

# **School of Computer Science and Engineering**

## Rumour Detection on Online Social Networks Based on Sentiment Analysis

**Project Plan**

**Student**

xxx

**Supervisor**

A/P YEO CHAI KIAT

1. **Project Objective**

This project aims to investigate and develop novel methodologies to enable detection of rumors from social media sources via sentiment analysis.

1. **Project Description**

The 21st century has seen the advent and tremendous growth of the use of social media in the world. Today, social media is ever-pervasive and is often used in tandem with traditional communication channels, such as print media and broadcast media, to broadcast information to the general public. Such information can include latest developments in war-torn countries, recent product launches, or even rumors of recent deaths of people of significance. As people have been increasingly reliant on social media streams for the latest news, combined with the fact that significant portions of content on social media streams are user-contributed, the onus has been on the consumers to weed out rumors from truthful pieces of information. Thus is the raison d’etre of this project – to automate detection of rumors on social media streams through the use of sentiment analysis techniques and machine learning methodologies.

1. **Project Scope**

The project encompasses the following tasks:

* Information Harvester Development

An information harvester will be developed to automate the collection and storage of relevant social media content from social media websites such as Twitter and Instagram.

* Literature Review

Literature review will be undertaken, taking cues from related academic papers in sentiment analysis, machine learning, psychology, and sociology.

* Feature Selection & Engineering

Experiments in extracting optimal features and creating new features will be performed in preparation for subsequent stages of the project.

* Sentiment Analysis using Machine Learning Techniques

Sentiment analysis techniques will be used in tandem with machine learning techniques to achieve the project’s core functionality of being able to detect rumors. Exploration into various sentiment analysis techniques and sentiment analysis libraries such as Natural Language Toolkit, OpenNLP, and Stanford NLP will be experimented with to provide a good output for the machine learning classifiers to use. Exploration into various machine learning techniques such as the use of neural networks, decision trees, and naïve bayes classifier will be experimented with to provide a good classification performance for detecting rumors. Full System Integration

A system, consisting of all previously mentioned components that have been created, will be developed into an integrated system, to provide an easy-to-use interface for users to easily gain insights of the achievements of the project.

* Report Writing & Oral Presentation

An Interim Report, a Final Report, and an Oral Presentation will be prepared, as per the Final Year Project requirements.

1. **Project Schedule**

The following Gantt chart details the planned project schedule.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | SEMESTER 1 2016/2017 | | | | | | | | | | | | | | | | | | | | SEMESTER 2 2016/2017 | | | | | | | | | | | | | | | | | |
| **Date Task** | **AUG** | | | | **SEP** | | | | **OCT** | | | | **NOV** | | | **DEC** | | | | **JAN** | | | | **FEB** | | | | **MAR** | | | | **APR** | | | **May** | | | |
| Project Planning |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **E X A M** | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **E X A M** | |  |  |  |
| Information Harvester Development |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Literature Review |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Feature Selection & Engineering |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Feature Selection R&D |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Feature Engineering R&D |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Sentiment Analysis using Machine Learning Techniques |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Sentiment Analysis Techniques R&D |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Machine Learning Techniques R&D |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Result Testing and Verification |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Interim Report Writing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Full System Integration |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Final Report Writing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Oral Presentation |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |