highlights and later the highlights were utilized in directed classifiers. Further, they created five help vector machine models for five personality attributes. Their exploratory outcomes demonstrate that the utilization of presence of mind learning with emotional and feeling information upgrades the exactness of the current systems which utilize just psycho-phonetic highlights and recurrence based investigation at lexical level [14]. In writers display a programmed personality attribute acknowledgment show in light of social system (Face book) utilizing users' status. They utilized machine learning calculations viz. Bolster

Vector Machine, Bayesian Logistic Regression (BLR) and Multinomial Naïve Bayes (MNB). In author created three machine learning calculations i.e. bolster vector machine, nearest neighbor with k=1 (kNN) and Naïve Bayes for deducing the personality attributes of users on the premise of their face book refreshes. A few characterization procedures were utilized to manufacture prescient personality models along the five personality measurements utilizing the etymological highlights of a dataset involved couple of thousand s requested from basic brain science understudies. Authors in fabricated personality acknowledgment show in both discussion and by means of Big. They abused two lexical assets as highlights, LIWC and MRC, and anticipated both personality scores and classes utilizing Support Vector Machines (SVMs) and M5 trees individually. They likewise detailed an extensive rundown of connections between Big5 personality characteristics and two lexical assets they utilized [15].

The Linguistic Inquiry and Word Count – LIWC (http://www.liwc.net) was utilized as an apparatus for etymological examination. In author took after the work exhibited indirect and built up personality recognition demonstrate for present day Greek with etymological highlights (like Part-of-Speech labels) and mental highlights (like in LIWC). They utilized SVM classifier for building the machine learning model, they exhibited that personality and dialect can be effectively ported from English to different dialects. In author utilized n-gram highlights from width of online journals a corpus of individual web-s for displaying four out of five personality measurements [16]. They manufactured their model with SMO and Naïve Bayes machine learning techniques. Their outcomes call attention to the significance of the component choice in expanding the classifiers exactness yielding 83%-93% for programmed includes choice. The relationship between users' personality and social system action has been the concentration of a few investigations in the last. In author removed word n-grams as highlights from a vast corpus of online journals with various element vector development settings, for example, the nearness/nonappearance of stop words or converse record recurrence. They found that bigrams, regarded as Boolean highlights and keeping stop words, yield great outcomes utilizing SVMs as learning calculation. Golbecketal proposed a model to anticipate personality from Facebook profile with semantic, (for example, word tally) and social system highlights (like companions check) information utilizing machine learning calculations [17]. They anticipated personality scores of 279 Facebook users, misusing both semantic highlights (from LIWC) and social highlights (i.e. companion tally, relationship status). In author additionally anticipated the personality of 279 Twitter users with the assistance of LIWC, auxiliary highlights (i.e. hastags, connections) and assumption highlights, and utilizing a Gaussian Process (GP) as learning calculation. Tomlinson et al. contemplated the Conscientiousness quality to recognize objective, inspiration, and the way the creator sees control over the depicted circumstances. They played out the examination of occasion structures of printed client announcements in a Facebook dataset In writers utilized system highlights (like adherents, following, and so on.) to assemble M5 rules based learning model for the expectation of personality scores of 335 Twitter users [18].

Authors in introduced a broad investigation of the system qualities (i.e. for example, size of companionship organize, transferred photographs, occasions went to, times client has been labeled in photographs) that relate with personality of 180000 Facebook users. They anticipated personality scores utilizing multivariate direct relapse (mLR), and detailed great outcomes on extraversion [19]. Ross et al. spearheaded the investigation of the connection amongst personality and examples of social system utilize. They guessed numerous connections amongst personality and Facebook highlights, including (1) constructive connection amongst Extraversion and Facebook utilize, number of Facebook companions and relationship with Facebook gatherings; (2) constructive connection amongst Neuroticism and uncovering private information on Facebook; (3) constructive relationship amongst Agreeableness and number of Facebook companions; (4) constructive connection amongst Openness and number of various Facebook highlights utilized; (5) adverse connection amongst Conscientiousness and general utilization of Facebook. In author built up a machine learning model for neuroticism and extraversion utilizing semantic highlights, for example, work words, deictic, evaluation articulations and modular verbs. In author utilized different feeling dictionaries like NRC hash label feeling vocabulary and NRC feeling dictionary for the personality identification and discovered key change in the exactness of the PRT framework [20].

