

## Student Psychometric Analysis Through Machine Learning

**Anagha Shailesh Kulkarni<sup>1\*</sup>, Kumudavalli M.V<sup>2</sup>, Vanitha<sup>3</sup>**

<sup>1,2</sup>Dept. of Computer Applications, Dayananda Sagar College of Arts, Science & Commerce

<sup>3</sup>Computer Science Department, Seshadripuram PU College, Yelahanka

*\*Corresponding Author: email@hotmail.com, Tel.: +00-12345-54321*

DOI: <https://doi.org/10.26438/ijcse/v7si9.13> | Available online at: [www.ijcseonline.org](http://www.ijcseonline.org)

**Abstract**— Psychology can be defined as the mental characteristic or attitude of a person, especially those affecting behavior in a context. Every psychological test has an objective and standardized measurement of a sample behavior. Here, sample of behavior refers to an individual's response on a situation or task which is prescribed or predefined before the task. In order to make the psychological test cost, time, and efficiency effective, we focus on building a Machine Learning classification model which predicts if the student who took the survey belongs to an Introvert category or an Extrovert category. The questions in the survey will focus on an individual features to build a model and questions prepared by the domain expert will carry particular weight, and depending on the answers we will predict which category the person falls under. The performance of the model is predicted by taking Accuracy and confusion matrix into consideration. It is induced for the betterment or ease of analyzing a personality in order to enhance individual's strength, enhancing professional and personal skills.

**Keywords**—: *psychology, behavior, machine learning, skills.*

### I. INTRODUCTION

Psychology can be applied in today's high complex working, education or on any base. It does not be mandatory that only one approach may work for different people. That's why psychologists are working on different field and are focused on identifying different methods to ensure better understanding about how people absorb, retain information and react to the situation at that instinct of time [1]. The main objective of this study is to examine the psychology of an individual with different approaches and strategies in order to make an effective change necessary for an individual. Psychology test are taken on variety of scale such as pre-schools, primary and secondary school, college's as well as universities and cooperate companies. Test's have variety of format via, verbal, written or digital. It can be in form of questioner, survey or bases on trends etc.

#### A. Psychology test

Psychology testing can be also called as psychometrics is a systematic method to quantify psychophysical behavior, problems faced, abilities, IQ and psychological performance is predicted. Types of test included:

#### I. Interpersonal needs inventory [IPNI]

Profiles interpersonal needs, indicating balance found to be related [2, 3]. The three main interpersonal needs are

- i. Affection / Openness
- ii. Control

#### iii. Inclusion

#### I. Big 5

These dimensions are considered to be traits that make up the individual personality [4, 5]. The five dimensions are:

- i. Openness
- ii. Conscientiousness
- iii. Extraversion
- iv. Agreeableness
- v. Neuroticism

#### II. MBTI

It's a way of analyzing psychological preferences in how or what way people perceive or react to the world around them [6, 7].

Types:

- i. Extraverted (E) vs. Introverted (I)
- ii. Sensing (S) vs. Intuition (N)
- iii. Thinking (T) vs. Feeling (F)
- iv. Judging (J) vs. Perceiving (P)

### II. RELATED WORK

Dominic B. et al.[8], have given an insight about Machine learning approaches for clinical psychology and psychiatry which clearly focuses on learning statistical functions from multidimensional data sets to make generalized predictions. Since the student psychological study is an essential metric in current scenario, and machine learning has gained its

popularity in variety of fields , the current study has been conducted to through some light into the inter disciplinary domain.

**III. DATA SET DESCRIPTION**

The survey was held for the passing out students. A set of questions were prepared based on the three popular personalities test via, IPNI, Big 5 and MBTI. The questioner had major dimensions which makes a personality. Each question was followed by 6 point scale i.e. never, rarely, sometimes, occasionally, often, usually.

Three types of personality test were combined. The major focus was on 2 factors i.e. extrovert and introvert. Based on their approach to the given situation the answers were marked on 6 point scale. In response to 6 point scale the values were calculated on predefined formula.

Sample of Survey questions used for the data collection purpose:

1. Do you like to be a member of committee?
2. Do you like others to guide you?
3. Do you want people to support during crises?
4. Do you like people to consult you?
5. Do you enjoy working for yourself rather than working for group or committee?
6. You enjoy having a wide circle of acquaintances
7. You spend your leisure time actively socializing with a group of people, attending parties, shopping, etc
8. Am the life of the party?
9. Don't talk a lot
10. Keep in the background.

