

PERSONALITY PREDICTION MODEL USING DEEP LEARNING

Harshit Agarwal
School of Computer Science
Galgotias University
Greater Noida
aagarwalharshit10@gmail.com

Prof. Dr. Meenakshi Sharma
School of Computer Science
Galgotias University
Greater Noida
minnyp@gmail.com

Rachna Jha
School of Computer Science
Galgotias University
Greater Noida
rachna14025@gmail.com

Abstract — The project depends on distinguishing the character of a singular utilizing AI calculations and large 5 models. The character of a human assumes a significant part in his own and proficient life. These days, numerous associations have likewise begun shortlisting the competitors in view of their character as this builds the proficiency of the work in light of the fact that the individual is working in what he is great at as opposed to what he is compelled to do. The Big Five model is otherwise called the Five-Factor Model (FFM) and OCEAN model was created in the mid 1980s as per numerous mental hypotheses. Whenever measurable examination is applied to character study information, a few words used to depict the individual and these words give a rundown of the general person or character of the individual precisely. Character as seen as a brand name of mindfulness, conduct and energetic models from natural and ecological parts. It shows individuals as they quarrel about contemplations, ways of behaving and sentiments. Character attributes don't change in that frame of mind as they give the sensation of being out of control the conspicuous characteristics of an individual with a predictable quality as opposed to mirroring a specific character. Character contrasts ought not out of the ordinary to forestall the justification behind the bar or struggle in the subject of work or schooling. Along these lines, in enrolling an everyday schedule of work, one must be isolated.

In our undertaking a 70:30 train test break not entirely settled and supported subsequent to contrasting the exactness of the different train test fragment particulars. Presently subsequent to taking care of the models with some portion of the data set test, the exactness acquired by the various models is thought about. We considered two measurements in view of which model was picked. These are two measurements for precision and execution. In view of these variables, Logistic Regression was decided to be the principal model for this undertaking. We have attempted to join both character forecasts utilizing ML calculations like K-Mean and Logistic relapse to foresee the character of an individual and utilized the term recurrence calculation to get the ability where the individual is great at. Clients can without much of a stretch recognize his character and his specialized ability from this model or framework.

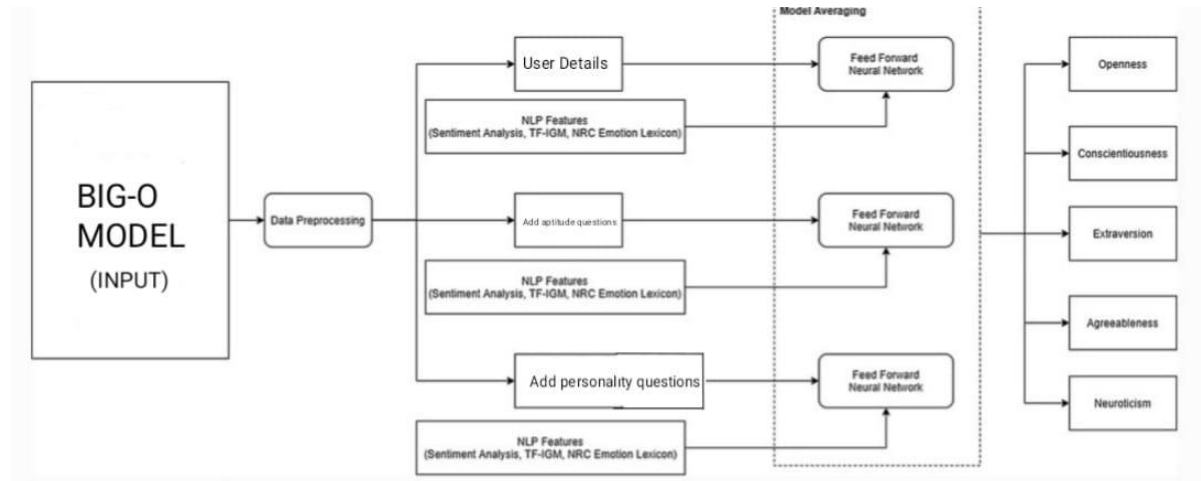
Keywords: Big-five model, KNN, CNN, Personality- prediction, machine learning, deep learning

1. Introduction

The framework module contains a client module and an overseer module.

