AN ITERATIVE METHOD OF SENTIMENT ANALYSIS FOR RELIABLE USER EVALUATION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Jingyi Hui

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2019

Purdue University

Indianapolis, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF THESIS APPROVAL

Dr. Shiaofen Fang, Chair

Department of Computer and Information Science

Dr. Arjan Durresi

Department of Computer and Information Science

Dr. Yuni Xia

Department of Computer and Information Science

Approved by:

Shiaofen Fang

Head of the School Graduate Program

To my parents who support me with selfless and unconditional love.
To my husband Huang for being here with me throughout the entire program.
To my baby Mia without whom this work would have been completed one year earlier...

ACKNOWLEDGMENTS

I wish to express my greatest gratitude to my advisor Dr.Shiaofen Fang for giving me the opportunity to start my graduate study and a chance to work with him. This work would not have been possible without his encouragement, guidance and support.

I would also like to thank Dr.Arjan Durresi and Dr.Yuni Xia for serving in my graduate committee. Their valuable advice, timely reviews and thoughtful insights helped me during the creation of this document.

Finally, I would like to specially thank Yefeng Ruan who provided assistance with the dataset used in this research.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
LIST OF FIGURES	ix
ABBREVIATIONS	x
GLOSSARY	xi
ABSTRACT	xii
1 INTRODUCTION	1
1.1 Background	1
1.1.1 Twitter	3
1.1.2 Apple Inc.(AAPL)	4
1.1.3 Yahoo! Finance	4
1.2 Research Objectives	4
1.3 Assumptions	5
1.4 Limitations	5
1.5 Delimitations	5
1.6 Organization	6
2 LITERATURE REVIEW	7
2.1 Usage of Social Media Information Mining	7
2.2 Credibility & Trust of information in SNSs	9
2.3 Existing Trust Management Framework of Social Network Data	10
2.4 Visual Analytics	11
3 FRAMEWORK AND METHODOLOGY	13
3.1 Pearson Correlation Coefficient	14
3.2 Iterative Trust Evaluation Method	14
3.2.1 Prediction with Accumulated Trust Coefficient	15
3.3 Visualization for Denoising	16
4 DATA EXPLORATION AND ANALYSIS	19
4.1 Data Collecting	19
4.2 Data Distribution	20
4.3 Data Preprocessing	22
4.3.1 Sentiment for tweets	22
4.3.2 Interpolation of Datasets	23

Page

4.3.3 Derivation of Stock Price 23
4.3.4 Normalization of Datasets 24

5 EXPERIMENTS, RESULTS, AND EVALUATION 27
5.1 Experiments 27
5.1.1 Active User Filtration 27
5.1.2 Data Spliting for Test 28
5.2 Model Evaluation 29
5.3 Iterative Method with Visualization Aided Denoising 32
5.4 Comparison with Other Trust Index 34

6 CONCLUSION 39
6.1 Summary 39
6.2 Discussion 39
6.3 Future Work 40

REFERENCES 42

LIST OF TABLES

Table — Page

3.1 Description of Social Sentiment Dataset for Trust Evaluation 14

3.2 Description of Ground Truth Dataset for Trust Evaluation 15

4.1 Description of Tweet Dataset 19

4.2 Description of AAPL Stock Price Dataset 20

5.1 Training Data & Result with $Threshold : 150$ 29

5.2 Average Training Results with Different Thresholds 30

5.3 Prediction Result with Threshold: 150 32

5.4 Segmented Training Data & Result with $Threshold : 150$ 35

5.5 Comparison between Iterative Evaluation and Follower Numbers as Trust Scores $Threshold : 150$ 37

LIST OF FIGURES

Figure Page

1.1 Most popular social networks worldwide as of April 2019, ranked by number of active users (in millions) 2

1.2 Average time spent engaging with/connected to social network/service during a typical day 2

1.3 Percentage of respondents trust each publisher 3

3.1 The Iterative Trust Evaluation Model with Visualization Aided Segmentation 13

3.2 Example: Visualization for an user's sentiment and ground truth. The abnormal part is selected by a red rectangle. 17

4.1 Distribution of User Number Grouped By Number of Tweets Posted by Each User (Y axis is logarithmic scaled with base 10) 21

4.2 Distribution of Follower Number for Users (Y axis is logarithmic scaled with base 10) 21

4.3 Close Price of AAPL 01/01/2015 - 08/31/2015 22

4.4 Example: User Sentiment & Stock Price Before Normalization 25

4.5 Example: User Sentiment & Stock Price After Normalization 25

5.1 The relation between number of active accounts and thresholds 27

5.2 The configuration of experiments data for training and testing 28

5.3 The relation between thresholds and final correlations after training . . . 30

5.4 The Comparison between Original, Trained, and Predicted Correlations . 31

5.5 Segmented User Sentiment & AAPL Stock Price Derivation 33

5.6 The Comparison between Original, Trained, and Predicted Correlations with Segmentation 36

5.7 Distribution of Number of Followers with $threshold : 150$ 36

5.8 Comparison between Iterative Evaluation and Follower Numbers as Trust Scores 38

ABBREVIATIONS

SNS	social network site
AVG	average
WSJ	Wall Street Journal
UGC	user-generated content
LDA	latent Dirichlet allocation
NLP	Natural Language Processing
GIS	Geographical Information System
LIWC	Linguistic Inquiry and Word Count
GPOMS	Google-Profile of Mood States
DJIA	Dow Jones Industrial Average
ELM	Elaboration Likelihood Model
PCC	Pearson correlation coefficient
CoE	Coefficient

GLOSSARY

Sentiment	a view of or attitude toward a situation or event; an opinion
Ground Truth	refer to information provided by direct observation (i.e. empirical evidence)
Timestamp	the time at which an event is recorded by a computer, not the time of the event itself
Naive Bayesian	a simple, yet effective and commonly-used, machine learning classifier.