III. PROPOSED METHODOLOGY

The proposed system can gather set of tweets from persons. After that this text can be processed into vector data. Stratification method will categorize user’s text into a considered data set. The test results were predictions for each and every Big Five Personality traits [21]. The primary and secondary personality characteristics were gained from the amalgamation of two traits. The system proposed was a web application.

3.1 Personality Detection Model

Information mining procedures assume an essential part in removing connection designs amongst personality and assortment of user’s information caught from different sources. For the most part, two methodologies were embraced for examining personality qualities of social system users. The primary approach utilizes an assortment of machine learning in view of social calculations to fabricate models system activities just. The second one expands the personality-related highlights with etymological prompts [22].

This intrigue is because of the way personality recognition that is likewise extremely valuable in social system investigation and supposition mining that is huge and creating fields of research. Online social systems are composed of immense archives information which is appropriate for personality acknowledgment; still, there are a few issues in utilizing them for building such models. (1) Social system information is for the most part not freely accessible, (2) if information is unlabeled, (3) it is extremely hard to comment on with personality judgments and (4) Generally, it is in a variety of dialects. Concept diagram of proposed model is shown in figure 2.

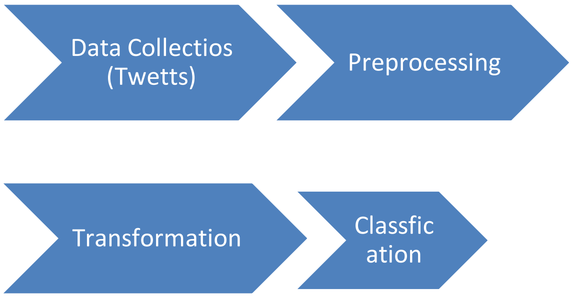

Figure 2: User Vs Personality Trait

1) Data Collection: For exhibiting the framework, we require tweets posted by an individual(s). For this, tweets are acquired utilizing Twitter API. Twitter API gives an entrance to twitter information containing information about the users, tweets distributed by a client, list items on twitter and so on. Tweet protest is in .json organized.

2) Pre-handling module: The tweets are first gotten from the tweet protest. The framework at that point removes meta-traits from the tweets. The information separated can be isolated into social behavior and syntactic information.

3) Transformation module: This module changes the "multi-mark issue into twofold arrangement issues". This module gets the Meta attributes removed from the past module. Utilizing this information, it develops an element vector. Each position in the vector compares to a meta-quality.

4) Classification module: A Multilayer Perceptron (MLP) Neural Network is utilized for characterization. There are five neural systems (classifiers), one for every personality characteristic. The yield of every classifier is either '1' (yes) or '0' (no) contingent upon whether the vectors coordinates or not, further surmising that the individual has the personality quality or number.

In this segment, we propose our system after significant investigation of past work done in this field and bits of knowledge in light of writing review of particular research papers in setting to personality forecast from Twitter, which includes a personality expectation show in view of Logistic Regression Classifier. Since, personality expectation is an arrangement issue, our model advances another approach of anticipating one's personality utilizing Logistic Regression calculation with a limited blunder work utilizing stochastic inclination drop which tends to specific impediments that ran over to specialists already [23]. The usage of our model is depicted in a diagrammatic flowchart of different advances performed in this examination. This approach conveys the possibility to defeat the disadvantages of past work done. We have additionally assessed our model with various assessment measurements, for example, exactness, review and precision and furthermore contrasted our outcomes and Naïve Bayes Classifier utilized preceding foresee user's personality by mapping phonetic highlights of tweets into various personality classes to which a client might possibly follow and architecture of the proposed model shown in figure 3.

Structural Features

Through a Twitter application, we are able to collect information about the user's egocentric network. We first obtained a list of friends. We were interested in density, and Twitter provides some information about links between a user's friends. A separate query must be made for each pair of users to determine if they are or are not friends. It was not possible to submit a query for each pair of friends because the Twitter application would timeout; Twitter limits the time an application can run, and since each query is sent over the network, performance becomes an issue [24]. Thus, we sampled 2,000 unique pairs of friends from a user's egocentric network and used that to determine the density of the network, i.e. what percentage of possible edges between friends exist.