Client login will be utilized by candidates to actually take a look at their characters. Rundown of client exertion endeavors view Results. The candidate needs to make and present their CV by finishing the CV structure. The CV configuration ought to be done in the way endorsed by the program. Chairman sign-in will be utilized by the recruiting organization to survey the character and specialized capacity of the individual to be chosen by the overseer who can see every one of the subtleties of the enrolled people.

The overseer can see the consequences of individual competitors which might make it more straightforward for the manager to choose the up-and-comer they need. Subtleties of the individual to be recorded on the outcomes page incorporate name, age, address, character and their expert abilities to be perceived after effectively transferring a CV, the individual to be composed can proceed with their web-based character based appraisal. The survey comprises of 4 inquiries every one of these 5 images are given to the client for every 2 question mark and will have a sum of 20 inquiries 4 inquiries relating to every character attribute. There are 700 informational collections and 300 test sets utilized for this data set, each characteristic is given a limit of 8 places. In light of the client's response to each question in the element tag doled out to it.



Whenever factual investigation is applied to character study information, a few words used to portray the individual and these words give a rundown of the general person or character/qualities of the individual precisely.

- Open to Experience: It includes different aspects, similar to mindfulness, creative mind, responsiveness, inclination to interest and assortment
- Reliability: This quality is utilized to depict the steadiness and watchfulness of the individual. The quality depicts how respectful and productive an individual is.
- Extraversion: It is the quality that portrays how all that competitors can associate with individuals that is the means by which great are his/her interactive abilities.
- Appropriateness: It is a quality that examinations the singular conduct in light of the liberality, compassion, helpfulness and capacity to change with individuals.
- Neuroticism: This quality for the most part depicts an individual to have mind-set swings and has outrageous expressive power.

2. Ease of use

Character is viewed as a brand name of mindfulness, conduct and energetic models from natural and ecological parts. It shows individuals as they squabble about contemplations, ways of behaving and sentiments. Character characteristics don't change in that frame of mind as they give the sensation of being out of control the conspicuous attributes of an individual with a reliable quality as opposed to mirroring a specific character. Character contrasts ought not out of the ordinary to forestall the justification for the bar or struggle in the subject of work or instruction. Along these lines, in enrolling an everyday schedule of work, one must be isolated.

Consequently, to foresee the singular character attributes we fostered an AI model in particular, the Myers-Briggs Type Indicator (MBTI). This venture follows the MBTI guideline as a manual for assist with recognizing client character in view of the accompanying character rules: Introvert (I) and Extrovert (E), Sensor (A) and Intuition (N), Thinking (T) and Sense (F), Sight (P) and Judgment (J). The mix of the over four kinds of character qualities will prompt sixteen character types, for example, "INFJ" or "ENFP" and so on. In our model we have utilized calculations like KNN, Logistic Regression, XG Boost. We took the data set from the Kaggle source. In our model we start by bringing in information from Kaggle. Then, at that point, add it to the information examination to Check whether there are absent or missing sums in the information base and the investigated information is remembered for the pre-information examination to clear the information. After the cleaning system, the information is shipped off the designing element lastly by contrasting the calculations and others we select the best calculation in our model that can foresee the character of every person.

3. Literature Review

An informational index containing characteristics, for example, Type and Post is brought into the venture. Information should be

examined to track down the missing qualities if accessible, to grasp the sort of cost information in the characteristics, to know the number of lines that are in the data set and to track down conditions between modifiers assuming they exist. Information examination was moreover

performed to break down the evaluations of each kind of individual in the information type.

Fundamental handling of information was performed to extricate inaccurate words, accentuation marks, copy letters and word-stops, to refine the information to increment. Fundamental handling comprises of chosen word evacuation, lemmatization execution and change of MBTI 4 parallel code into double. In the Selective Word Removal certain words might make the machine cheat. Lemmatization carries every one of the names to their starting point. For instance, running, playing and thinking becomes running, playing and thinking in grouping. Changing the MBTI character over completely to twofold includes doling out every component a 0 or 1 to be effectively deciphered by the model. Here I, N, F, J are totally relegated to 0, while E, S, T, P are completely allocated to 1. In this way, the ENFP character type is converted into [1,0,0,1].