ABSTRACT

Hui, Jingyi M.S., Purdue University, August 2019. An Iterative Method of Sentiment Analysis for Reliable User Evaluation. Major Professor: Shiaofen Fang.

Benefited from the booming social network, reading posts from other users over internet is becoming one of commonest ways for people to intake information. One may also have noticed that sometimes we tend to focus on users provide well-founded analysis, rather than those merely who vent their emotions. This thesis aims at finding a simple and efficient way to recognize reliable information sources among countless internet users by examining the sentiments from their past posts.

To achieve this goal, the research utilized a dataset of tweets about Apples stock price retrieved from Twitter. Key features we studied include post-date, user name, number of followers of that user, and the sentiment of that tweet. Prior to making further use of the dataset, tweets from users who do not have sufficient posts are filtered out. To compare user sentiments and the derivative of Apples stock price, we use Pearson correlation between them for to describe how well each user performs. Then we iteratively increase the weight of reliable users and lower the weight of untrustworthy users, the correlation between overall sentiment and the derivative of stock price will finally converge. The final correlations for individual users are their performance scores. Due to the chaos of real world data, manual segmentation via data visualization is also proposed as a denoise method to improve performance. Besides our method, other metrics can also be considered as user trust index, such as numbers of followers of each user. Experiments are conducted to prove that our method out performs others. With simple input, this method can be applied on a wide range of topics including election, economy, and job market.

1 INTRODUCTION

This chapter gives an overview to this research and the paper, and describes the motive of the research derived from daily life. It points out the significance of recognizing useful information from social networks, and also the limitation of the research. The organization of this paper is demonstrated at the end of this chapter.

1.1 Background

From 2000 onward, billions of internet users have blended social network sites (SNSs) such as Facebook, Instagram, LinkedIn into both of their work and life according to boyd *et al*. [1]. Although, almost all of the SNSs contain personal profiles and list of friends, technical features and user bases of each site may diverge greatly. As carriers of opinions and news, social networking sites like *Facebook*, and micro blogs like *Twitter* are 2 of the major categories of SNSs (Figure 1.1).

Lerman [2] claimed that SNSs are playing an important role in information dissemination, search, and expertise discovery. Explosion of information and data via SNSs facilitates a revolution in remote learning, entertainment, and opinion sharing. By clicking and typing, exchanging thoughts with people becomes more convenient than ever, wherever they are and whenever it is. With ubiquitous social media, people tend to look for advice or read comments from others as their own decision-making support. A survey from *GlobalWebIndex* exhibits that from year 2012 - 2018, the average time people spent on social websites increased from 1 hour 30 minutes to 2 hours 22 minutes(Figure1.2). Segmented by age, the youngest user group(16-24) ranks on the first place with an AVG usage time for more than 3 hours, which suggests the time for people spent on SNSs is likely to keep increasing in the following years.

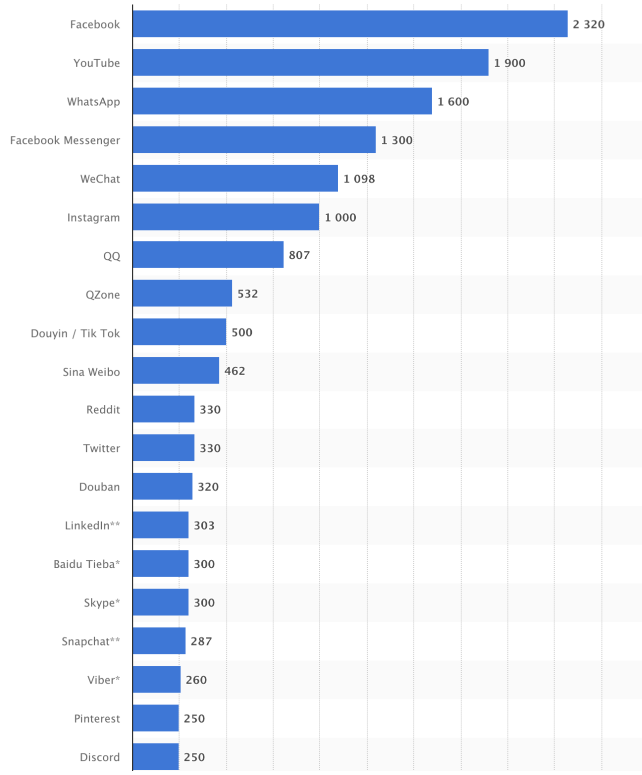

Figure 1.1. Most popular social networks worldwide as of April 2019, ranked by number of active users (in millions)