**IV. METHODOLOGY**

Step 1: The two questionnaires related to the factors were prepared.

Step 2: Data was collected from students of certain age group with different background in terms of perception, thinking and attitude.

Step 3: The parameters (never, rarely, sometimes, occasionally, often, usually) were defined for scaling based on which the numerical value is assigned.

Factors extracted from the tests are described in Table 1.

Table 1. Factor extraction criteria

Test	Factor	Score	Observation
IPNI	Extension	0 to 10 10 to 20 20 to 30	Introvert Ambivert Extrovert
Big 5	Extroversion	0 to 10 10 to 20 20 to 30	Introvert Ambivert Extrovert
MBTI	Extraverted (E) vs. Introverted (I)	0 to 10 10 to 20 20 to 30	Introvert Ambivert Extrovert

Step 4: The data is imported which is originally in CSV format.

Step 5: The data is checked to find if there is any NaN value present (empty value present). If yes, then the row with NaN value is removed.

Step 6: The problem is a classification problem and hence classification model like Logistic regression, Decision Tree, Random Forest is being chosen to train the model.

Step 7: since the data was imbalanced (in the received data of the current study we had Ambivert - 36, Extrovert - 9, Introvert- 4 ), to balance it gave more weight to Extrovert and Introvert so that Overfitting of a model can be avoided.

Step 8: The data was split into train and test, where train data size is 80% of the whole data and the test data size of 20%.

Step 9: Different models were applied on the data to check which has a better performs to the input data.

**V. RESULTS AND DISCUSSION**

Various tests like IPNI, Big 5, MBTI were combined to get a custom made questionnaire too analyze the personality of the outgoing students. The data set was used to train the model, and it was clearly found that the Random forest classification model performed very well as the number of columns were less. Random Forest gives an accuracy of 90% for the given test data.

The efficiency of the model is expressed as Accuracy, Precision and Recall matrix.

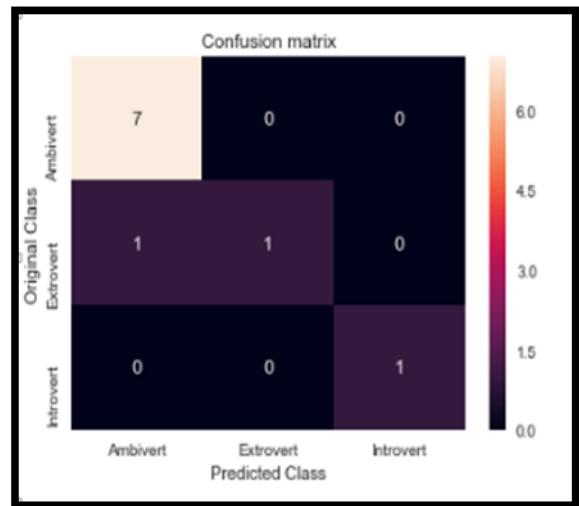


Figure 1. Accuracy matrix

Figure1 describes the Accuracy which is just the ration of correctly predicted observations to the total observation in the data set.

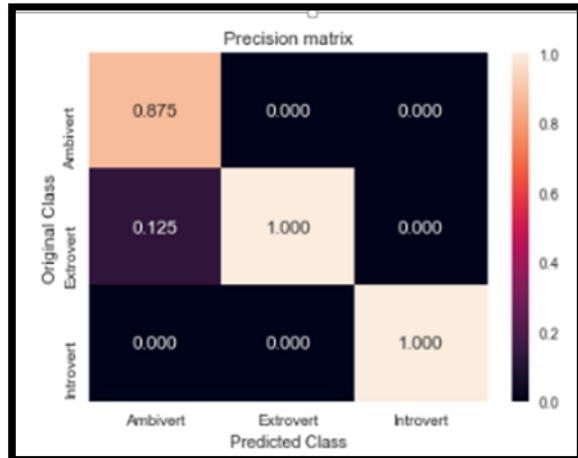


Figure 2: Precision Matrix

Figure 2 gives the precision matrix, which is the ratio of correctly predicted positive data to total predicted positive data.

