Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De

Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

November 09, 2023

Table of Contents

1. Introduction
2. Problem Statement
3. Methodology
    1. Dataset Overview
    2. Preprocessing
    3. Exploratory Data Analysis
    4. Models and Evaluation
4. Future Work and Conclusion

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Introduction

This study aims to contribute to the larger goal of making mental health care more accessible and effective.

The motivation behind this research is twofold.

- It aims to advance the diagnostic capabilities for depression, enabling faster and more accurate identification of the disorder through automated means.
- It seeks to push the boundaries of NLP and ML applications in psychiatry, demonstrating how these tools can interpret the subtleties of human language to identify complex mental states.

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Problem Statement

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

The primary objective of this research is to devise an ML framework capable of discerning patterns indicative of depression within clinical interview transcripts.

The problem statement revolves around two main axes.

- The effective processing of unstructured text data to extract meaningful features and the subsequent application of various classification models to accurately predict depression.
- The overarching goal is to explore the efficacy of different ML models and to establish which among them can serve as reliable predictors for depression, based on linguistic cues.

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Dataset Overview

```
DAIC_Dataset/

        301_P.zip/
              301_AUDIO.wav
              301_video_files
              301_TRANSCRIPT.csv

        302_P.zip/
              302_AUDIO.wav
              302_video_files
              302_TRANSCRIPT.csv

        ...

        400_P.zip/
              400_AUDIO.wav
              400_video_files
              400_TRANSCRIPT.csv

        train_split.csv
        test_split.csv
```

Figure 1: Dataset Structure

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Preprocessing

Pseudocode Here is a pseudocode representation of the preprocessing routine:

```
procedure PREPROCESS_DATA
    for each batch in dataset_batches do
1       UNZIP_BATCH(batch)
        for each patient_dir in batch do
2           transcript_data <- EXTRACT_TRANSCRIPT(patient_dir)
            depression_score <- RETRIEVE_SCORE(patient_dir, split_data)
3           CONSOLIDATE_DATA(transcript_data, depression_score)
        end for
4       CLEANUP_BATCH(batch)
    end for
end procedure
```

Figure 2: Preprocessing Pseudocode

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Exploratory Data Analysis

Uniform Manifold Approximation and Projection (UMAP):
- Reduces word embeddings to 2D space, preserving data structure.
- Faster, scales well, and shows distinct clusters for words like "depression" and "anxiety".

Principal Component Analysis (PCA):
- Linear technique that orders axes by variance.
- Reveals the spread of embeddings, with words like "attended" and "jobs" significant in variance.

t-Distributed Stochastic Neighbor Embedding (t-SNE):
- Preserves local structures, revealing clusters at many scales.
- Words related to "depression" and "anxiety" appear below the x-axis, while positive factors above.

Word Cloud:
- Identifies prominent themes with words like "really," "know," "laughter," "things".

