NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
College of Engineering

# Personality Recognition

## From Text Based on the MBTI Model

Student: Yi Jia Xin Joceline  Supervisor: A/P Erik Cambria

### Project Objectives:

Personality is the individual differences in patterns of thought, feeling, and cognition formed from biological and environmental factors. It has been shown that personality propels a person's behaviour in society and affects individual outcomes such as happiness and subjective well-being. Traditionally, personality recognition is done through personality tests, which involve questionnaires and are usually conducted by qualified psychologists.

The aim of this project is to explore machine learning and deep learning techniques for automatic personality recognition from textual data based on the Myers-Briggs Type Indicator (MBTI) Model.

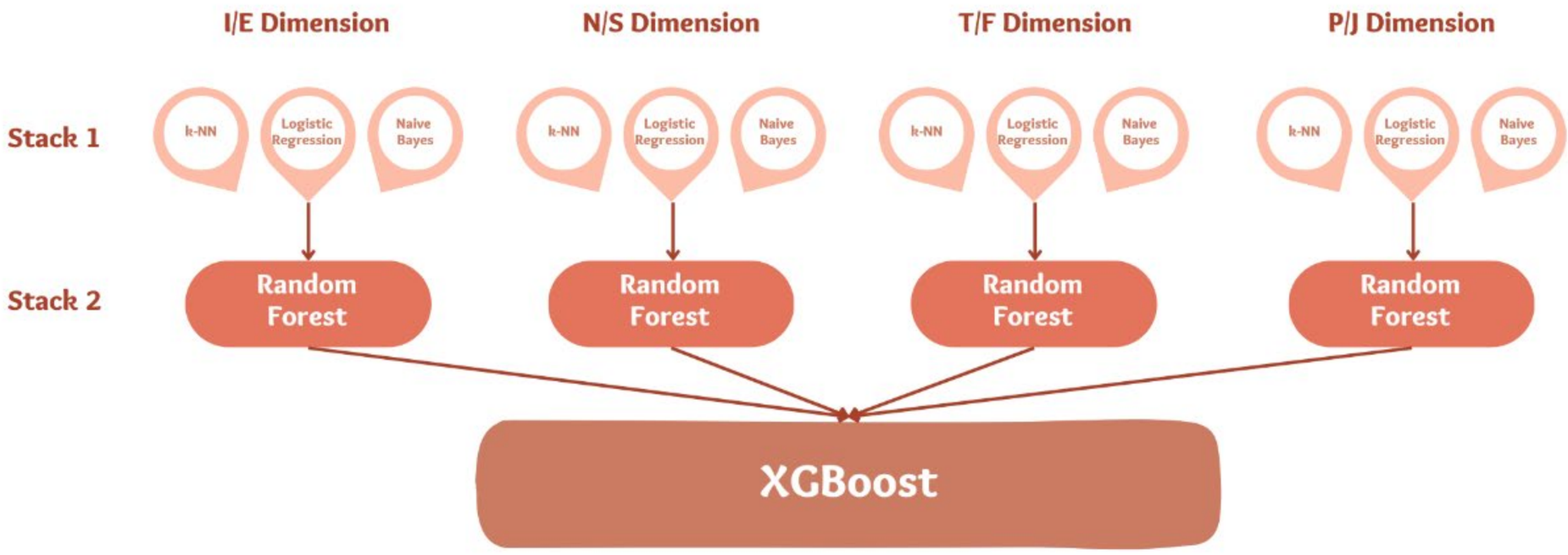
### Methods:

- Feature Engineering (Count Vectorization, TF-IDF, Word Embeddings – Word2Vec, GloVe)
- Machine Learning (Naïve Bayes, k-NN, Logistic Regression, Random Forest)
- Neural Networks (RNN, LSTM, GRU, CNN)
- Transfer Learning (Fine-tuned BERT, DistilBERT)
- Ensemble Learning (Boosting, Stacking, Bagging)

### Results:

| Binary Models | I/E Axes | N/S Axes | T/F Axes | P/J Axes | Multiclass Models | Accuracy |
|---|---|---|---|---|---|---|
| Logistic Regression | 61.34 | 59.82 | 66.69 | 64.31 | Stacking Ensemble 2 (with no under sampling) | 25.51 |
| Naïve Bayes | 60.90 | 59.42 | 67.78 | 63.79 | Stacking Ensemble 5 | 16.97 |
| Random Forest | 60.52 | 56.46 | 65.22 | 61.97 | Stacking Ensemble 3 | 16.93 |

### Proposed Solution:




Stacking Ensemble 4