Twitter Sentiment Detection

Researcher: Loyiso Jiya
Co-researcher: Daiyaan Sataar
Supervisor: M Ghaziasgar
Co-supervisor: R Dodds

- ➔ Recap
- ➔ Implementation
- ➔ Training
- ➔ Project progress
- ➔ Demo

Introduction

- Recap
- Implementation
- Training
- Project progress
- Prototype demonstration

➔ Recap
➔ Implementation
➔ Training
➔ Project progress
➔ Demo

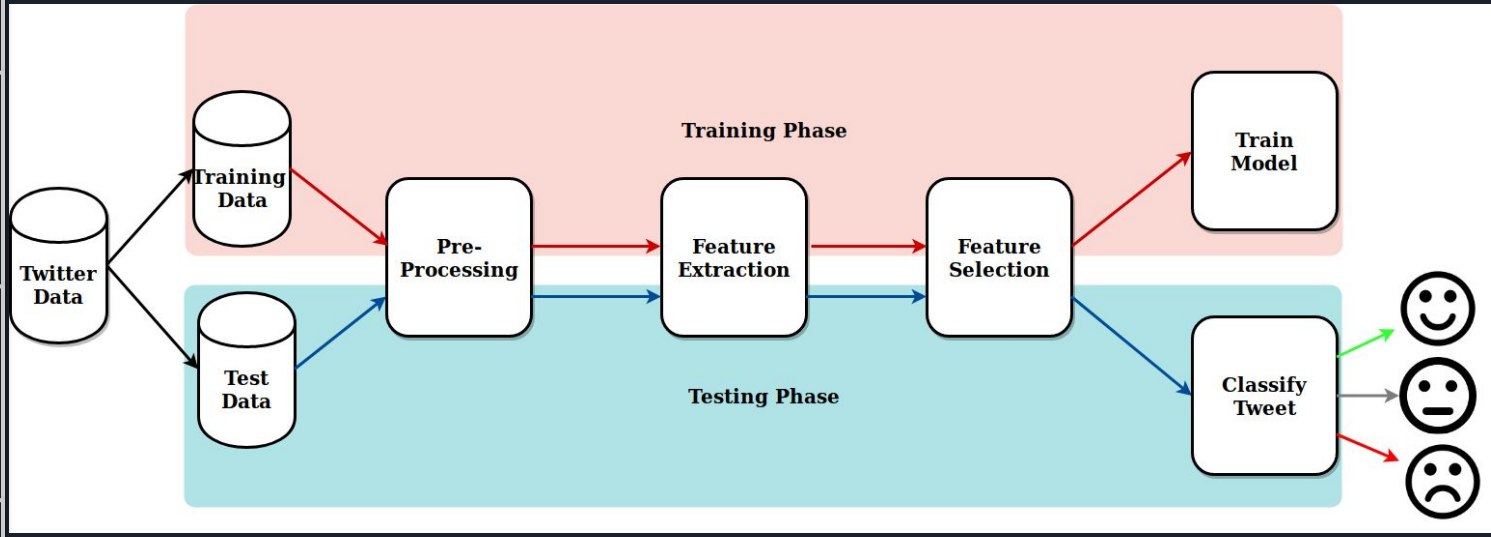
Recap

- Description and Aim:
  - Determining overall sentiment/feeling a text conveys
  - Attempts to classify the emotion or sentiment conveyed by text
  - Leverage social media data to gain insights previously unachievable
  - Mental illness detection [1], movie review sentiment detection [2]
  - Companies gauge how their products or services are received by consumers

- Term Two Work:
  - Pre-processed data
  - Developed high-level detail of implementation phase

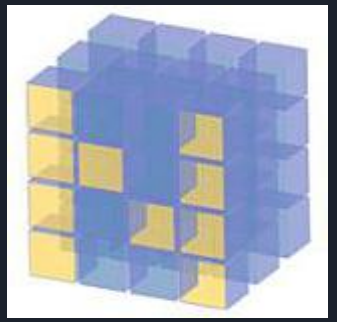
➔ Recap
➔ Implementation
➔ Training
➔ Project progress
➔ Demo