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

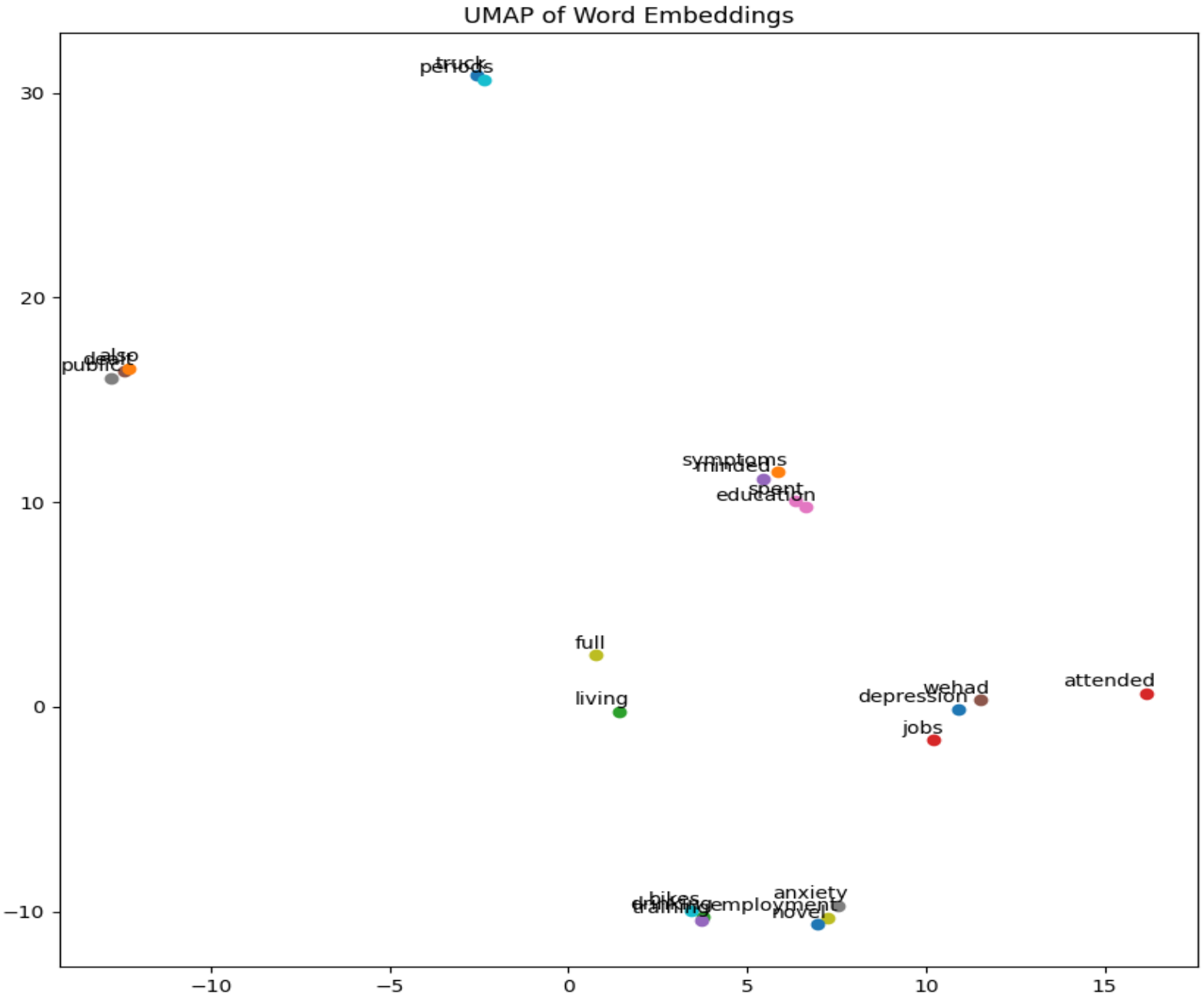
Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