Personal Information

Users provide a wealth of personal information. We collected everything available, even though some features would turn out to have no use in our analysis. The raw data included features like the user's name, birthday, relationship status, religion, education history, gender, and hometown. Most of this information was not required, so some users did not in clude all information. Where possible, we created additional features that indicated whether or not the user had included the information (e.g. was a religion or hometown provided or not), or how many items were listed (e.g. how many educational experiences were listed). These added features turned out to be much more useful and predictive than the original raw data. For example, from 279 users, 111 listed a religion. Within those 111 people were 82 different entries. This creates a space too sparse to do any statistical analysis, but just knowing if a person listed a religion or not reveals insights into what they are willing to share.

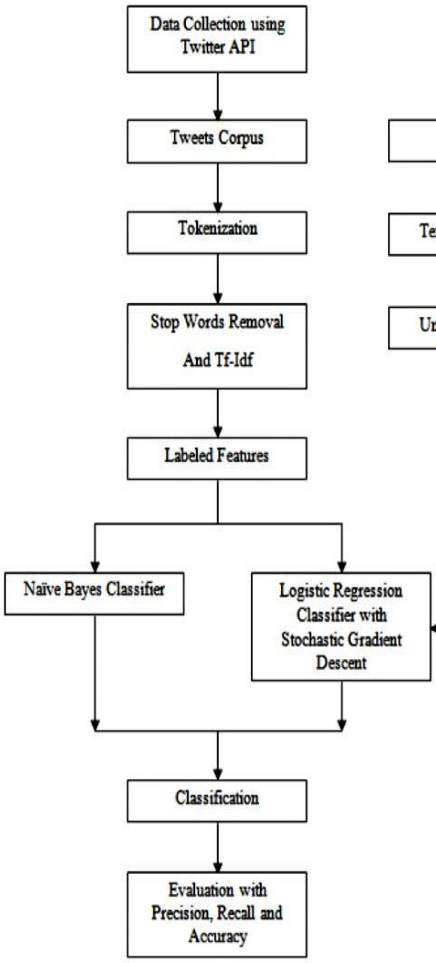

Figure 3: Proposed Architecture Diagram

Activities and Preferences

Providing lists of personal activities or favorite things has always been a part of Twitter. Users list favorite TV shows, movies, music, book, quotes, as well as political and organizational affiliations and favorite activities. As was the case with religion described above, the space is far to sparse to do any analysis over the actual entries in these fields, so we created more companion measures. For lists of favorite things and activities, we counted the number of characters in the entry, roughly measuring how much information the user provided in each field. This included values of 0 for users who did not supply any information. For organizational affiliations, we counted the number listed and for political affiliations, we simply measured whether it was shared or not [25].

Language Features

Similar to the activities and likes described above, users also have opportunities to share more personal written information through the ”About Me” and ”blurb” text in their profiles, and through status updates. We collected these entries and also added features to measure the character length of each entry. Previous research has shown that linguistic features can be used to predict personality traits. Since there is text available on users' Twitter profiles, there is potential to apply these linguistic analysis methods to help predict personality. However, the text samples used in earlier studies are much larger than are available to us through Twitter. Data collected in was used in both studies mentioned above. They had three separate sources of text, ranging from an average of 1,770 words to over 5,000 words per person. We are in a much different position pulling text from Twitter profiles. We combined status updates, the ”About Me” text, and the ”blurb” text from the profile into a single string so there would be enough text to analyze. The average number of words in these three strings combined was 26.6. Many users did not have enough words for linguistic analysis. Fifty-four subjects had no text at all in these three fields, and 112 had fewer than 10 words. We eliminated these users to the text analysis statistics would not be too noisy. This left us with 167 subjects with an average of 42.6 words per person. Note that the elimination of subjects with too little text should have a limited impact on our results. For each personality factor, a two-tailed Student's t-test showed no significant difference in the personality score between users with 10 words or more and users with fewer than 10. Following the methods used in as well as other studies of Twitter behavior, such as, we utilized the Linguistic Inquiry and Word Count (LIWC) tool to analyze the text. LIWC produces statistics on 81 different features of text in five categories. These include Standard Counts (word count, words longer than six letters, number of prepositions, etc.), Psychological Processes (emotional, cognitive, sensory, and social processes), Relativity (words about time, the past, the future), Personal Concerns (such as occupation, financial issues, health), and Other dimensions [26].

