



Twitter Sentiment Detection

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

→ Introduction

→ Architecture

→ Pre-processing

→ Project plan

→ References

Introduction

- Project description
- Architecture
- Pre-processing
- Project plan
- Prototype demo

→ Introduction

→ Architecture

→ Pre-processing

→ Project plan

→ References

Introduction

- Description:
 - Text classification problem
 - Attempts to classify the emotion or sentiment conveyed by text
- Aim:
 - 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

→ Introduction

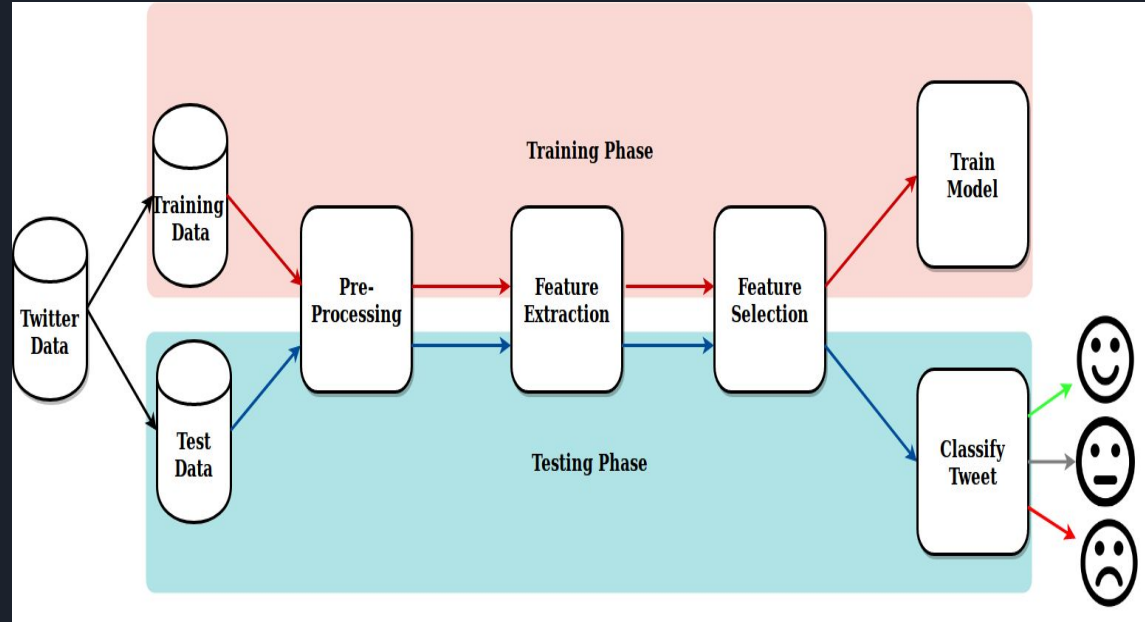
Architecture

→ Architecture

→ Pre-processing

→ Project plan

→ References



→ Introduction

→ Architecture

→ Pre-processing

→ Project plan

→ References

Pre-processing

- Data set :
 - Stanford Twitter corpus[3]
 - the polarity of the tweet (e.g. 0 = negative, 2 = neutral, 4 = positive)
 - the id of the tweet (e.g. 2087)
 - the user that tweeted (e.g. robotickilldozr)
 - the text of the tweet (e.g. Lyx is cool)

→ Introduction

→ Architecture

→ Pre-processing

→ Project plan

→ References

Pre-processing

- **Tokenization:**
 - **Text:** After sleeping for four hours, he decided to sleep for another four
 - **Result:** { 'After', 'sleeping', 'for', 'four', 'hours', 'he', 'decided', 'to', 'sleep', 'for', 'another', 'four' }
- **Stopword removal:**
 - **Text:** This is a story of the cow that jumped off a building to disprove gravity.
 - **Result:** ['This', 'story', 'cow', ' , 'jumped', 'building', 'disprove', 'gravity',]

→ Introduction

→ Architecture

→ Pre-processing

→ Project plan

→ References

Pre-processing

- Stemming, etc.
 - **Text:** Leading begins with following.
 - **Result:** Lead begin with follow.
- Spelling correction:
 - **Text:** 'amaaziing', 'caar', 'mussage', 'hte'
 - **Result:** 'amazing', 'car', 'message', 'the'
- End goal is to have a feature vector that has a meaningful contribution to the prediction

→ Introduction

→ Architecture

→ Pre-processing

→ Project plan

→ References

Project plan

- Term 2:
 - Research relevant techniques
 - Pre-process text
 - Pre-processing demonstration
- Term 3:
 - Prototype development
 - Training and optimization of models
- Term 4:
 - Testing and evaluation
 - Integrate model with UI

→ Introduction

→ Architecture

→ Pre-processing

→ User interface

→ References

References

1. S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, “Detecting depression and mental illness on social media: an integrative review,” *Current Opinion in Behavioral Sciences*, vol. 18, pp. 43–49, 2017.
2. K. Chakraborty, S. Bhattacharyya, R. Bag, and A. E. Hassanien, “Comparative sentiment analysis on a set of movie reviews using deep learning approach,” in *International Conference on Advanced Machine Learning Technologies and Applications*. Springer, 2018, pp. 311–318.
3. T. Inc. (2018) Quarterly results. [Online]. Available: <https://investor.twitterinc.com/financial-information/quarterly-results/default.asp>



Thank you.

Any Questions?