Synopsis – Predicting MBTI

Name of Team Members: Nikhil Subhash Shelke (44)

Manas Sharad Torane (62)

Rohan Vishnu Pawar (24)

Pranay Sanjay Satpute (39)

Class: BE

Div: B

Domain: Machine Learning

Topic and Sponsorship: MBTI (Myers-Briggs Type Indicator)

Abstract : The MBTI (Myers-Briggs Type Indicator) is a widely known approach to personality classification. Datasets for the machine learning approach to personality classification using MBTI are highly imbalanced. Handling imbalanced data sets is a significant open problem with a considerable impact on machine learning methods. This paper presents the results of applying different techniques and suggests their best in mitigating the challenge of imbalanced MBTI datasets. Supervised learning algorithms such as Logistic Regression, Random Forest, and advanced models like BERT and LSTM will be employed to achieve high accuracy. The system has applications in career counseling, mental health support, personalized recommendations, and human resource management. Even though most techniques could be used and implemented to some other problems and areas, like images and sound processing, natural language processing has enough challenges to focus on natural language processing and the specific issue of the MBTI datasets.

Keywords: Personality Prediction, MBTI, NLP, Psychological Profiling, Text Classification, Machine Learning

Challenges identified :

- Mapping linguistic patterns accurately to MBTI dimensions
- High overlap between certain personality types
- Data imbalance across different MBTI categories
- Ensuring the model is not biased by topic or writing style
- Interpretability of predictions for psychological relevance
- Ethical concerns related to psychological profiling

Novelty or Industrial Application: -

- Useful for job-matching platforms, online counselors, and self-help
- Applicable to social media analysis, HR tech, and educational platforms
- Deployable as a lightweight API or browser extension for personality detection

Base IEEE/ Springer / Equivalent publication (paper URL):

https://ieeexplore.ieee.org/document/9596742

List of References:

- "MBTI Personality Prediction using Text Classification Techniques", IEEE, 2020
- "Predicting Personality from Social Media using NLP", Springer, 2021
- "Personality Detection via Linguistic Analysis of Text", ACM, 2019
- "Neural Approaches for MBTI Classification from Text", Elsevier, 2022
- "BERT-based Deep Learning Model for MBTI Personality Type Detection", Springer, 2023