Lab Activity 3 Report

NLP in Psychology: Personality and Emotion Detection from Social Media Posts Using GPT & Embeddings

. Introduction

Psychological insights are often reflected in the language people use on social media. This assignment explores whether advanced Natural Language Processing (NLP) techniques can identify personality traits and emotional tendencies using text alone.

Using MBTI personality data, this project applies Word2Vec embeddings, BERT classification, and GPT-based emotional reasoning to map user text to cognitive traits and emotional state indicators. The work supports applications in mental wellbeing aligned with SDG 3.

---

2. Dataset Information

- Dataset: Myers Briggs Type Indicator(MBTI) (mbti_1.csv)
- Dataset Link:- https://www.kaggle.com/datasets/datasnaek/mbti-type

- Total Users: 8,675

- Each row: A user with combined Reddit posts

- Introversion (I) – Extroversion (E)
- Intuition (N) – Sensing (S)
- Thinking (T) – Feeling (F)
- Judging (J) – Perceiving (P)

- Labels: 16 MBTI personality types converted into four binary axes:
    - I vs E
    - N vs S

    - T vs F
    - J vs P
- Labels encoded into multi-label binary vectors suitable for machine learning.

---

3. Methodology

A) Personality Classification Using BERT

- Model: bert-base-uncased
- Problem Type: Multi-label sequence classification
- Max sequence length: 512
- Training setup:
    - 3 epochs
    - Batch size: 8
    - Binary cross entropy loss

B) Word2Vec Embedding Analysis

- Technique: Gensim Word2Vec
- Embedding dimension: 100
- Used to analyze semantic similarity of psychological concepts such as:
    - Logical reasoning (Thinking trait)
    - Emotional language (Feeling trait)

C) Emotional Context Generation Using GPT-Style Instruction

- Analysis of personality predictions
- Psychological interpretation of user language

- Supports emotion understanding along with personality traits

---

4. Results

Model Performance (BERT)

Metric	Value
Validation Accuracy	0.5850
F1-Micro Score	0.7909

F1 score indicates strong classification reliability.

Word2Vec Findings

Similarity results show clear differences in the semantic neighborhoods of:

- “logic”: reasoning, intuition, understanding
- “feel”: emotional and expressive words

Example Prediction Output

Input Sample Post:

> "I struggle to see the purpose of small talk. I spend most of my time analyzing abstract theories and trying to optimize ideas."

Predicted MBTI Type: INTP
Expected MBTI Type: INTP
Trait confidence values (logits) show strong presence of:

- Introversion (I)
- Intuition (N)
- Thinking (T)

- Perceiving (P)

---

5. Discussion

Key observations:

- Personality influences vocabulary and communication style.
- BERT is effective in detecting personality from large text samples.
- Word2Vec verifies that psychological traits are reflected in semantic clusters of words.
- GPT-based contextual reasoning provides additional mental health insights.

---

6. Real-World Applications

- Online mental health screening
- Early detection of emotional distress patterns
- Personalized communication assistants
- Hiring and educational tools tailored to cognition styles
- Behavioral monitoring for digital well-being

---

7. Conclusion

The experiment successfully:

- Classified personality traits using BERT
- Analyzed semantic language patterns using Word2Vec
- Generated emotional context using GPT-style analysis

- Demonstrated how NLP can detect psychological signals from social media posts

8. References

- Gensim Word2Vec Documentation
- MBTI Dataset (Kaggle)