

INFO3406 Project Stage 1

Zhu Linzi Zhou Nic

September 2018

Contents

| | | |
|---|---------------------|---|
| 1 | Section 1: Problem | 2 |
| 2 | Section 2: Approach | 2 |
| 3 | Section 3: Data | 2 |
| | Appendices | 3 |
| A | Graphs | 3 |
| B | Something else | 3 |

1 Section 1: Problem

Words convey emotions, feelings and attitude. Different words convey these affects to varying degrees. The problem this project is aiming to solve is seeing if there is a link between titles of Movies and how succesful they are.

The questions we will be answering in Stage 2 are:

1. Is there are link between the intensities of a lexicon being used in the title of a movie and the critical acclaim?
2. Is there a link between movie titles and movie success based on the emotion the leixon used in their title is aiming to convey?

2 Section 2: Approach

The approach we are going to be taking in Stage 2 is performing statistical analysis on our dataset. We will attempt to answer the questions we set out by doing explaratory factor analysis and t-tests.

3 Section 3: Data

The data was acquired for the lexicons from <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>. The da

The tools we used were:

- python
- pgAdmin4
- SPSS
- Github
- weka

The Python programming language was used to clean the data. The tables were created using a combination of pgAdmin4; to create the tables and psql; to copy the data from the CSV files into the tables. The data from the lexicon dataset did not readily come in CSV form. As such python was used to clean the data and load it onto the our database.

Appendices

A Graphs

Figures of graphs are going to go here.

B Something else

Things that are not graphs will go here.