Past concentrate on character expectation has been finished by utilizing virtual entertainment Facebook and a few highlights such as LIWC highlights, SNA highlights, time-related elements, and others 4

. Their examination is practically the same with our own

particularly for the dataset (250 dataset from my Personality) and the highlights (LIWC and SNA highlights). Another research in character forecast in view of Facebook status were finished by utilizing two methodologies, for example, open vocabulary DLA (Differential Language Analysis) and LIWC highlights 5

. By utilizing Facebook, an exploration characterizing

highlights with sack of-words and token (unigrams) approaches were led also. Other review was done to make a character forecast framework by involving Twitter with LIWC and MRC as highlights 6

. All referenced above investigates did character expectation by involving web-based entertainment in English in light of Big Five Character models. Late exploration was led to make a character expectation framework involving Twitter in Bahasa in view of Big Five Personality models 7,8

. Other examination on character forecast was finished utilizing profound learning procedure to group Big Five Personality models from online entertainment Facebook

Contrasted with this multitude of approaches, this exploration has concentrated to catch character of an individual from numerous mix of profound learning designs with model averaging. Scientist likewise utilized NLP highlights as extra elements to profound learning engineering, acquired from psycho-etymological and fundamental phonetic highlights.

4. Methodology

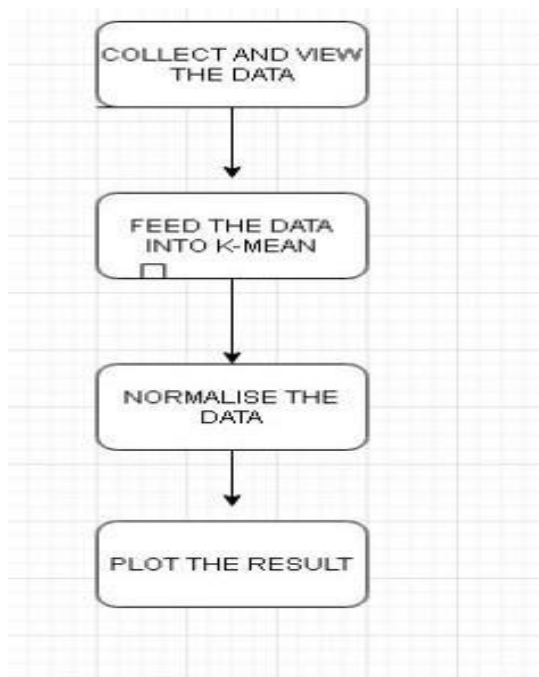
In our project a 70:30 train test break rate was determined and approved after comparing the accuracy of the different train test segment specifications. Now after feeding the models with part of the database test, the accuracy obtained by the different models is compared. We considered two metrics based on which model was chosen. These are two metrics for accuracy and performance. Based on these factors, Logistic Regression was chosen to be the main model for this project.

In short, all behavioral classes have had a specific impact on the prediction of individual trait points. However, behaviors related to social interaction and application use have been shown to be very important for models.) .

4.1 Data

The first dataset called the data-final dataset, which comprises of 1 million clients with a sum of 20 thousand situations with. This dataset is gathered through kaggle, permitting clients to take part in mental examination by filling in the character survey [7]. The dataset is an extended dataset as exploration done is gathered through multiple surveys . In this drawn out adaptation, the dataset is added with gathered information bringing about a sum of 5lakh clients with 20538 situations with. Equivalent to past research, the gathered information will be commented on by brain science master.

Furthermore, all datasets will be isolated into three sorts, in particular preparation set, test set, and approval set where the proportion of the appropriation is 70% preparation set and 15% test and 15% approval.