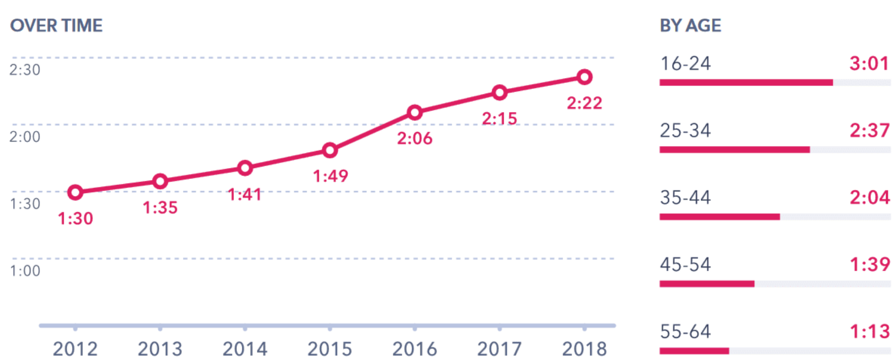

Figure 1.2. Average time spent engaging with/connected to social network/service during a typical day

In the meanwhile, doubts toward the trustworthiness of SNS news/posts arose among the public. Tapia [3] argued that a large portion of messages from social media have been deemed as untrustworthy, especially those contributed by average citizens or non-professionals. A survey carried out by *Sharethrough* using *Qualtrics* completed in September 2017. By asking 1,052 U.S. residents, age from 18 to 50, the research yielded that compared with traditional publishers like *Time Inc.* and *WSJ*, the credibility of social media falls far behind(Figure1.3). Locating reliable information has become a unparalleled curse on the other side of the benefits of SNSs stated by Metzger [4].

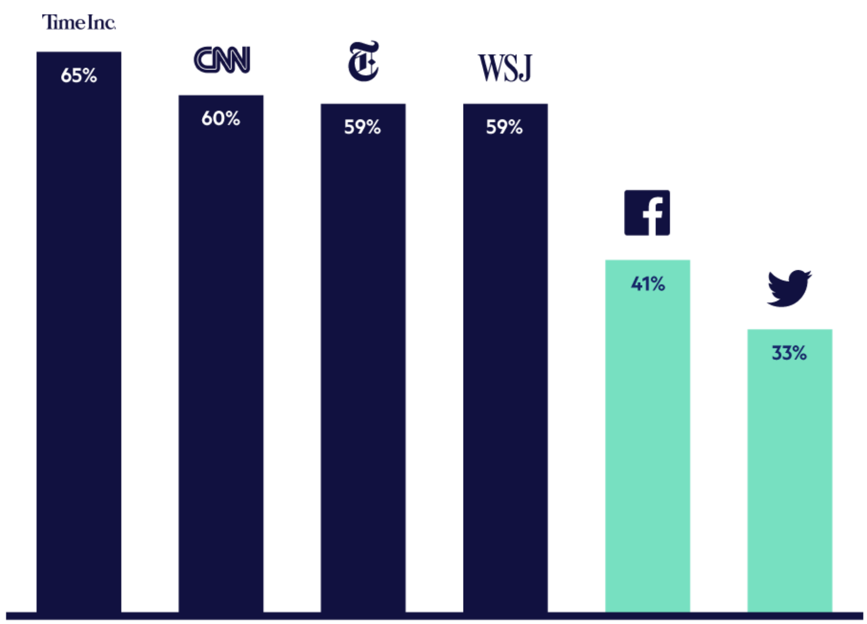

Figure 1.3. Percentage of respondents trust each publisher

1.1.1 Twitter

Founded in 2006, Twitter (http://twitter.com) is an American online news and social networking service on which users post and interact with messages known as

”tweets”, and soon gained popularity around the world. By April 2019, Twitter has over 320 million monthly actively users, and about 500 million tweets are sent everyday. As an advantage of Twitter's restrict: a post can contain at most 280 characters for almost all languages, 140 characters for Chinese, Korean, and Japanese, sentiment analysis can easily be conducted on tweets. With vast quantity of data, Twitter becomes a popular platform for data analysts to perform ad-hoc analysis.

1.1.2 Apple Inc.(AAPL)

Apple Inc., as one of the tech giants along with Google, Facebook, and Amazon, is a multinational company headquartered in California, US. Apple is famous for its R&D ability in consumer electronics and softwares and has 1.3 Billion active devices worldwide. By the year 2015, Apple's market capital reached $650 billion in total. Based on such a large quantity of user group and market value, there is sufficient attention and discussions on SNSs and can be utilised by this study.

1.1.3 Yahoo! Finance

As a part of Yahoo!'s network, Yahoo! Finance provides finance news, data, and commentary including stock quotes. In this research, the history stock data which serves as ground truth was retrieved from Yahoo! Finance.

1.2 Research Objectives

The research objectives are as follow:

- To design a method to identify trustworthy SNS information sources for general SNS users.
- To explore how the parameters will affect the performance used in this new method.

- Establish an approach to eliminate/lower the effect from noisy data.
- To examine if this new approach improves performance compared with using existing metrics.

1.3 Assumptions

Several important assumptions are made in order to simplify the study:

- Only active SNS users who post their opinions frequently may be considered as reliable information providers.
- When a SNS user pays close attention to some topic, the user will promptly post his/her opinion on SNS according to the events happening in the real word.
- The sentiment of a SNS post represents the user's judgement and expectation towards the related topic.

1.4 Limitations

The following limitations of this study should be noted.

- Among various SNSs, this research only makes use of the dataset from Twitter.
- Limited by the size of sample dataset, this study does not provide a complete picture of all SNSs users who give opinions about the discussed topic.
- This study is limited by the age and time span of dataset, from 01/01/2015 to 08/31/2015.

1.5 Delimitations

- In this research, we only study SNS users whose number of post is above some threshold. The setting of the threshold will also be explored.

- The application of the study can be generalized to any topic if there is a metric related to that topic can serve as the ground truth.

