

Reddit Suicide Post Dataset Classification Prediction

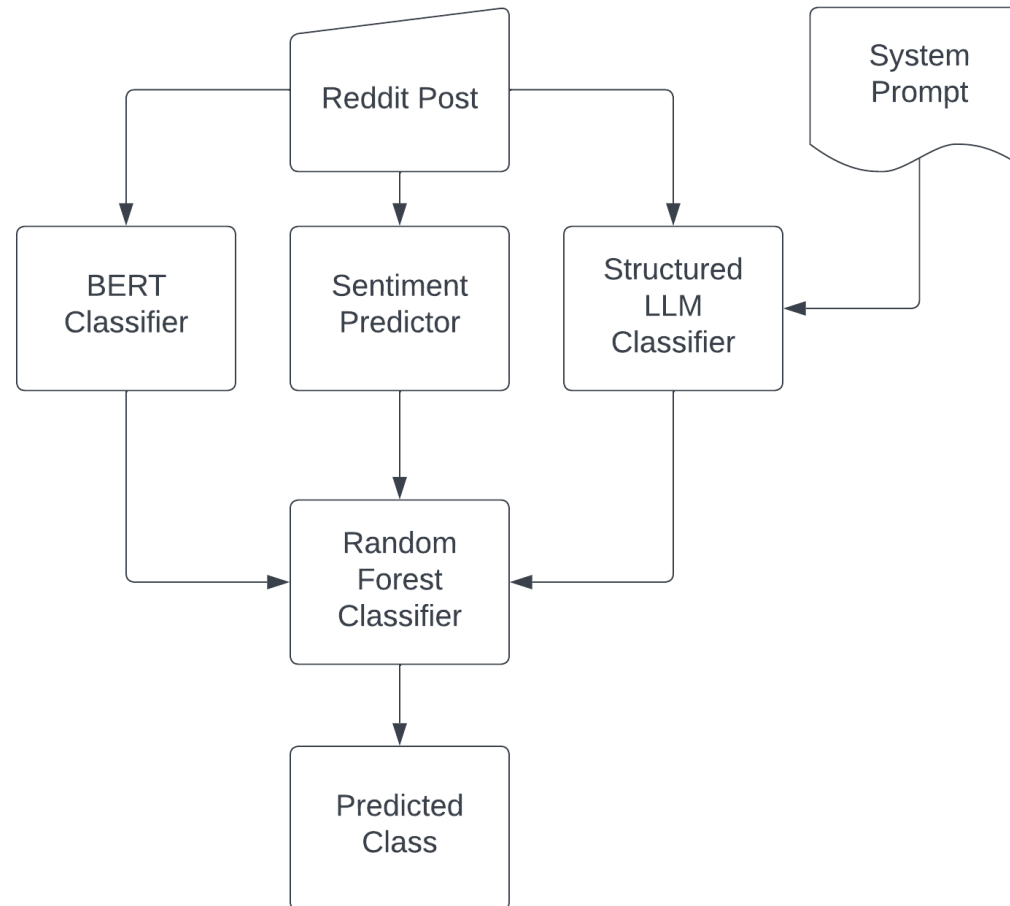
Tiernan Lindauer

10/6/2024

Approaches

Approach Used	Validation Accuracy
1. Sequential MLP w/ embedding input (5K, 30K parameter models) 20% dropout	19%
2. BERT Classifier (128 tokens max, 3 epochs)	42%
3. BERT Classifier (512 tokens max, 4 epochs)	43%
4. Sentiment + Post Length Predictor w/ Logistic Regression	27%
5. Function calling <code>gpt-4o-mini</code> (0-shot, 128 tokens max)	40%
6. Ensemble #2 + #4 + #5 via RandomForestTree	45%
7. Longformer Classifier (4096 tokens max)	WIP

Ensemble Structure



Limiting Factor for Accuracy

- Not enough training data to generalize properly, leading to high overfitting
- Very long posts (which include an entire Reddit thread) create a significant amount of data & complexity (nuisance) of the text
- BERT truncation leaves valuable information on the table
 - Specialized techniques like Longformer and BigBird can help
- Sequential MLP struggle with long sequences because they can't capture context well over long spans
- LLM is using zero-shot because of the long sequences and highly nuisanced use case
- Subjectivity in classification (Attempt/Behavior/Ideation/Indicator/Support)