4.2 Preprocessing

The model results show that social interaction and behavior have a significant impact on model performance in the domain ($\beta = 0.027$, $CI95\% = [.026, .028]$) and facet levels ($\beta = 0.019$, $CI95\% = [.019, .020]$). Application usage was secondary (domains $\beta = 0.014$, $CI95\% = [.013, .015]$, $\beta_{facets} = 0.014$, $CI95\% = [.014, .015]$) followed by

day and night activity ($\beta_{domains} = 0.013$, $CI95\% = [.012, .014]$, $\beta_{facets} = 0.011$, $CI95\% = [.011, .012]$), total phone activity ($\beta_{domains} = 0.006$, $CI95\% = [.005, .007]$, $\beta_{facets} = 0.004$, $CI95\% = [.004, .005]$), and music (domains $\beta = 0.001$, $CI95\% = [.000, .002]$, $\beta_{facets} = 0.001$, $CI95\% = [.001, .002]$).

The behavioral phase of the movement was very important in predicting the size of the Big Five personality traits ($\beta_{domains} =$ We provide additional, experimental results

for a randomized sample search, showing which combinations of behavioral categories were most predictable in general, in our database.

Mainly here we used procedure and implementation of big-five model alongside with KNN and CNN.

4.3 Feature Extraction

As far as measurable elements, this examination utilizes various methodologies contrast with the past exploration [37] which use TF-IDF as term weighting factor, rather TF-IGM is presented in research. TF-IGM consolidate another factual model to gauge the heaviness of each class in text exactly. The weight expresses the significance or word commitment to the class of records. Moreover, this technique can isolate mark classes in a printed information, particularly for information that has more than one name.

Subsequently, this technique is truly appropriate for use in character forecast which permits an individual to have more than one character. The TF-IGM worth can be determined by searching for the TF esteem and the IGM esteem. TF addresses the heaviness of a word, where the number of words that show up in a report. In the interim, IGM is valuable for estimating the strength of a word in recognizing one class and another. The computation of TF and IGM values can be depicted in the accompanying equation-

TF=Total appearance of a word in a document/ Total words in a document

(1)

IGM=1+λ(Total appearance of a word in a document/Total appearance of a word in each class)

(2)

TF-IGM=TF*IGM TF-IGM=TF*IGM

(3)

λ address a customizable coefficient, which use to keep the general harmony between the worldwide and nearby factors in the heaviness of a term. In addition, TF-IGM worth will go from zero to 1. Each word in an archive will be counted with a TF-IGM esteem then the words will be arranged by the biggest worth. Words that have incredible worth will be utilized as an element for making arrangement models since it tends to be accepted that these words contain the significant importance of a record with a particular class name. Finally, the utilization of semantic investigation and NRC feeling vocabulary as corresponding elements in anticipating attributes of an individual as were likewise utilized in this review. The two strategies utilize the open jargon approach, which require a predefined corpus in tracking down the logical element from a text information.

The softmax function and scaled down function denoted by the following formula.

$$\text{Scaled}(x) = \frac{QKdk}{\sum_{j=1}^K \text{Scaled}(x_j)} \quad (4)$$

$$\text{Softmax}(x) = \frac{\exp(x_i)}{\sum_{j=1}^K \exp(x_j)} \quad (5)$$

As per past profound learning writing [21, 28], the unweighted averaging may be a sensible troupe for comparative base students of equivalent execution. In this examination the model averaging (unweighted) can be determined by joining the softmax probabilities from three unique arrangements model. The mean class the likelihood is determined as follow:

$$y_{i,k} = y_{i1,k} + y_{i2,k} + y_{i3,k} \quad \forall k \in [1..K] \quad (6)$$

$$y = \arg \max_k (y_{i,k}) \quad (7)$$

where KK is the number of classes, and yy is the predicted label for a sentence. For loss function, a cross entropy loss denoted by the following formula was used.

$$\text{CrossEntropyloss} = -(y \log(p) + (1-y) \log(1-p)) \quad (8)$$

where y is the genuine name worth and p is the anticipated character of from a sentence. To augment the exhibition of the model fabricated, the boundary tuning cycle will be completed. Matrix search technique will be utilized to perform rehashed look in observing ideal boundaries that will deliver the most extreme degree of prescient execution. A portion of the boundaries to be changed are the cluster size, age, and learning rate.

