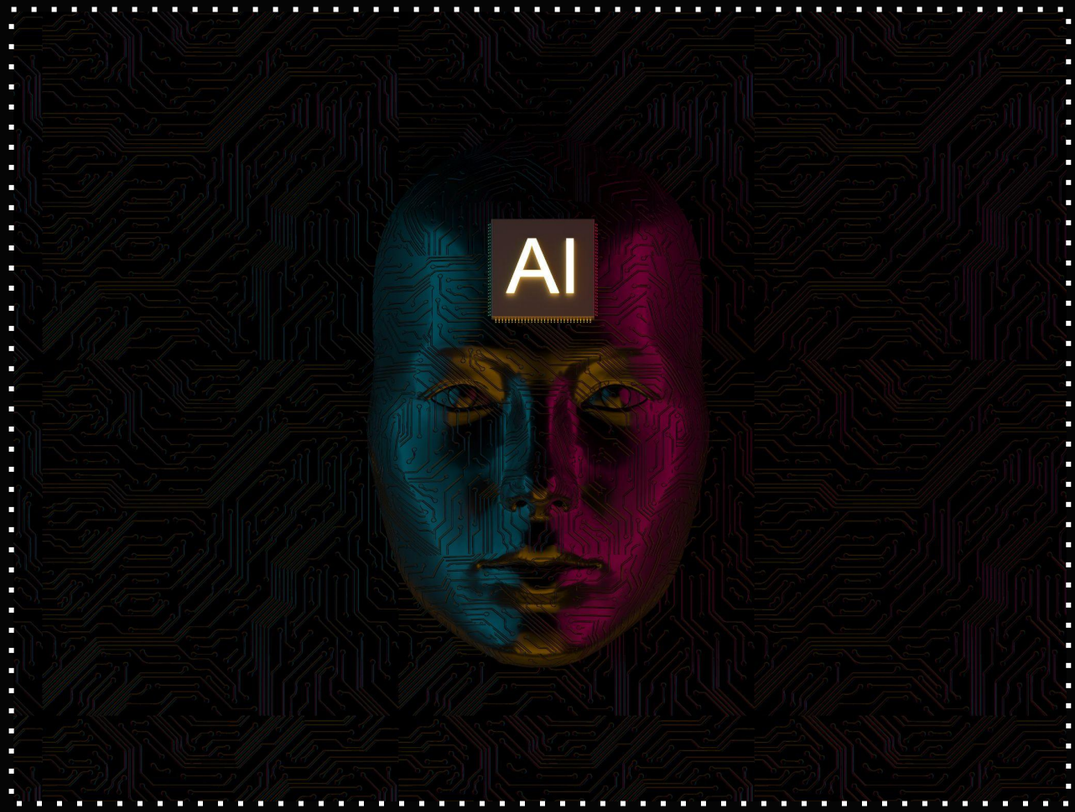
Essay Classification Problem: Student-written or LLM-generated?

CPSC444 - AI | *Prof. Hu*

Tytus Felbor | *Spring ‘25* | *May 9, 2025*



Problem Statement

Classification Goal

Identify essays as student-written or LLM-generated

Importance

Ensure academic integrity by detecting AI content

Challenges

Severe class imbalance, text input, test variability

Dataset Overview

Size & Features

1,378 essays total; text and binary labels

99.8% student-written, severe imbalance

Preprocessing

GloVe 100D embeddings averaged per essay

Variable-length texts handled

Methodology

Model

Minimal ANN with 100D input,

64-node hidden layer

Oversampling

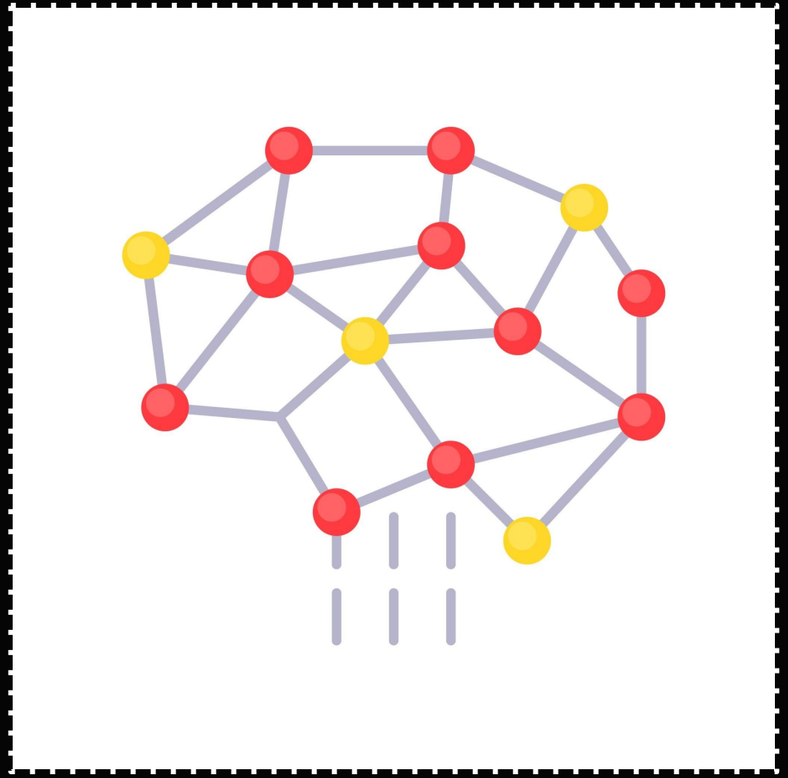
Random oversampling to 578

LLM-generated (29.6%)

Metrics

F1-Score primary

Accuracy secondary



Implementation

Tools

Google Colab, TensorFlow, NLTK, scikit-learn

Pipeline

Preprocess → Oversample → Split (80:20) → Train ANN

Tuning

Manual tuning of learning rate, batch size, dropout

4 combinations tested

Efficiency

Training time ~20–40 minutes on Colab free tier

Results

Best Model Params

Learning rate=0.001, batch size=32, dropout=0.2 or 0.3

Validation Performance

F1-Score 0.996 (threshold 0.7)

Accuracy 0.997

Insight

Excellent minority detection, dataset size limits generalization

Conclusion

Summary

Simple ANN achieved near-perfect F1 and Accuracy

Strengths

Efficient, robust preprocessing, high performance

Limitations

Small dataset, possible overfitting to oversampled data

Future Scope

1

Expand Dataset

Add more LLM-generated essays for balance

2

Model Complexity

Increase hidden layer size or add second layer

3

Hyperparameter Tuning

Test more combinations to improve performance

Any Questions?

Thank you for your attention