Internal Twitter Statistics

A number of features were available that described a user's experience, settings, and history with Twitter. This included the user ID, an integer value that corresponds to when the user joined the network (lower values indicate an earlier join time), the unix timestamp of their last profile update, the number of notes (short messages) posted, and other features that proved less useful such as the URL of the profile picture, whether or not their profile was blocked (no one's was), whether the person was an app user (everyone was), and if they had provided a status update (everyone had). Personality and Profile Correlations We had 279 subjects who completed the personality inventory, but we only used data from the 167 subjects who had at least 10 words among all of their text fields so we could perform a linguistic analysis. Demographic information was pulled from their Twitter profiles. Among these subjects, the average age was 31.2 years (std dev 8.7). Of those reporting gender, 68 were female and 61 were male (38 did not report). In terms of location, 82.6% (138) were from the United States with the remaining subjects coming from India (8), Australia (7), Italy (7), and others (7). We found many weak correlations between users' profile features and personality scores. This echos previous results of linguistic analysis and personality found in [27]. These are reported in table and statistically significant correlations ($p < 0.05$) are bolded. Below, we discuss some of the more interesting relationships.

Predicting Personality

Our feature set for each user included all meaningful features. We excluded those which could not be quantified (e.g. picture URL), for which the value was the same for all users (e.g. if their profile was blocked), or where the data was so sparse that it would not be predictive (e.g. personal website URL). Where possible, we included our companion statistics on these features (e.g. while the actual website URL was not used, a feature indicating presence or absence of the URL was included). Linguistic features were included as described above. We also added five additional features. We ran a multiple linear regression analysis for each personality factor, producing a vector of weights for each feature. The dot product of the weight vector and the feature vector was computed for each user and for each personality feature to create five composite features. In total, we had 74 features per user. To predict the score of a given personality feature, we performed a regression analysis in Weka with a 10-fold cross-validation with 10 iterations using two algorithms: M5′Rules, a rule-based variation of the M5′ algorithm [28], and Gaussian Processes.

Discussions

The question that arises from this research is how the results can be used. Drawing on research results that connect personality type to behavior and preferences, there is potential to integrate previous personality results into social media as a way to enhance the accuracy of certain features or the user's experience. Research on interface preference and personality type showed that users preferred interfaces designed to represent personalities that most closely matched their own. This has significant implications for this work. With the ability to infer a user's personality, social media websites, e-commerce retailers, and even ad servers can be tailored to reflect the user's personality traits and present information such that users will be most receptive to it. For example, the presentation of Twitter ads could be adjusted based on the personality of the user. Similarly, product reviews from authors with personality traits similar to the user could be highlighted to increase trust and perceived usefulness by the user. Customized website "skins" could be created for different user personality types, as suggested.

Our methods provide a straightforward way to obtain personality profiles of users without the burden of tests, and this will make it much easier to create personality-oriented interfaces. This same idea can be extended even further to advertising. While results of integrating personality to marketing have been mixed, some work has demonstrated connections between marketing techniques and consumer personality. For e-commerce marketers, both those who advertise on Twitter and elsewhere, utilizing social media profiles as a way to determine consumer personality can make it easy to implement existing techniques that benefit from this knowledge of consumer background. Recommender systems may also benefit from integrating predicted personality values. Results showing correlations between personality and music taste are well established in the literature. Inferring personality traits from Twitter profiles may allow recommender systems to improve their accuracy by recommending music, and possibly other items, that are tailored to the user's personality profile.

IV. RESULTS AND DISCUSSION

By and by, to specify monotonously, that personality expectation from tweets still goes under semantic examination and is available to even now an immense range of potential outcomes to enhance its course of nature after some time as more institutionalization will be presented in removing significant surmising from content. This approach supplements explore business related to not just versatile expectations utilizing machine learning and Twitter yet additionally proposal methodologies for users and customized client involvement with promotions and substance recommendation. However, our working usage of the personality expectation demonstrates utilizing NLTK toolbox, Scikit-learn and Python can be informed with the accompanying forbidden information. Comparative Analysis of proposed model with Naïve Bayes is displayed in Table 1 and graphs are shown in figure 4 precision, accuracy and recall respectively.

