**CSE523 Machine Learning**

**Prof. Mehul Raval**

# **Stress Detection**

**Weekly Report**

**Date: 25-03-2023**

| **Name** | **Enrolment Number** |
| --- | --- |
| Mohsinali Vijapura | AU2040080 |
| Manan Vadaliya | AU2040264 |
| Pruthviraj Dodiya | AU2040175 |
| Mitsu Sojitra | AU2040157 |

**Task performed this week:**

* Explored different feature extraction techniques such as Bag of Words, TF-IDF, and Word Embeddings to convert raw text data into numeric vectors.
* Preprocessed the data by removing stop words, stemming, and converting text to lowercase.
* Trained several machine learning models such as Logistic Regression, Naive Bayes, Random Forest, and Support Vector Machines using the preprocessed data and evaluated their performance using cross-validation.
* Tuned hyperparameters of the best-performing models using grid search to improve their performance.

**Challenges and Solutions**

* One of the main challenges was dealing with imbalanced data where the number of samples in the stressed class was much smaller than the number of samples in the non-stressed class. We addressed this issue by using oversampling techniques such as Synthetic Minority Over-sampling Technique (SMOTE) to generate synthetic samples for the stressed class.
* Another challenge was selecting the optimal feature extraction technique and machine learning model for our problem. We addressed this by experimenting with multiple techniques and models and selecting the ones that gave the best performance.

**Task to perform end in next week**

* Evaluate the performance of the best model on the test set and report the final accuracy, precision, recall, and F1 score.
* Analyze the feature importance to gain insights into which words or phrases are most indicative of stress.
* Write the final report summarizing the project, including the problem statement, methodology, results, and future work.