- The trust weight generated by our method for each user is only applicable to that specific related topic, and cannot be used as a reference in other topics.

1.6 Organization

This thesis contains 6 chapters, organized as follows:

Chapter 2 is a review of related works. Several subjects are studied including the importance of SNSs in the information diffusion, credibility & trust of information in social media, and existing state-of-art SNS trust management framework.

Chapter 3 introduces the framework and methodology used in this study. It consists of two parts: first, the detailed design of our iterative algorithm; second, a visualization aided denoise method for noisy data.

Chapter 4 is about data exploration and preprocessing. It starts with a description of sample data used in this study, then gives the procedures we go through to melt and reshape the dataset.

Chapter 5 provides the experiment results, along with evaluations. Also, we tried to give interpretations of all these results via intuitional observation of original data.

Chapter 6 not only presents conclusions and discussions of this research, but also proposes some potential works can be done in future.

2 LITERATURE REVIEW

This chapter is a review of related works about social network information mining & analysis. It starts with a summary of current status of SNS information mining, including why mining SNS is considered valuable and necessary, then follows researchers' comments and findings of credibility of SNS data. The last part introduces a couple of trust frameworks from recent research.

2.1 Usage of Social Media Information Mining

Mining information from social networks has become a novel field of study in modern data analysis. Liu *et al.*'s investigation on brand-related user-generated content (UGC) on SNSs [5]. By applying Natural Language Processing(NLP) algorithms: Latent Dirichlet Allocation (LDA) and sentiment analysis on posts from Twitter, they developed a framework which can provide insights for brand managers in various business operations. With geotagged tweets, Lai *et al.* generated mobile users' local interests by merging Geographical Information System(GIS) with LDA algorithm, aiming at improving the accuracy of outdoor targeted advertising [6].

Social media mining is not only used in business, but also benefits studies of politics and economy. A paradigmatic example of the great success of social network data analysis is the 2012 U.S. presidential election. The analysts in Obama's campaign draw action patterns for potential swing state voters and concluded topics and messages may persuade voters to support Democratic Party. Take a step further, insights of other subjects can be found from SNS election data. A computational public opinion mining approach is developed by Karami *et al.* to explore economic issues from social media during an election [7]. This economics-based opinion mining approach combines two data mining methods: Linguistic Inquiry and Word Count

(LIWC) and LDA. Karami states that this method can offer a better understanding of public opinions on some specific topics.

Besides, social network can even help to predict stock price. With properties of high volatility, dynamics and turbulence, stock price prediction is considered one of the most difficult tasks. It does not simply depend on financial indicators, natural and politic issues may also influence stock market implicitly or explicitly. With the thriving of mining and modeling techniques of social media data, researchers are exploring an approach from a new perspective. Dual sentiment analysis is used by Naren in stock prediction [8]. By analysing financial news articles with timestamp retrieved from *Yahoo! Finance*, Gidófalvi proposed his method of predicting stock price movement with a naive Bayesian text classifier [9]. Bollen *et al.*'s study proved that Twitter mood – tracked by OpinionFinder and Google-Profile of Mood States(GPOMS) can significantly improve the prediction of Dow Jones Industrial Average(DJIA) [10].

There are a couple of reasons for the vast usage of social networks information mining. In Mao's study [11], survey, news media, social feeds and search engine data are investigated as source data of stock price prediction. Compared with others, survey is considered most expensive to conduct. Since survey quality may vary due to respondents' biases and truthfulness, and the result is usually lagging, its sentiment indicator cannot statistically significantly predict market performance. While, Google Insight Search volume, Twitter Investor Sentiment and sentiment from traditional/on-line news media are found to be statistically significant predictors. For social feeds data, Mao randomly sampled 15% - 30% of total tweets from July 2010 to September 2011, and defined two indicators: 1) Twitter Investor Sentiment (TIS); 2)Tweet volumes of financial search terms (TV-FST). By analysing the correlations between these two indicators and the DJIA, the indicators reveals a statistically significant Pearson correlation coefficient of 0.62 over weekly values. Although, Public mood indicators extracted from social networks has been used to predict stock market fluctuations, according to Mao, different types of web data are used to predict different financial indicators, and it is not clear which sentiment indicator has the best predictive power over which specific indicator.

2.2 Credibility & Trust of information in SNSs

The development of internet makes social websites ubiquitous. Pew Research (2014) shows that at least 30% of adults in the U.S. are consuming news on Facebook, and 78% of them claim themselves exposed to news while they are using Facebook for other purposes. On the other side, Gronke *et al.* believes the trust in mainstream media has declined in the past decade [12], more and more audiences are drifting away from traditional news publishers to social media. On SNSs, users get news feeds either from subscription of news organization, or posts created by other average users (friends). In this process, professional journalists and friends work as censors, and may evaluate the news content. So far, few studies have been done to assess the credibility of news delivers on social media [13].

Tapia *et al.* conducted a research to find out whether micro-blog information from Twitter is trustworthy to fit the needs of disaster response [3]. For professionals of disasters and emergency response, social feeds are considered as rich sources of timely data that may offer valuable information affected individuals and respondents. On the other hand, affected users may offer local specifics in disasters to keep outsiders informed via social media. Disasters and emergency responders have developed standard centralized operating mechanisms to response crisis in the past. Lack of vetting standards, challenges arise when responders try to integrate information from unprofessional observers with current operating mechanisms. The major concern is about trust due to the veracity, accuracy, and legitimacy of data.