Table 1: Comparative analysis of Naïve Bayes and proposed model

No. of features	Naïve Bayes Model			Proposed Model (Logistic Regression)		
	P	**A**	**R**	**P**	**A**	**R**
100	64.91	68.91	66.91	65.21	70.25	67.26
500	66.48	71.48	68.28	67.21	72.24	68.43
1000	68.39	74.39	69.79	69.45	76.78	69.99
1500	71.47	79.21	72.47	72.21	81.21	73.19
10000	76.36	83.36	78.36	77.23	84.45	78.97
15000	78.89	85.19	79.99	81.92	87.84	82.64
20000	82.57	88.57	82.87	84.34	89.66	84.13

* P- Precision, A- Accuracy, R- Recall

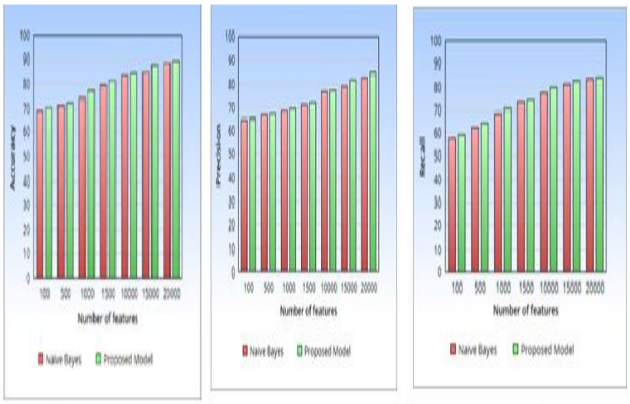

Figure 4a Figure 4b Figure 4c

Figure 4. Comparative graphs for Naïve bayes and proposed model

V. CONCLUSION AND FUTURE SCOPE

In this paper, we have shown that users' Big Five personality traits can be predicted from the public information they share on Twitter. After processing this data, we found many small correlations in the data. Using the profile data as a feature set, we were able to train two machine learning algorithms - m5sup 'Rules and Gaussian Processes - to predict each of the five personality traits to within 11% of its actual value. With the ability to guess a user's personality traits, many opportunities are opened for personalizing interfaces and information. We discussed some of these opportunities for marketing and interface design above. However, there is much work to be pursued in this area. One area that deserves attention is the connection between personality and the actual social network. We considered two structural features - number of friends and network density - but we did not look at personality scores between friends. Understanding the connections between personality, tie strength, trust, and other related factors is an open space for research. By improving our knowledge of these relationships, we can begin to answer more sophisticated questions about how to present trusted, socially-relevant, and well-presented information to users.

REFERENCES

[1] Cantandir, I. Fernandez-Tobiaz, A. Belllogin, "Relating personality types with user preferences in multiple entertainment domains", EMPIRE 1st Workshop on Emotions and Personality in Personalized Services, **2013.**

[2] L.R. Goldberg, J.A. Johnson, H.W. Eber, R. Hogan, M.C. Ashton, C.R. Cloninger, "The International personality item pool and the future of public domain personality measures", Journal of Research in Personality, **vol. 40, no. 1, pp. 84-96, 2006.**

[3] F. Mairesse, M. Walker, M. Mehl, R. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text", Journal of Artificial Intelligence Research (JAIR), **vol. 30, no. 1, pp. 457-500, 2007.**

[4] Sucharitha Y, Vijayalata Y, Prasad VK. Analysis of Early Detection of Emerging Patterns from Social Media Networks: A Data Mining Techniques Perspective. InSoft Computing and Signal Processing 2019 (pp. 15-25). Springer, Singapore.

[5] P.T. Costa, R.R. McCrae, "Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI)", Psychological Assessment Resources, **1992**

[6] Sucharitha Y, Prasad VK, Vijayalatha Y. Emergent Events Identification in Micro-Blogging Networks Using Location Sensitivity.