5.Implementation

5.1 Dataset

The first dataset called the data-final dataset, which comprises of 1 million clients with a sum of 20 thousand situations with. This dataset is gathered through kaggle, permitting clients to take part in mental examination by filling in the character survey. Furthermore, all datasets will be isolated into three sorts, in particular preparation set, test set, and approval set where the proportion of the appropriation is 70% preparation set and 15% test and 15% approval.

	EXT1	EXT2	EXT3	EXT4	EXT5	EXT6	EXT7	EXT8	EXT9	EXT10	EST1	EST2	EST3
0	4.0	1.0	5.0	2.0	5.0	1.0	5.0	2.0	4.0	1.0	1.0	4.0	4.0
1	3.0	5.0	3.0	4.0	3.0	3.0	2.0	5.0	1.0	5.0	2.0	3.0	4.0
2	2.0	3.0	4.0	4.0	3.0	2.0	1.0	3.0	2.0	5.0	4.0	4.0	4.0
3	2.0	2.0	2.0	3.0	4.0	2.0	2.0	4.0	1.0	4.0	3.0	3.0	3.0

5.2 Implementation

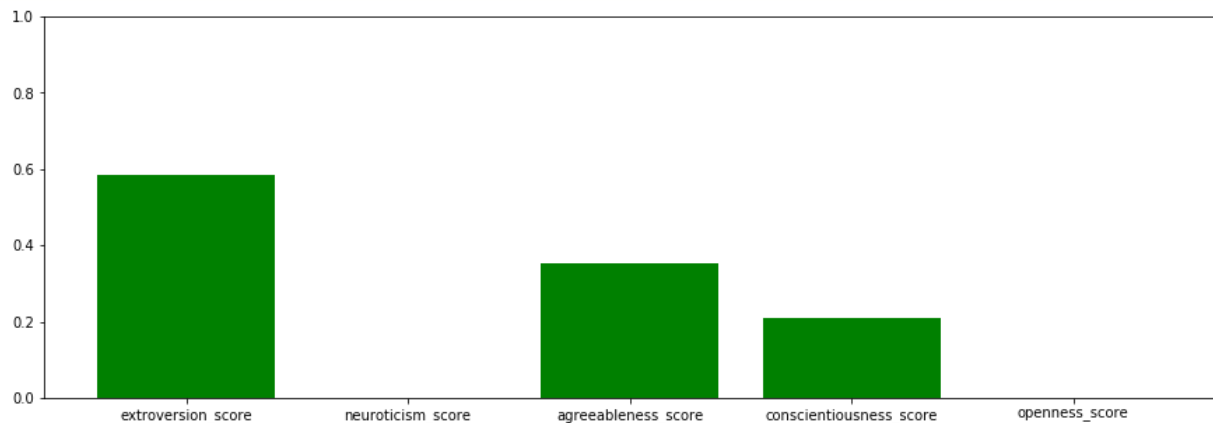
This section describes about the training and testing of the model. The proposed prediction model is evaluated using a PC having JUPYTER NOTEBOOK environment with i7 processor and 8GB RAM on data-final dataset. To train the model batch-size of 51akh users input is used.

5.2.1 Training & Model Building

The training phase includes the data preprocessing, augmentation and model building of the network. As we have already discussed , to implement the model KNN is used and the data is loaded by using the panda class. To label the data properly we created an array in which data is stored in column and rows and are stored in the respective train_dataset and validation_dataset, they are then labeled as openness, conscientiousness, extraversion, agreeableness and neuroticism respectively for the training stage. The numpy python library is used for displaying the smart progress bars which show the progress of the code execution.

```
all_types = {'one':one, 'two': two, 'three' :three, 'four':four, 'five':five, 'six': 'nine': nine, 'ten': ten} all_types_scores = {}
for name, personality_type in all_types.items(): personality_trait = {} personality_trait['extroversion_score'] =
personality_type[0] - personality_type[1] personality_trait['neuroticism_score'] = personality_type[0] - personality_type[1]
personality_trait['agreeableness_score'] = personality_type[0] - personality_type[1] + personality_type[2]
personality_trait['conscientiousness_score'] = personality_type[0] - personality_type[1]
personality_trait['openness_score'] = personality_type[0] - personality_type[1] all_types_scores[name] = personality_trait
```

```
all_extroversion = [] all_neuroticism =[] all_agreeableness =[] all_conscientiousness =[] all_openness =[] for
personality_type, personality_trait in all_types_scores.items():
all_extroversion.append(personality_trait['extroversion_score'])
all_neuroticism.append(personality_trait['neuroticism_score'])
all_agreeableness.append(personality_trait['agreeableness_score'])
all_conscientiousness.append(personality_trait['conscientiousness_score'])
all_openness.append(personality_trait['openness_score'])
```



