Unraveling MBTI Types from Online Posts

Salma Zainana   Daneshvar Amrollahi

Stanford University

Introduction

Understanding human personality traits is crucial for applications like personalized content recommendation. The **Myers-Briggs Type Indicator (MBTI)** classifies people into one of 16 personality types based on four binary attributes. Our algorithm analyzes individual online forum posts, using Naive Bayes and GPT-4 models to predict the person's overall MBTI type. We also use Neural Networks to predict each of the four MBTI attributes of the individual.

Data Preprocessing

1. **Data cleaning**: remove irrelevant characters (e.g., punctuation), eliminate stopwords (common words lacking meaningful context), lemmatization (simplifies words to their base form), and Synthetic Minority Over-sampling Technique a.k.a. SMOTE (creates synthetic instances of the minority class).
2. **Feature Extraction**:
   - **TF-IDF (Term Frequency-Inverse Document Frequency)** is a statistical measure used to evaluate the importance of a word within a document relative to a corpus, resulting in a sparse vector representation of documents that emphasizes unique terms.

$$\text{TF}(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d}$$

$$\text{IDF}(t, D) = \log\left(\frac{\text{Total number of documents } D}{\text{Number of documents containing term } t + 1}\right) + 1$$

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \times \text{IDF}(t, D)$$

   - **Word2Vec** is a technique that leverages neural networks to map words into a high-dimensional vector space. It learns embeddings that cluster similar words together in this space, enabling the capture of complex word relationships and similarities.

Models

- **Naive Bayes**: Assumes all features are independent of each other within each class, simplifying calculations.

$$P(\text{MBTI}_k|\mathbf{x}) = \frac{P(\text{MBTI}_k)P(\mathbf{x}|\text{MBTI}_k)}{P(\mathbf{x})}$$

$$P(\mathbf{x}|\text{MBTI}_k) = \prod_{i=1}^{n} P(x_i|\text{MBTI}_k)$$

$$P(x_i|\text{MBTI}_k) = \frac{\text{Number of times } x_i \text{ appears in MBTI}_k + \alpha}{\text{Total number of words in MBTI}_k + \alpha \times D}$$

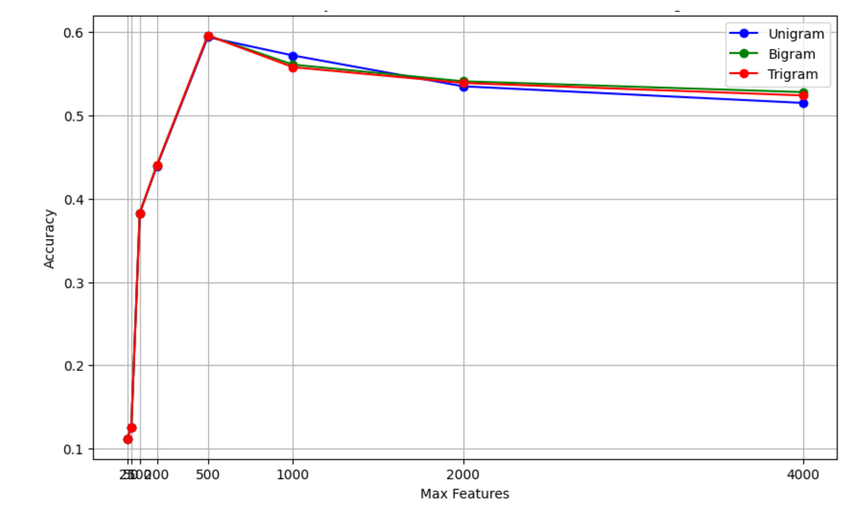
- **Neural Networks**: Tested various models including SimpleRNN, LSTM with Bidirectional LSTM, LSTM with Conv1D, and Conv1D alone
- **GPT**: Large Language Models (LLMs) like GPT use Transformer architecture for text generation, learning from vast text corpora through pretraining and fine-tuning for specific tasks. They employ self-attention to understand context and generate coherent text.

Results

- **GPT-4**: Experiment was conducted on $15 \times 16 = 240$ data points, with 15 data points for each personality type. Achieved accuracy **71%** in correctly predicting all 4 attributes and **88%** in predicting 3 correct attributes.
- **Neural Networks**:

MBTI Dichotomy	Best Threshold	Accuracy	ROC-AUC Score	G-Mean Score	F1-Score (avg/total)
Extrovert vs Introvert	0.93	0.7804	0.68	0.51	0.72
Intuition vs Sensing	0.33	**0.8599**	0.69	0.39	0.80
Thinking vs Feeling	0.59	0.7925	**0.87**	**0.79**	0.79
Judging vs Perceiving	0.68	0.6594	0.68	0.61	0.64

- Naive Bayes



Discussion

- Despite higher test set accuracy with LSTM (65%) compared to our primary model (59%), their ROC-AUC scores around 0.5 for each MBTI dichotomy suggest these architectures may not effectively generalize, and highlighting the necessity of using diverse metrics to evaluate true model performance.
- Naive Bayes is a simple yet effective classification method, typically bounded by a maximum achievable accuracy

Future Work

- **Refine Preprocessing & Sentiment Analysis**: Enhance text preprocessing to include sentiment analysis for capturing language nuances related to MBTI types.
- **Tackle LSTM Overfitting**: Explore novel regularization techniques and model adjustments
- **Fine-tuning GPT** to see what is the maximum achievable accuracy by an LLM for this task

szainana@stanford.edu

github.com/daneshvar-amrollahi/mbti-cs229

daneshvar@cs.stanford.edu