UMAP



Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

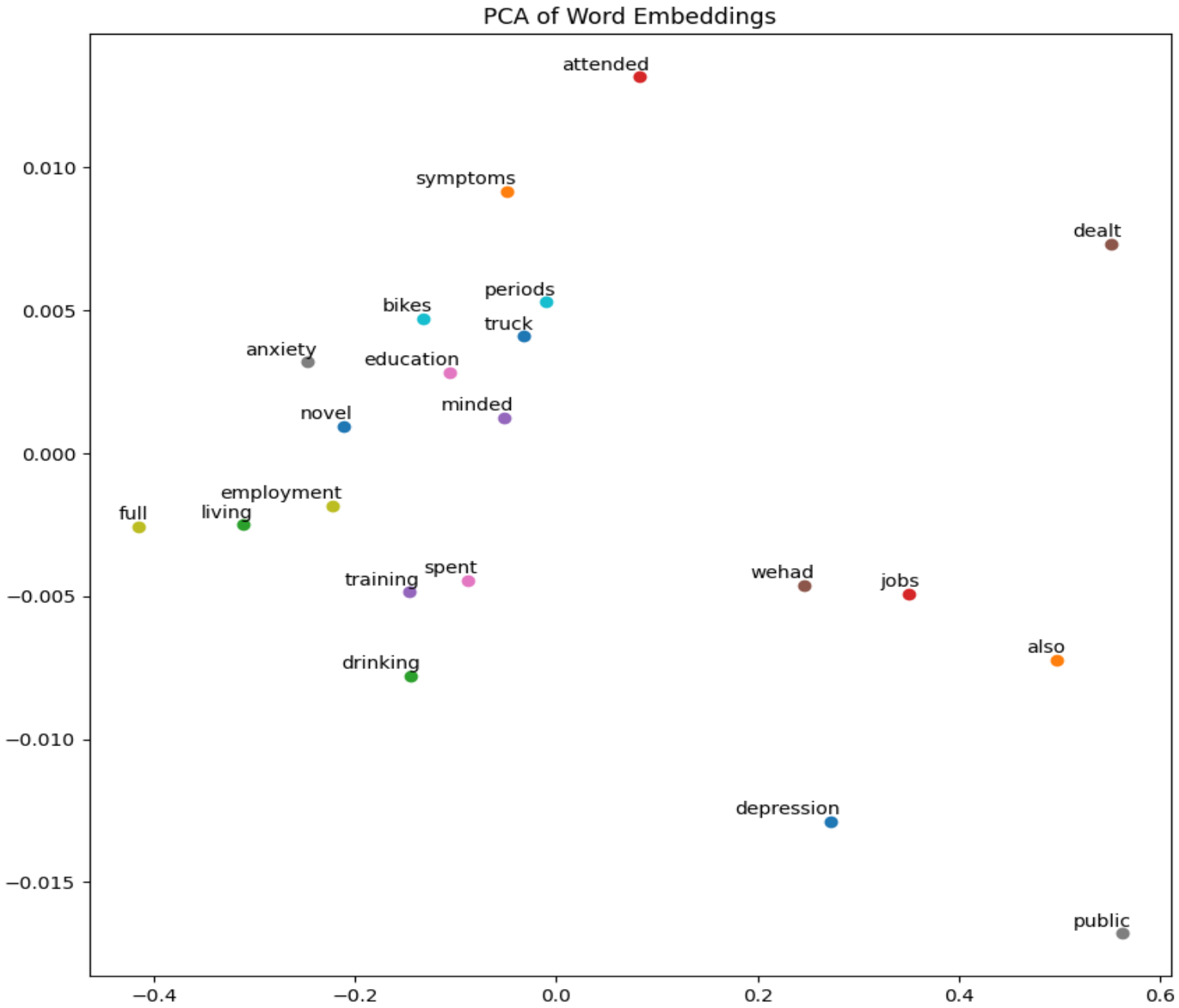
Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

PCA



Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

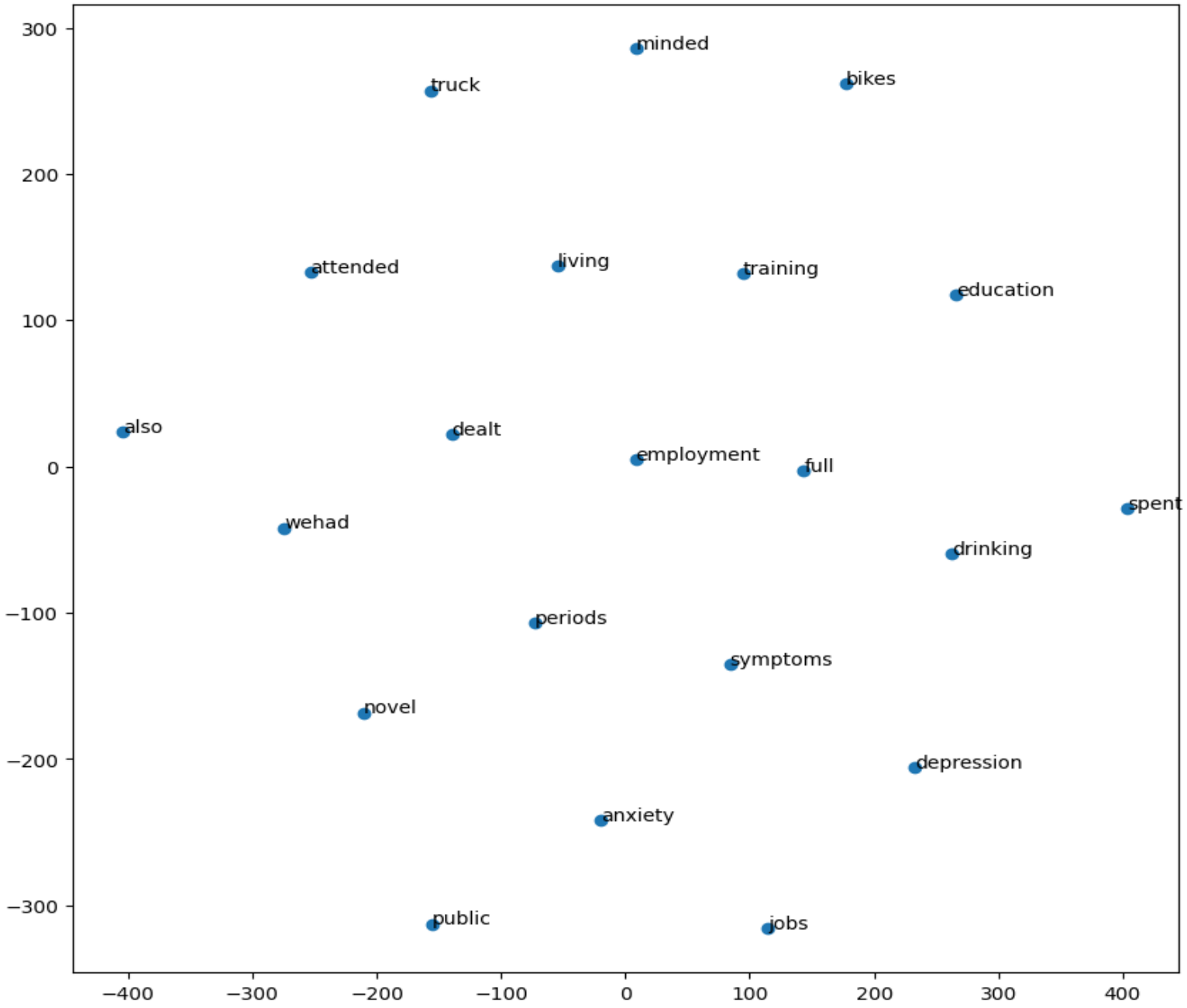
Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