[7] J. Golbeck, C. Robles, K. Turner, "Predicting personality with social media", CHI'11 Extended Abstracts on Human Factors in Computing Systems, **pp. 253-262, 2011.** Access at ACM

[8] Yadala Sucharitha, Y. Vijayalata and V. Kamakshi Prasad*, "Predicting Election Results from Twitter Using Machine Learning Algorithms", Recent Advances in Computer Science and Communications (2020) 13: 7. https://doi.org/10.2174/2666255813999200729164142

[9] M. Hall, E. Frank, H. Holmes, P. Pfahringer, P. Reutemann, I.H. Witten, "The WEKA data mining software: an update", SIGKDD Explorations, vol. 11, no. 1, **2009**. Access at ACM

[10] Reddy PC, Babu AS. Survey on weather prediction using big data analystics. In2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT) 2017 Feb 22 (pp. 1-6). IEEE.

[11] N.M.A. Lestari, I.K.G.D. Putra, A.A.K.A. Cahyawan, "Personality types classification for Indonesian text in partners searching website using Naive Bayes methods", International Journal of Software and Informatics Issues, **2013**.

[12] Arifin, K.E. Purnaha, "Classification of emotions in Indonesia text using K-NN method", International Journal of Information and Electronics Engineering, **vol. 2, no. 6, 2012**

[13] Reddy PC, Sureshbabu A. An applied time series forecasting model for yield prediction of agricultural crop. InInternational Conference on Soft Computing and Signal Processing 2019 Jun 21 (pp. 177-187). Springer, Singapore.

[14] A. Firoj, E.A. Speanov, G. Riccardi, "Personality traits recognition on social network – Facebook", ICWSM Workshop on Personality Classification, **2013**

[15] F. Iacobelli, A. Cullota, "Too neurotic not too friendly: structured personality classification on textual data", ICWSM Workshop on Personality Classification, **2013**.

[16] Reddy PC, Babu AS. A novel approach to analysis district level long scale seasonal forecasting of monsoon rainfall in Andhra Pradesh and Telangana. International Journal of Advanced Research in Computer Science. 2017 Nov 1;8(9).

[17] F. Celli, F. Pianesi, D. Stillwell, M. Kosinski, "Workshop on computational personality recognition (shared task)", Proceedings of WCPR13 in conjunction with ICWSM-13, **2013**.

[18] Daniele Quercia§ Michal Kosinski† David Stillwell∗ Jon Crowcroft§, Our Twitter Profiles, Our Selves: Predicting Personality with Twitter, Horizon Digital Economy Research grant (EP/G065802/1).

[19] Reddy PC, Sureshbabu A. An Adaptive Model for Forecasting Seasonal Rainfall Using Predictive Analytics. InInternational Journal of Intelligent Engineering and Systems 2019 (pp. 22-32).

[20] Settanni M, Azucar D, Marengo D. Predicting individual characteristics from digital traces on social media: A meta-analysis. Cyberpsychology, Behavior, and Social Networking. 2018 Apr 1;21(4):217-28.

[21] Ozer DJ, Benet-Martinez V. Personality and the prediction of consequential outcomes. Annu. Rev. Psychol.. **2006** Jan 10;57:401-21.

[22] Roberts BW, Kuncel NR, Shiner R, Caspi A, Goldberg LR. The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. Perspectives on Psychological science. 2007 Dec;2(4):313-45.

[23] Soto CJ. How replicable are links between personality traits and consequential life outcomes? The Life Outcomes of

Personality Replication Project. Psychological Science. 2019 May;30(5):711-27.

[25] Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ. 2016 Sep 29;4:e2537.

[26] P Chandrashaker Reddy* and A Suresh Babu, “An Enhanced Multiple Linear Regression Model for Seasonal Rainfall Prediction”, International Journal of Sensors, Wireless Communications and Control (2020) 10: 3.

[27] Ferwerda and M. Schedl, "Enhancing music recommender systems with personality information and emotional states: A proposal", CEUR Workshop Proc., **vol. 1181, pp. 36-44, 2014.**

[28] Ferwerda, E. Yang, M. Schedl and M. Tkalčič, "Personality Traits Predict Music Taxonomy Preferences", Proc. 33rd Annu. ACM Conf. Ext. Abstr. Hum. Factors Comput. Syst. - CHI EA ‘15, **pp. 2241-2246, 2015.**

