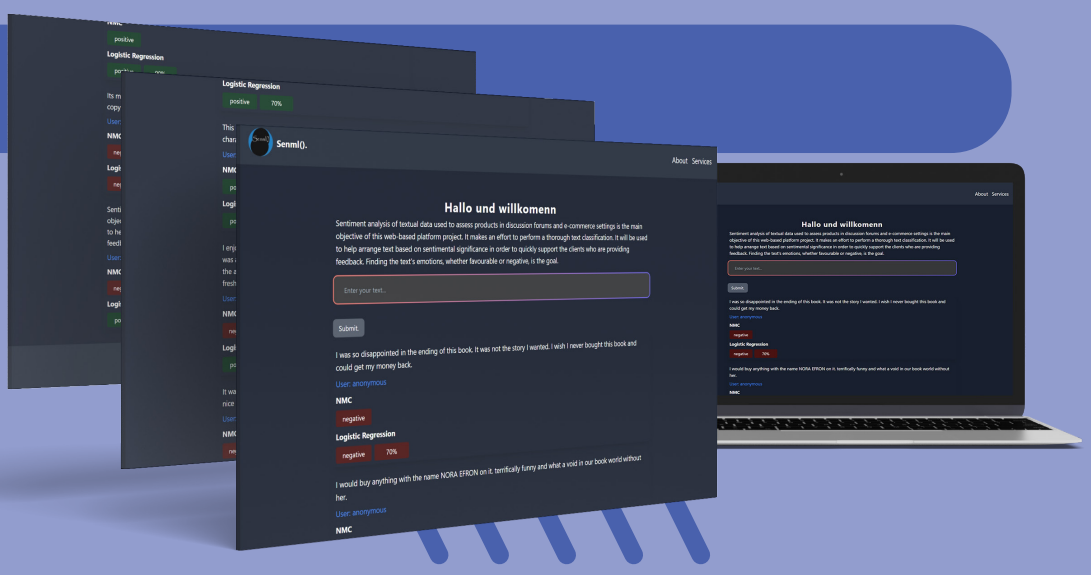
# Sentiment Analysis Text Classification

& model comparison

with Mlpack C++

**Student:** John Kelechukwu Obi - 20140237
BSc. Computer Engineering
**Supervisor:** Asst. Prof. Dr. Cem Kalyoncu



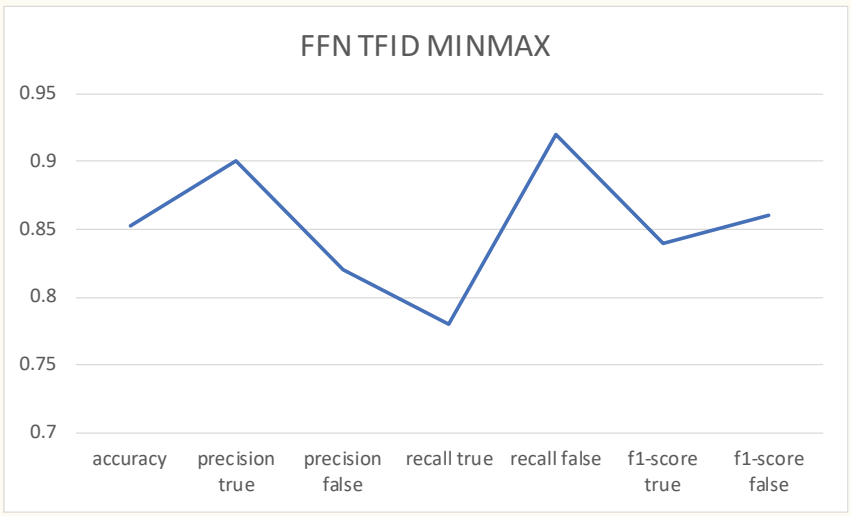

## Introduction

A comprehensive yet simple implementation of a sentiment analysis system thats gets textual inputs and classifies them on their sentimental weights to determine if they were positive or negative.

## Methodological process:

1. **Preprocessing,**
   Json parsing and stopword removal, stemming, arbitrary character removal, Thresholding, Equalizing and limiting
2. **Feature extraction,**
   Bag of words encodings, TFID encodings, Scalar methods utilization,
3. **Modelling,**
   FFN, Logistic Regression, Random forest, Nearest mean classifer, k-nearest neighbor
4. **Evaluation,**
   Accuracy, Precision, Recall, F1-score
5. **Deployment,**
   Containerisation

## Graph showing metrics from one of the best perfoming model





## Risks and Ethics

**Text model accuracy transparency,**
The showcase of the model's accuracy and introduction of bias helps keeps the model from stray away from its intended path.

**Data Security/Privacy,**
The texts are gotten from random users and there are no links of the texts that points to the actual user.

**Scalability,**
This project is actual quite very scalable as it requires less to add new features.

**Deployment,**
Scalable and yet easily deployable. Two in one bundle with zero environmental issues.