t-SNE



Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Word Cloud

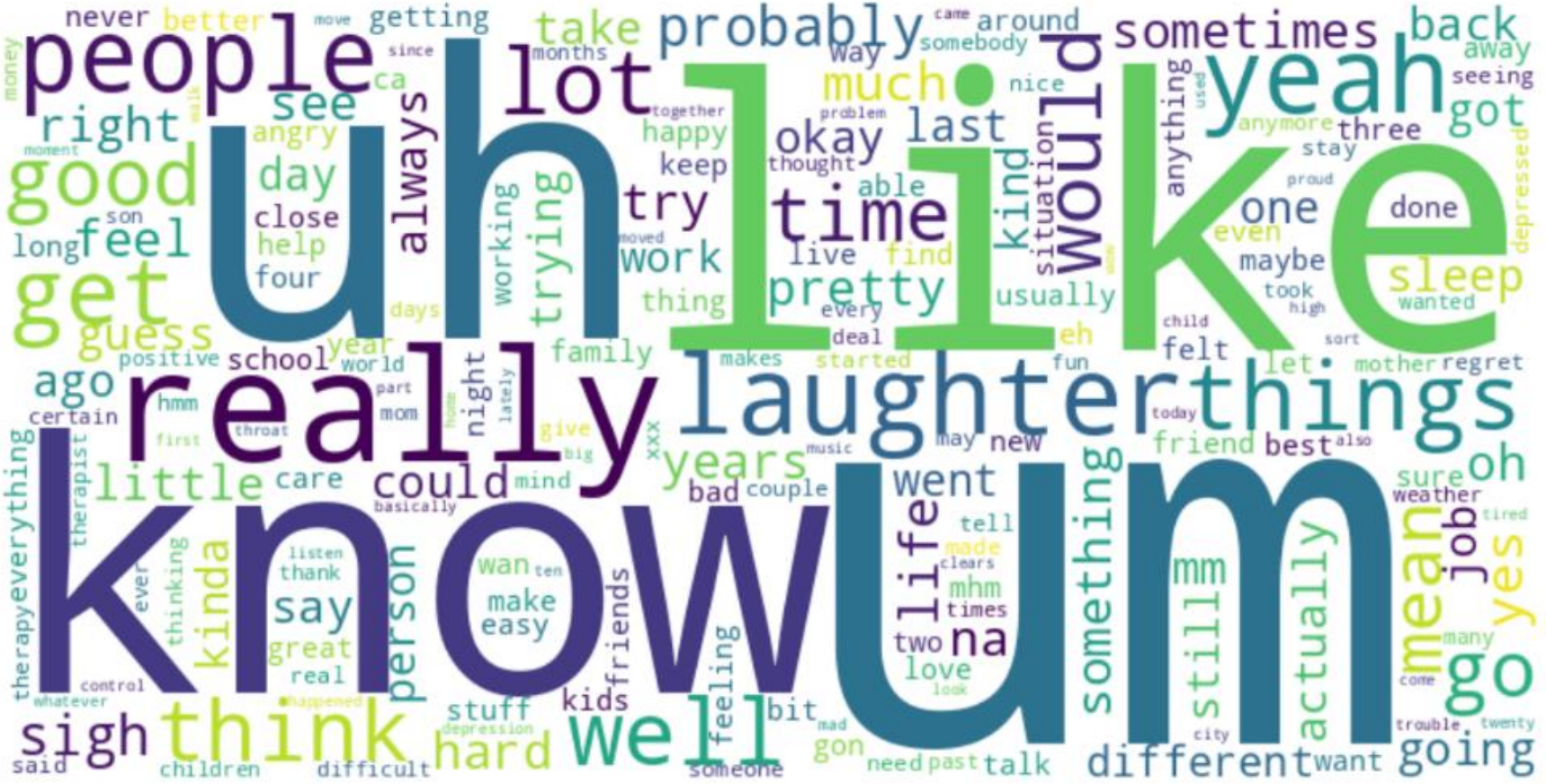


Figure 3: Most Frequent Words

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Models and Evaluation

Logistic Regression with TF-IDF:

- **Choice Reason:** Efficient with linearly separable, high-dimensional data.
- **Output & Implication:** Achieved an F1 score of 0.5, indicating moderate performance potentially due to linear assumptions and insufficient complexity for the dataset's nuances.

Multinomial Naive Bayes with TF-IDF:

- **Choice Reason:** Effective with text classification due to assumptions of feature independence and performance with discrete data.
- **Output & Implication:** F1 score of 0.56 suggests balanced precision and recall but moderate due to potential flaws in the feature independence assumption in real-world data.

Support Vector Machine (Linear & RBF Kernel) with Word2Vec:

- **Choice Reason:** Good for both linearly separable data (Linear kernel) and non-linear cases (RBF kernel).
- **Output & Implication:** RBF kernel SVM scored an F1 of 0.56 versus 0.31 for linear SVM, showing better but still average performance, possibly due to sensitivity to hyperparameters and high-dimensional feature space complexity.

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Models and Evaluation (contd…)

BERT (Bidirectional Encoder Representations from Transformers):

- **Architecture:** Utilizes bidirectional context learning, effective for learning from given transcript.
- **Fine-Tuning:** Adapted to domain specifics; crucial parameters included epochs (3), learning rate (2e-5 with warm-up and decay), and batch size.
- **Performance:** Achieved high model performance with an F1 score of 0.69, balancing training time against overfitting.

DistilBERT (Distilled BERT):

- **Architecture:** A smaller, more efficient version of BERT, removing every other attention layer through knowledge distillation. Used due to small dataset.
- **Hyperparameter Tuning:** Included a softmax temperature of 2 for optimal knowledge transfer from BERT.
- **Performance vs. Efficiency:** F1 score slightly lower at 0.56, trading off some accuracy for faster speed and reduced resource usage, suitable for resource-constrained environments.