To determine trustworthy sources from social media, Metzger [4] studied traditional media sources and pointed out that credible information is often provided by those received professional training and education or have jobs requiring specific experience. The complexity in ways to find out credible data has been increased due to the emergence of tremendous social data. According to Metzger, the greatest challenge is to find the desired information from among possible sources, in other words, locate the most trustworthy information providers.

2.3 Existing Trust Management Framework of Social Network Data

As the trust issue being proposed, scholars and researchers have developed various methodologies and frameworks to evaluate the credibility for social network data.

After the investigation of Facebook pages, Li *et al.* summarized the factors that may effect information credibility [14]. Li stated that credibility research started with an interest in how people become persuaded. The Elaboration Likelihood Model (ELM) developed by Petty and Cacioppo [15] demonstrates 2 routes to affect receiver's attitude toward information sources: Central route requires user's evaluation on content and argument strength of information; Peripheral route depends on information-irrelevant factors. The significance of effect on information consumers are adjusted by consumers motivation and ability: central route is usually used by consumers with better ability to evaluate information credibility, otherwise, peripheral route is used. Based on ELM, Li developed a credibility prediction model for social media platforms. The model defines 5 major factors under 2 credibility dimensions, the structure and components are illustrated as follow:

1. Medium Credibility: Medium Dependency, Interactivity, Medium Transparency

2. Message Credibility: Argument Strength, Information Quality

The above-mentioned factors will be moderated by personal expertise to obtain the final information credibility.

Using cognitive heuristics is a common approach to study social information credibility. Metzger [4] concluded 6 types of heuristics that can be used in credibility judgments, listed as follow:

1. Reputation: Also known as name recognition, is also a subset of "authority" heuristic. It is the most basic heuristic principle used to lower people's effort to process online information.

2. Endorsement: Shows that people tend to trust information and sources credited by others. It also reduce people's workload in filtering out unreliable information.

3. Consistency: Verify information consistency across various sources. This method requires more effort than others, and is also regarded as a variant of the endorsement and reputation heuristics.

4. Self-confirmation: It means that people tend to believe in information which confirms their preexisting knowledge and consider the information incredible if it refutes their existing beliefs.

5. Expectancy Violation: When a source fails to meet a user's expectation, it will be judged unreliable. For example, presence of typos or grammatical errors.

6. Persuasive Intent: This refers to the tendency that people consider information not credible when they feel the information is biased. It is often found in commercial information, especially for unexpected advertising.

2.4 Visual Analytics

Visualization is the science of analytical reasoning facilitated by interactive visual interfaces [16]. Begin with maps and charts in 17th century, graphical presentation for quantitative data has been used for centuries [17]. A visual presentation of data enable users intuitively explore a large quantity of data. With visualization, user can easily get the big picture, grasp difficult concepts, and identify new patterns.

Much work has been done on visualise text data/sentiment. The simplest and most common form of text visualization is word cloud, which is a concise and fun way to summarize contents from text and website. Cui *etal* designed a state-of-art method [18] which can dynamically construct cloud words and ensures semantic coherence and spatial stability at the same time. Dou [19] argues that combining topic modeling algorithms with matrix visualization can be used as a topic-driven visualization method to reveal correspondences between topics & terms, topics & documents. For time series text data, river/stack graph is frequently used to plot evolving topics. Besides, the theme river metaphor can also portray the trend of topics over time. Hao and Ben-Avi [20] developed an multi-dimensional sentiment visualization application for Reddit. The application first analyse the text by IBM Watson Alchemy API, and get sentiment scores in 5 different emotions dimensions including joy, sad, anger, fear, and disgust. The application then demonstrate the 5 scores respectively in a bar chart.

3 FRAMEWORK AND METHODOLOGY

The Iterative Method for Reliable User filtration is used to find out a trust-weighted information source combination for common social website users. The framework consists of two parts: (1)the main body is trust evaluation, which assigns coefficients for each user based on their trust performance; (2)data visualization is used to facilitate reducing noise from real world.

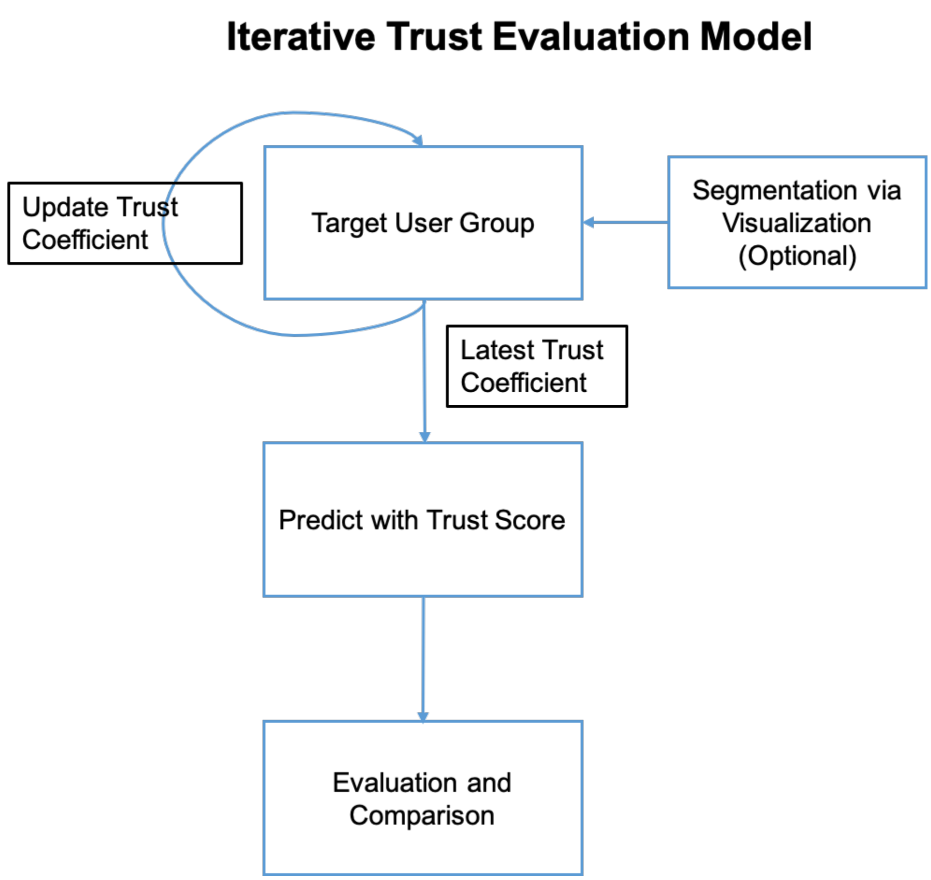

Figure 3.1. The Iterative Trust Evaluation Model with Visualization Aided Segmentation

3.1 Pearson Correlation Coefficient

To find out the trust score for each user, we need to define a method for trust evaluation. Intuitively, users whose opinions can positively relate to ground truth are regarded as trustworthy sources. In existing works, a couple of methods have been used to study the relation between social sentiments and stock market indicators [10, 21], and Pearson Correlation Coefficient (PCC) is one of them. In our model, we use the PCC as the measurement to describe the extent to which two variables are linearly related. For each user within some time period, we will calculate the correlation $Corr_{ST}$ between the user's sentiment with quantized ground truth. Let T denotes ground truth value, and S denotes user sentiment, the PCC within date i to j can be calculated with Equation 3.1:

$$Corr_{ST} = \frac{\sum_{d=i}^{j}(S_d - \overline{S})(T_d - \overline{T})}{\sqrt{\sum_{d=i}^{j}(S_d - \overline{S})^2}\sqrt{\sum_{d=i}^{j}(T_d - \overline{T})^2}} \tag{3.1}$$

3.2 Iterative Trust Evaluation Method

Suppose a user are following N other social network users who give opinions about the target topic. From day i to day j, we can retrieve sentiments for each user each day, and the ground truth for each day. In the beginning, each source in the target group is treated equally, and their trust coefficients are assigned to 1. Example data is demonstrated in Table 3.1 and Table 3.2:

Table 3.1.
Description of Social Sentiment Dataset for Trust Evaluation

$User$	Day_i	Day_{i+1}	$Day_{...}$	Day_j	$TrustCoe$
$User_1$	$senti_{u_1 d_i}$	$senti_{u_1 d_{i+1}}$	$senti_{...}$	$senti_{u_1 d_j}$	$TrustCoe_{u_1}$
...	...	...	...	...	...
$User_n$	$senti_{u_n d_i}$	$senti_{u_n d_{i+1}}$	$senti_{...}$	$senti_{u_n d_j}$	$TrustCoe_{u_n}$

Table 3.2.
Description of Ground Truth Dataset for Trust Evaluation

$User$	Day_i	Day_{i+1}	$Day_{...}$	Day_j
$Truth$	$Truth_{d_i}$	$Truth_{d_{i+1}}$	$Truth_{...}$	$Truth_{d_j}$

In this method, we keep updating the trust coefficients for each user and renewing correlation between the overall sentiment and ground truth until the overall correlation converge – the absolute value of previous and current correlation is smaller than threshold ϵ. When updating the correlation efficient for each user, first we need to calculate the correlation coe_{user} between user sentiments and truth value, then multiply the previous coefficient coe_{old} by the current coefficient plus one (Equation 3.2). We add 1 to coe_{user} since the the PCC is in range $[-1, 1]$. When PCC equals -1 means variants are negatively related, and $+1$ stands for a perfect linear relation. Intuitively, when a user's sentiments are negatively related to the ground truth, we consider this user is not trustworthy (or we say it is not a useful information source), and vice versa.

$$coe_{new} = coe_{old} * (coe_{user} + 1) \tag{3.2}$$

In each iteration, we also store an overall sentiment for the target user group simply by adding up each users trust-weighted sentiment for each day. Finally, the overall sentiment is used to calculate the overall correlation between user group and ground truth. See the following Algorithm 1 for details.

3.2.1 Prediction with Accumulated Trust Coefficient

When the overall PCC converges, we get a set of trust coefficients for each user. To validate this model with hold out data, we multiply users' sentiment with the trust coefficients we get from iterative step for each user respectively, and calculate the overall PCC then compare with the unweighted one.

Algorithm 1 Iterative Trust Evaluation Algorithm

procedure ITERTRUST(S, T) ▷ S: sentiment, T: ground truth
  $overallSenti \leftarrow [0, ..., 0]$ ▷ length of dataset
  $preCorr \leftarrow inf$
  $corr \leftarrow 0$
  while $|preCorr - corr| > \epsilon$ **do** ▷ Converge condition
    for $user\ in\ S$ **do**
      $userCorr \leftarrow Corr(user, T)$
      $TrustCoe \leftarrow TrustCoe * (userCorr + 1)$
      $overallSenti \leftarrow overallSenti + user * TrustCoe$
    $preCorr \leftarrow corr$
    $corr \leftarrow Corr(overallSenti, T)$
  return $corr$ ▷ return the final correlation

3.3 Visualization for Denoising

When using social media data to support decision making, we need to take several issues into consideration. The most obvious and important one is data density, which is also the reason for us to filter out inactive users. Although we assumed SNS users tend to share opinions online, in the real world, unlike professional media, social media users do not necessarily share their opinions about some topic. It means sometimes even if a topic related event happens, a social website user may not post his/her comments about it. This is a critical issue especially when we study problem. Therefore, users do not give adequate information should not be selected into the target group. The second issue is: a user's analysis ability may vary with time. If the time range we investigated is long enough, it is unfair to assign a single trust coefficient to an information source for the whole time. Observing the evaluation metric, if the variation of trust score is significant enough, it is reasonable to divide the data into corresponding parts and calculate trust values independently.

Visualization can bring great simplicity when dealing with massive data analysis. User can easily get the big picture and see the details at the same time, thus to quickly identify abnormal issues. To lower the effect caused by these noise, we introduced visualization aided segmentation for sentiment dataset. For the sentiment dataset, we found very few users would write posts every day. Among all the users we collected, less than 1% users posted over 100 tweets during the first 8 months in 2015. Figure 3.2 shows an example of an user's sentiment and ground truth. We can see that the user's sentiment remains the same in the first 25 days, while in the rest of 218 days, his/her sentiment fluctuates frequently. We consider the first 25 days an abnormal period, while overall this user is very active. With data visualization, we can easily spot this abnormal, and make a segmentation at the 25th day. This user then have 2 trust scores from 2 segments respectively. This segmentation can be applied multiple times on a data instance, as long as we consider abnormal periods exist.

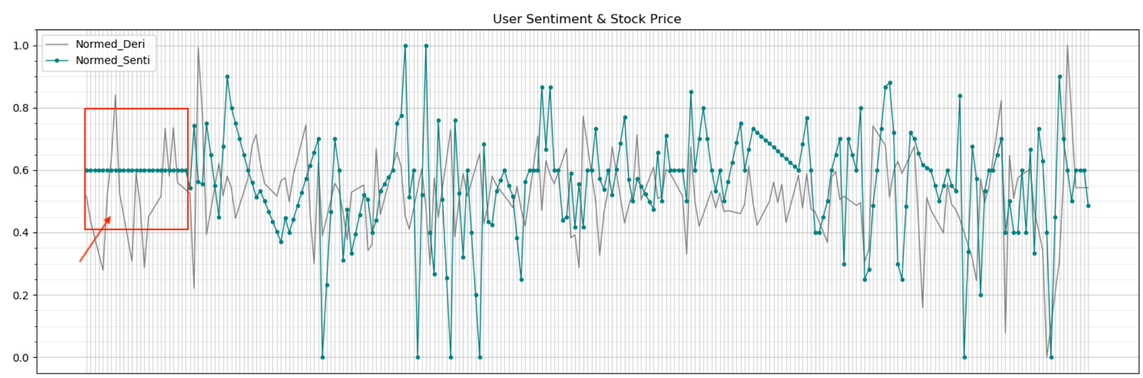

Figure 3.2. Example: Visualization for an user's sentiment and ground truth. The abnormal part is selected by a red rectangle.

Draw on the iterative trust evaluation method, with visualization, we can input a series of segmentation points. In this version, before calculating trust scores for each user, we need to check whether there is any segmentation point for this user. If there is, we need to calculate trust score for each segment, and store different trust scores for each of the periods (unlike the first version, each user only have one score). The upgraded algorithm is demonstrated as follow (Algorithm 2):

Algorithm 2 Iterative Trust Evaluation with Visualization Aided Denoising

```
procedure ITERTRUST(S, T, Seg)    ▷ S: sentiment, T: ground truth, Seg: segmentation, list of list
    overallSenti ← [0, ..., 0]    ▷ length of dataset
    preCorr ← inf
    corr ← 0
    while |preCorr − corr| > ε do    ▷ Converge condition
        for user in S do
            if Seg[user] ≠ None then    ▷ if there are segments for a user
                for period in Seg[user] do
                    userCorr[period] ← Corr(user[peroid], T)
                calculate corr for each period
            else
                userCorr ← Corr(user, T)
        TrustCoe ← TrustCoe ∗ (userCorr + 1)
        overallSenti ← overallSenti + user ∗ TrustCoe
    preCorr ← corr
    corr ← Corr(overallSenti, T)
return corr    ▷ return the final correlation
```

4 DATA EXPLORATION AND ANALYSIS

4.1 Data Collecting

The input data of the experiment consist of 2 parts. Both of these datasets are within time interval from 01/01/2015 to 08/31/2015, 243 days in total:

1. The twitter dataset is a ".txt" file contains 113615 tweets about AAPL (Apple's ticker symbol) from Twitter. The tweets are are posted by Tweeter users following any of the 3 official financial accounts, and retrieved through Twitter's open API and Twitter library use following procedures:
   (a) Select followers of financial related accounts: StockTwits, FinancialTimes, and MarketWatch as targeted user groups. Followers of these accounts are considered highly related to the selected topic.
   (b) Retrieve targeted user IDs and the tweets posted by them. Each of the tweets consists user ID, number of his/her followers, tweet text, and date & time of the post. Table 4.1 gives a detailed description of the data structure.