Recap



- Introduction
- Implementation
- Training
- Project progress
- Demo

Implementation



➔ Introduction
➔ Implementation
➔ Training
➔ Project progress
➔ References

Implementation

- Text preprocessing:
  - Pandas
  - BeautifulSoup
  - NLTK
  - Scikit learn
- Feature Extraction:
  - Term Frequency - Inverse Document Frequency
- Feature Selection:
  - Information gain using *mutual_classif_info*
- Demo creation:
  - Tkinter
  - Classes to represent each frame of prototype

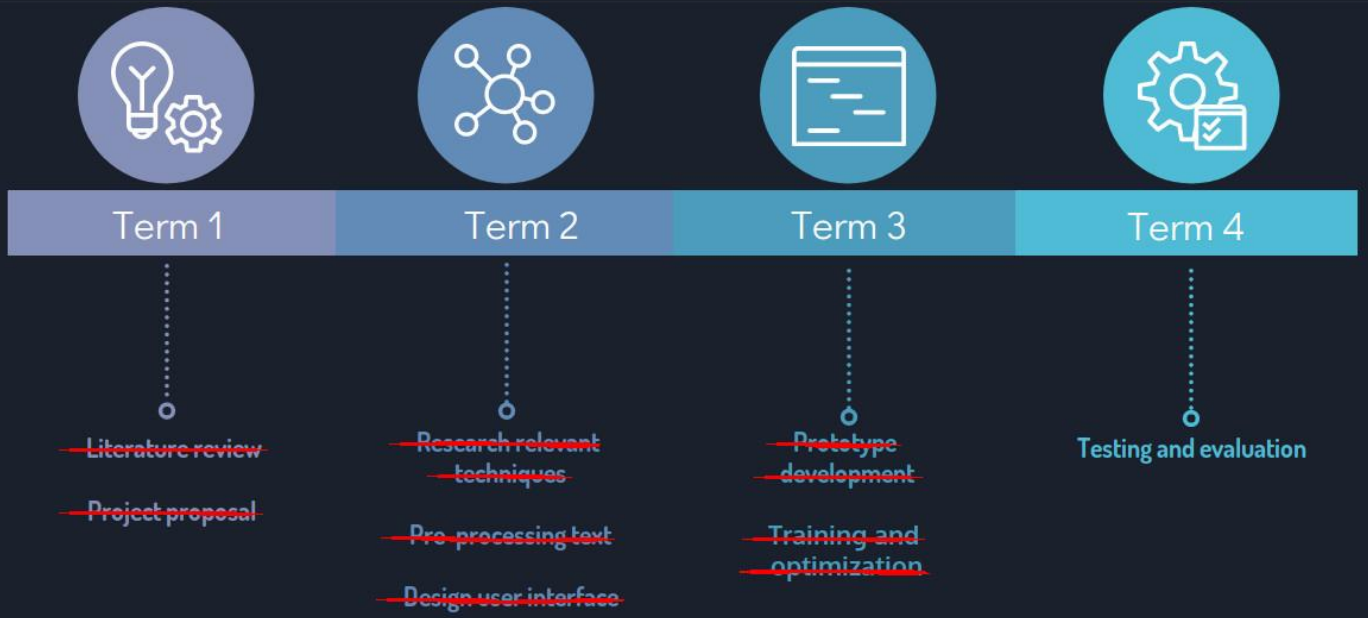
➔ Introduction
➔ Implementation
➔ Training
➔ Project progress
➔ Demo

Training

- Hardware:
  - Lab PC with GPU and 16GB RAM
- Model training:
  - Xgboost model trained
  - 250K samples out of 1.6M used for training
  - Grid search performed for optimum parameters
- Preliminary results:
  - F1-score 75% without feature selection
  - F1-score 76% with information gain feature selection

- ➔ Recap
- ➔ Implementation
- ➔ Training
- ➔ Project progress
- ➔ Demo

Project plan



- ➔ Recap
- ➔ Implementation
- ➔ Training
- ➔ Project progress
- ➔ **Demo**

Demo

Thank you.

Any Questions?