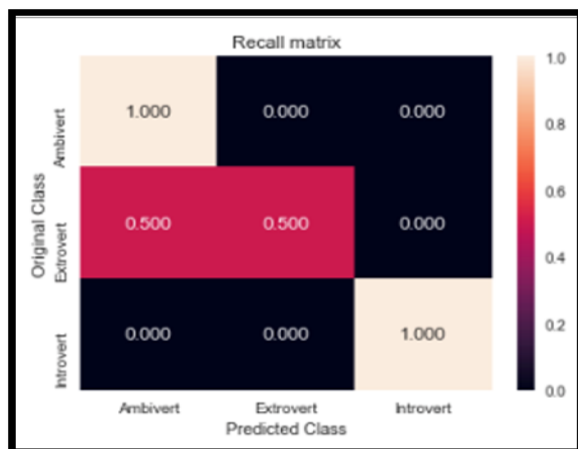


Figure 3: Recall Matrix

Figure 3 describes the Recall or sensitivity which is the ratio of correctly predicted positive data to all the observations of the actual class.

## VI. CONCLUSION AND FUTURE SCOPE

Psychology of human can be classified into n number of personalities. Every person has his/her own perception of reacting to the situation at that instance of time. There are many personality tests available online. All the tests available aren't cost, time and efficiency effective. In this analysis we have combined three major psychological test i.e. IPNI, Big 5, MBTI and a questionnaire is created. Based on the feedback data of the questionnaire a Machine Learning model is built and trained with few classification models and model performed well with Random Forest approach. The model efficiency is expressed with confusion matrix

## REFERENCES

- [1] Kumudavalli M.V., Anagha Shailesh Kulkarni, G. Ambrish, "3D Metric approach to study the factors affecting student's psychology on Education" International Journal of Computer Sciences and Engineering, Vol.7, Issue.2, pp.981-984, Feb-2019.
- [2] Michael Barkham, Gillian E. Hardy, Mike Startup, "The IIP 32: A short version of the Inventory of Interpersonal Problems", <https://doi.org/10.1111/j.2044-8260.1996.tb01159.x>, February 1996
- [3] Michael Barkham, Gillian E. Hardy, Mike Startup, "The structure, validity and clinical relevance of the Inventory of Interpersonal Problems", <https://doi.org/10.1111/j.2044-8341.1994.tb01784.x>, June 1994.
- [4] Jeromy Anglim, Sharon Grant, "Predicting Psychological and Subjective Well-Being from Personality: Incremental Prediction from 30 Facets Over the Big 5", Journal of Happiness Studies February 2016, Volume 17, Issue 1, pp 59–80
- [5] Hans Georg Wolff, Sowon Kim, "The relationship between networking behaviors and the Big Five personality dimensions", Emerald Group Publishing Limited
- [6] Alma M. McCarthy, Thomas N. Garavan, "Developing self-awareness in the managerial career development process: the value of 360 degree feedback and the MBTI", MCB UP Ltd
- [7] Passmore, Jonathan, "MBTI types and executive coaching", The Coaching Psychologist, 2006
- [8] Dominic B. Dwyer, Peter Falkai, and Nikolaos Koutsouleris Machine Learning Approaches for Clinical Psychology and Psychiatry Annual Review of Clinical Psychology Vol. 14:91-118, 2018.

## Authors Profile

**Dr. Kumudavalli M.V.**, is working as an Associate Professor at Dayananda Sagar College of Arts, Science & Commerce, Bangalore, India. Her research interests are Bioinformatics, Data Science, Operations Research, Networks etc. She has 14 years of academic experience. She has many research papers and certifications to her credits.



**Anagha Shailesh Kulkarni** is currently Pursing her degree. She is interested in psychology and NLP (Neuro Linguistic Programme). She developed her interest in human psychology and power of subconscious mind in early stage of her education. Initially she started conducting seminars and counseling to students from 8<sup>th</sup> to 10<sup>th</sup> grade. She is continuing her increase of knowledge base, interest and enhancing skills on regular bases.



**Vanitha S.** is currently working as Lecturer at Seshadripuram PU College, Yelahanka. Her research interests are in Iot, machine learning. Her academic experience is 12 years.