6. Result and Discussion

The aftereffects of the model that have been made will be assessed utilizing a few metric estimation approaches as follows:

a.F1 Measure

Estimation metric of a model which consolidates the typical upsides of accuracy and review creating score by thinking about an order blunder. These estimation measurements are best utilized when misleading negative and bogus positive qualities are significant. In the expectation of character, the bogus positive and misleading pessimistic qualities are considered to decrease prescient blunders since, in such a case that the forecasts are off-base, perhaps somebody can be set inconsistent with their character.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (9)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (10)$$

$$\text{F1 Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

b.Precision

Helpful as a proportion of the exhibition of a model, but this estimation centers around the all out information that is exactly anticipated, to be specific genuine positive and genuine negative. This estimation is great for class dissemination on adjusted information. In light of past examination, many utilize this estimation as assessment metric. Subsequently, to contrast the consequences of exploration and past examinations, this measurement is utilized.

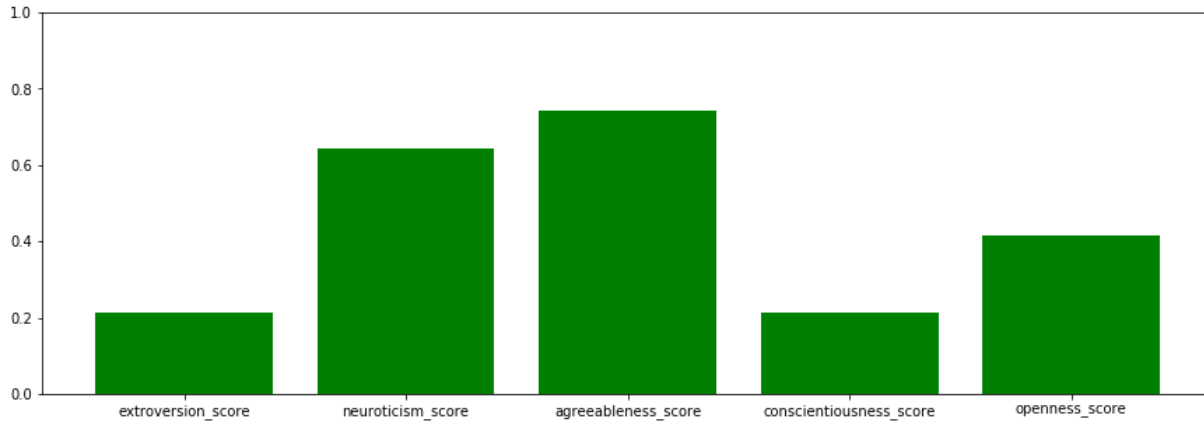
Accuracy=

$$\frac{(\text{True Positive} + \text{True Negative})}{(\text{True Positive} + \text{False Positive} + \text{True Negative} + \text{False Negative})} \quad (12)$$

The main result will fetch the predicted value for each type. The main concern was to improve accuracy, so we took a different algorithm to check for accuracy and chose the most accurate algorithm that gave the most accurate results.

Logistic Regression is designed to be divided into binary categories and does not support divisional functions that have more than two categories automatically. Dividing a multi-category data set into multiple binary data sets and inserting a binary

division algorithm into each is one way to use binary division algorithms in multi-stage situations. Instead of using multiple class layouts we use a single section to separate a personality feature that will be repeated in a loop 4 times to produce the final output. Tests and training is performed using test_train_spilt at different values to determine the best variables to divide the database into two parts.