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Models and Evaluation (contd…)

Model	Accuracy
Logistic Regression	0.50
Multinomial Naive Bayes	0.56
SVM (Linear Kernel)	0.31
SVM (RBF Kernel)	0.56
BERT	0.69
DistilBERT	0.56

TABLE 3.1: Model accuracy comparison

To evaluate the performance of the models, the F1 score was employed as a primary metric due to its balanced consideration of both precision and recall, making it more robust to class imbalances than accuracy alone.

BERT's superior context understanding capabilities, enabled by its transformer architecture, is reflected in its highest F1 score.

Challenges: The size of the dataset was too low not to overfit any model while learning adequate features at the same time. Additionally, the dataset possessed class imbalances. The limitations of the models used are intrinsically tied to their architecture.

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Models and Evaluation (contd…)

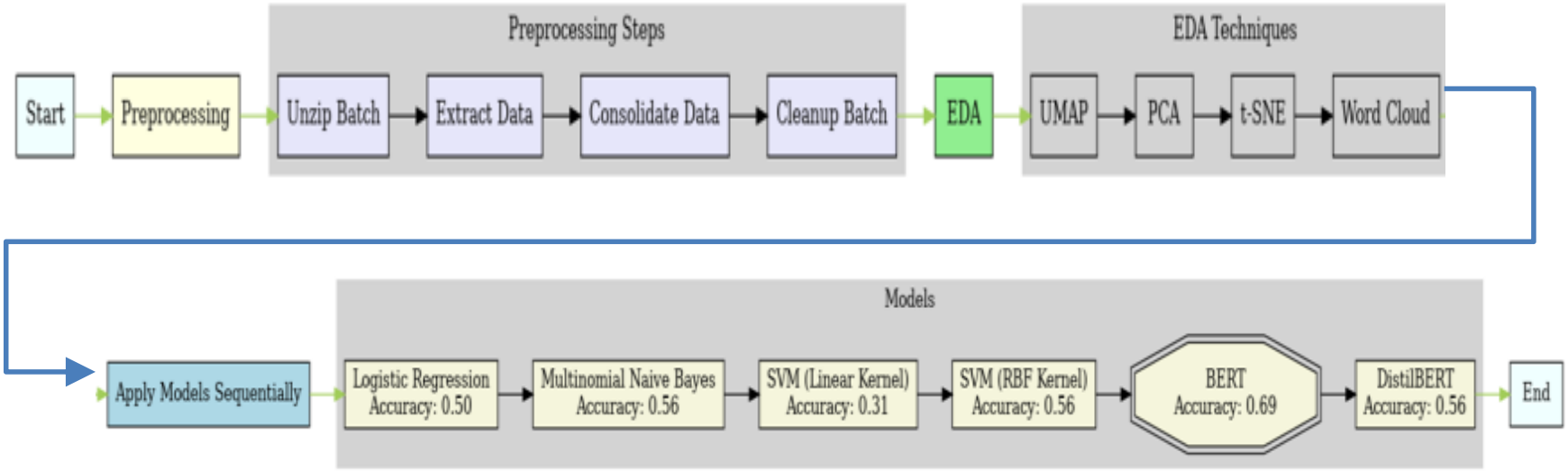


Figure 4: Process Pipeline

Machine Learning Approaches to Detecting Depression from Clinical Interview Transcripts

Srijanak De
19CS30047

Introduction

Problem Statements

Methodology
Dataset Overview
Preprocessing
Exploratory Data Analysis
Models and Evaluation

Future Work and Conclusion

Future Work and Conclusion

This study provides a detailed examination of text classification using machine learning and NLP, showcasing the strengths and limitations of models ranging from Logistic Regression to advanced neural networks like BERT.

Performance metrics indicate the superiority of deep contextualized models and the importance of feature extraction methods.

The research lays groundwork for future work focusing on unsupervised learning, sophisticated data augmentation, multilingual datasets, and efficient transformer models to enhance applicability and performance in NLP.