Fine-Grained Sentiment Analysis on Financial Microblogs

2018 NLP Project 1

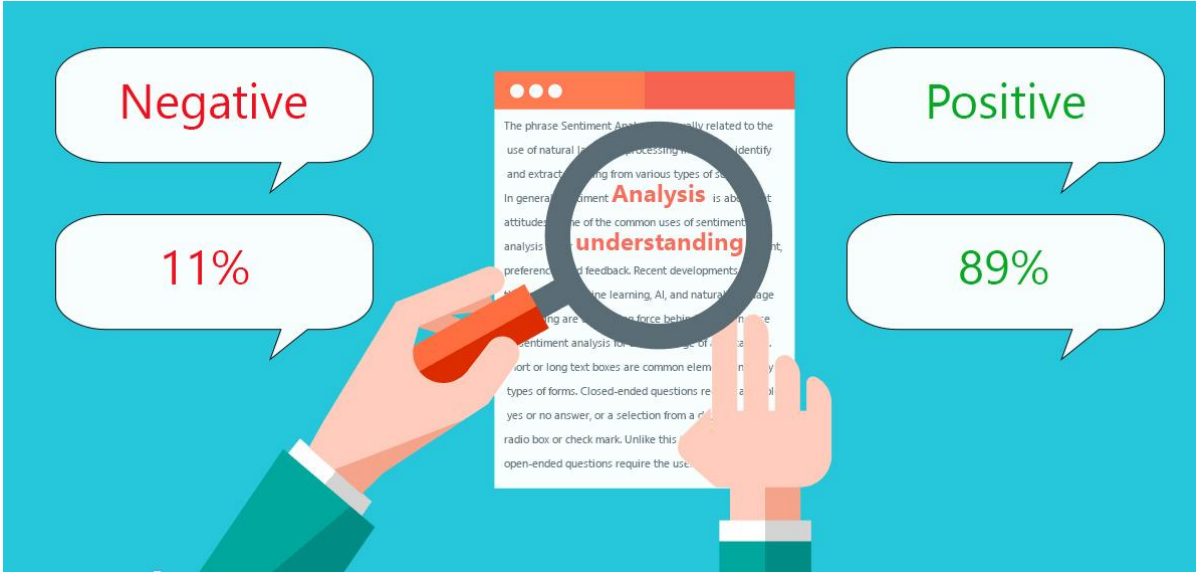陳信希 教授
顏安孜 陳重吉 {d04922005, d05922016}@ntu.edu.tw

CONTENTS

PART ONE
Task Description

PART ONE Task Description

Example:
Tweet: *$TSLA if $249.84 breaks we see $245 then $2*
Cashtag: *$TSLA*

Predict Sentiment Score = -0.519

Financial Microblog
➔ StockTwits

Cashtag
➔ $AAPL = Apple Inc.

Sentiment Analysis
➔Bullish/Bearish

Fine-Grained Sentiment Analysis
➔ Score between -1 to 1

PART TWO
Data

Test Set

634 instances

Json format
"tweet" : analyzed tweet
"target" : targeted cashtag
"snippet" : key snippet for
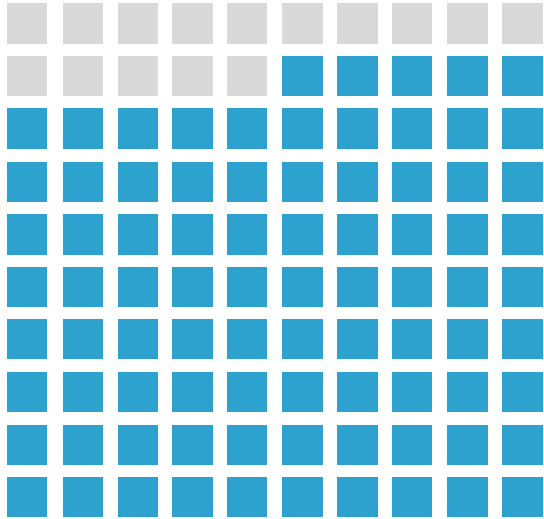targeted cashtag
"sentiment" : sentiment score

Training Set

1396 instances

PART THREE
Evaluation

PART THREE Evaluation
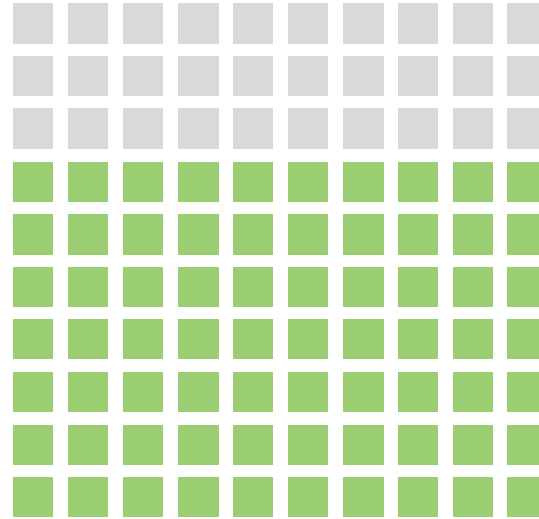
$ **F1**

3-Classes

Classify into
Bullish/Bearish/Neutral
Calculate
Micro-average F1 &
Macro-average F1

$ **MSE**

Fine-Grained Score

Calculate
Mean Squared Error

1. F1: http://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html
2. MSE: http://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_squared_error.html

PART FOUR
Report

PART FOUR Report

Up to 6 pages

1. Name and ID
2. Division of work
3. Methods
4. Evaluation
5. Discussion
6. Conclusion

Report

Describe

Write the proper comment for each part and function

Code

5 minutes

State your idea, methodology, evaluation and conclusion clearly and logically

Presentation

Proportion

Report: 60%
Presentation: 30%
Performance: 10%
Bonus: 1st~3rd 10%
4th~6th 5%

Grading Policy

*Written in Chinese or English (depend on your native language)
*Bonus depends on MSE results

Submission

Project1_team_<team number>.zip
    Report_team_<team number>.pdf
    Code_team_<team number>
        Readme(description of each script)
           $script_1$
      ...
           $script_n$

Ex:
Project1_team_0.zip
    Report_team_0.pdf
    Code_team_0
        Readme.txt
        functions.py
        main.py

SCHEDULE

Submit via CIEBA

Delay: -5% per day

Should you have any question, please let TAs know.

5/10 **Presentation in class**

5/9
Slide Submission Due

5/4 **Report & Code**
Submission Due

3/29
Project 1 release

Resource

NTUSD-Fin: http://www.nlg.csie.ntu.edu.tw/nlpresource/NTUSD-Fin/

Loughran and McDonald Sentiment Word Lists: https://www3.nd.edu/~mcdonald/Word_Lists.html

SentiWordNet: http://sentiwordnet.isti.cnr.it/

Sentiment and emotion lexicons: http://saifmohammad.com/WebPages/lexicons.html

Reference

Semeval 2017 Task 5: http://alt.qcri.org/semeval2017/task5/

S17-2138~S17-2155: http://www.aclweb.org/anthology/S/S17/

Learning Stock Market Sentiment Lexicon and Sentiment-Oriented Word Vector from StockTwits:
http://www.aclweb.org/anthology/K17-1031

THANK YOU

HAVE FUN!