All prescient models that have been made will be assessed utilizing the precision and f1 measure metric methodology. The aftereffect of the assessment utilizing the dataset, which shows that the most elevated exactness created from every quality overwhelmed by the proposed model which utilized model averaging technique and NLP factual highlights. The first and second most noteworthy precision is delivered in the Openness character model with 86.17% exactness and 0.912 f1 measure score, Neuroticism characteristic with a precision of 78.21% and f1 measure score 0.709. Notwithstanding, concerning Agreeableness quality the most noteworthy precision and f1 measure score of other character models is made by utilizing the XGBoost with expansion of NLP highlights, which is found in Agreeableness characters with exactness upsides of 72.33% and 0.701. Be that as it may, the subsequent worth contrasts somewhat for around 0.74% and 0.011 from the proposed model concerning the two measurements. By taking a gander at all the mean precision of the examinations on every calculation, it was observed that the proposed model engineering has the most noteworthy typical exactness and f1 score, which is 77.34% and 0.749 for character forecast framework utilizing dataset.

7. Limitations

Interest in the relationship between personality traits and physical health has been renewed in recent years. Theory and research in this area has been difficult to quantify the concept and approach. The current article briefly reviews these books and discusses the benefits and limitations of the five-dimensional human model as an integrated framework for human and health studies. The model has already been successfully used in a few cases, and many opportunities exist. Although there may be some limitations, the use of a five-dimensional model — and other aspects of current human theory and research — may make it easier to make progress in research on how humanity influences life.

8. Conclusion and Future Work

The main purpose of the project is to use user content to create a personal profile. Preliminary analysis included looking at attitudes of trends and analysis analysis. In the data collected on social media sites, there are many raw feelings that are also useful to predict personality. Then, in order to clean up the data, the text was pre-processed by removing links, numbers, punctuation marks, and sensitive words in context. Different machine learning models of different combinations are studied. These days, many parties have begun screening candidates based on their personalities as it increases efficiency as a person works better than he or she should. Our model includes algorithms such as KNN, XG Boost, random forest, Stochastic Gradient Descent and Logistic retrospective. These algorithms are comparable. It was found that the best results were obtained using a retrospective regression. However, the accuracy is almost the same after using XG Boost. Our model helps users to easily identify their personality in this model.

However, the same can be implemented in future for datasets with larger size for getting better results. The clusters using KNN can be beneficial in assisting professionals in differentiating the staff at an early stage and thus can help in increasing their efficiency. This will help in providing proper work flow in the industries and help the industries to grow.

9. References

1. Christian, H., Suhartono, D., Chowanda, A., & Zamli, K. Z. (2021). Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging. *Journal of Big Data*, 8(1), 1-20.
2. Tandra, T., Suhartono, D., Wongso, R., & Prasetio, Y. L. (2017). Personality prediction system from facebook users. *Procedia computer science*, 116, 604-611
3. Shen, J., Brdiczka, O., & Liu, J. (2013, June). Understanding email writers: Personality prediction from email messages. In *International conference on user modeling, adaptation, and personalization* (pp. 318-330). Springer, Berlin, Heidelberg.
4. Abood N. Big five traits: a critical review. *Gadjah Mada Int J Business*. 2019;21(2):159–86.
5. Shen, J., Brdiczka, O., & Liu, J. (2013, June). Understanding email writers: Personality prediction from email messages. In *International conference on user modeling, adaptation, and personalization* (pp. 318-330). Springer, Berlin, Heidelberg.
6. Ong, V., Rahmanto, A. D., Suhartono, D., Nugroho, A. E., Andangsari, E. W., & Suprayogi, M. N. (2017, September). Personality prediction based on Twitter information in Bahasa Indonesia. In *2017 Federated Conference on Computer Science and Information Systems (FedCSIS)* (pp. 367-372). IEEE.
7. Robey, A., Shukla, K., Agarwal, K., Joshi, K., & Joshi, S. (2019). Personality prediction system through CV analysis. *Int. Res. J. Eng. Technol*, 6(2), 2395-0056.
8. Kunte, A. V., & Panicker, S. (2019, November). Using textual data for personality prediction: a machine learning approach. In *2019 4th international conference on information systems and computer networks (ISCON)* (pp. 529-533). IEEE.
9. S. Bharadwaj, S. Sridhar, R. Choudhary and R. Srinath, "Persona Traits Identification based on Myers- Briggs Type Indicator(MBTI) - A Text Classification Approach," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018,pp. 1076-1082,doi: 10.1109/ICACCI.2018.8554828.
10. C. Y. Yaakub, N. Sulaiman and C. W. Kim, "A study on personality identification using game-based theory," 2010 2nd International Conference on Computer Technology and Development, 2010, pp. 732-734, doi: 10.1109/ICCTD.2010.5646417.
11. A.Ariyanto, E. C. Djamal and R. Ilyas, "Personality Identification of Palmprint Using Convolutional Neural Networks," 2018 International Symposium on Advanced Intelligent Informatics (SAIN), 2018, pp. 90-95, doi: 10.1109/SAIN.2018.8673353.
12. Azucar, Danny & Marengo, Davide &Settanni, Michele. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta- analysis. *Personality and Individual Differences*. 124. 150-159. 10.1016/j.paid.2017.12.018.
13. Aditi V. Kunte, Suja Panicker. "Using textual data for Personality Prediction:A Machine Learning Approach" , 2019 4th International Conference on Information Systems and Computer Networks (ISCON), 2019
14. Predicting Mental health disorders using Machine Learning for employees in technical and non-technical companies,2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE),Date of Conference: 10-11Dec.2020,DateAddedtoIEEEExplore: 08March021,I NSPECNumber: 20532983,DOI:10.1109/ICADEE51157.2020.9368923,Publisher:IEEE,Conference Location:Coimbatore, India
15. [7]B. Y. Pratama and R. Sarno, "Personality classification based on Twitter text using Naive Bayes, KNN and SVM," 2015 International Conference on Data and Software Engineering (ICoDSE), 2015, pp. 170- 174, doi: 10.1109/ICoDSE.2015.7436992.
16. Safitri, G., & Setiawan, E. B. (2022). Optimization Prediction of Big Five Personality in Twitter Users. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 6(1), 85-91.
17. Safitri, G., & Setiawan, E. B. (2022). Optimization Prediction of Big Five Personality in Twitter Users. *Jurnal RESTI (Rekayasa Sistem Dan Te*
18. Catal, C., Song, M., Muratli, C., Kim, E. H. J., Tosuner, M. A., & Kayikci, Y. (2017). Cross-Cultural Personality Prediction based on Twitter Data. *J. Softw.*, 12(11), 882-891.
19. Gardade, M. R., & Khambre, P. (2021). A REVIEW ON PERSONALITY PREDICTION VIA SOCIAL MEDIA USING DATA MINING. *INTERNATIONAL JOURNAL*, 6(1).
20. Dharshni, P., Pon, J., Monisha, M., Nayanthara, C., & Krishna Priya, G. Personality Prediction Based on User Behavior on

Social Media.

21. Kalghatgi, M., Manjula, R., & Nandini, S. (2015). Social-network-sourced big data analytics for personality prediction: A review. *International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)*, 14(2), 382-388.
22. Narwade, R., Palkar, S., Zade, I., & Sanghavi, N. Personality Prediction with CV Analysis.
23. Mishra, S., & Sagnika, S. (2020, March). A Review on Personality Prediction. In *International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy* (pp. 61-70). Springer, Singapore.
24. Ožura, A., Ihan, A., & Musek, J. (2012). Can the big five factors of personality predict lymphocyte counts?. *Psychiatria Danubina*, 24(1.), 66-72.
25. Sherman, R. A., Nave, C. S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of personality and social psychology*, 99(2), 330.
26. Chraibi, K., Chaker, I., & Zahi, A. (2020, December). Automatic Personality Prediction: A Systematic Mapping Study. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 2053-2060). IEEE.
27. Mukherjee, S., Ghosh, I. D., & Mukherjee, D. (2022). Big Five Personality Prediction from Handwritten Character Features and Word 'of' Using Multi-label Classification. In *Proceedings of the Seventh International Conference on Mathematics and Computing* (pp. 275-299). Springer, Singapore.