Table 4.1.
Description of Tweet Dataset

Attribute	Description
User ID	Unique user ID for Twitter platform
NO. of Followers	The number of followers for the user
Tweet	Text content of the post
Date & Time	WeekDay MM dd HH:mm:ss TimeZone YYYY

2. The financial dataset, used as the ground truth in our analysis, is a ".csv" file that records historical daily statistics of AAPL retrieved from Yahoo! Finance. The file contains 173 data instances instead of 243, due to the market does not open for trading during the weekend. The original data instances contain 8 attributes, but a couple of them are not relative to our research (like Volume, Price to Earnings Ratio, and Price to Sales Ratio). So we only keep 3 most important attributes, listed in the following Table 4.2:

Table 4.2.
Description of AAPL Stock Price Dataset

Attribute	Description
Date	Trading date
Open	Open price of AAPL of that day
Close	Close price of AAPL of that day

4.2 Data Distribution

In the Tweet dataset, there are 17171 different user IDs in total. Taking a look at the distribution of user number grouped by number of tweets posted by each of them (Figure 4.1), we can find that the numbers of tweets from each user vary within a wide range, from 1 to 2602. 89.73% of target users posted less than 10 tweets during that period, and less than 1% users have more than 100 posts. This distribution shows that, although users in this dataset are following a specific topic on Twitter, they do not provide much information for other users. Thus, we will ignore users with no more than 10 tweets in our study.

For the remaining 1594 users, we also inspect the number of followers for each of them, see distribution in Figure 4.2. Firstly, sort users in order of the number of their followers from low to high, then plot their follower number respectively. We can observe that most users have 100 to 10000 followers, and only a small portion of them have large quantities of fans.

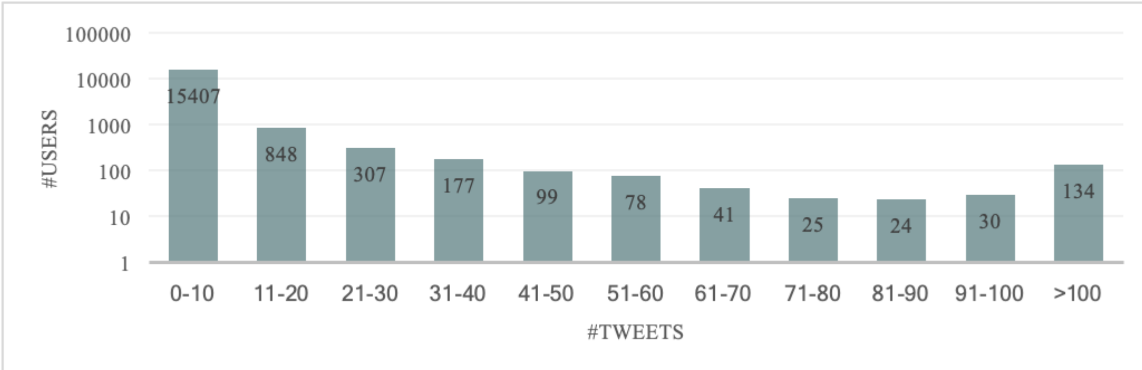

Figure 4.1. Distribution of User Number Grouped By Number of Tweets Posted by Each User (Y axis is logarithmic scaled with base 10)

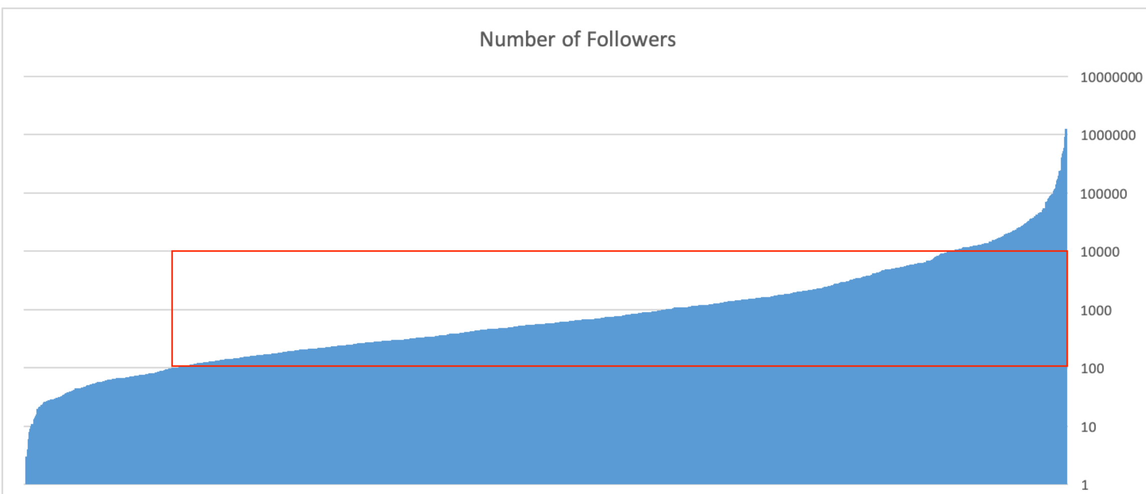

Figure 4.2. Distribution of Follower Number for Users (Y axis is logarithmic scaled with base 10)

To simplify the study, we focus on the close prices of AAPL as plotted in Figure 4.3, which is sufficient for observing the trend of stock price. In the first 8 months in 2015, Apple's stock price moved up from January to mid of February, then peaked

4 times in fluctuation until the mid of July, finally followed by a drop to its lowest price during this period.

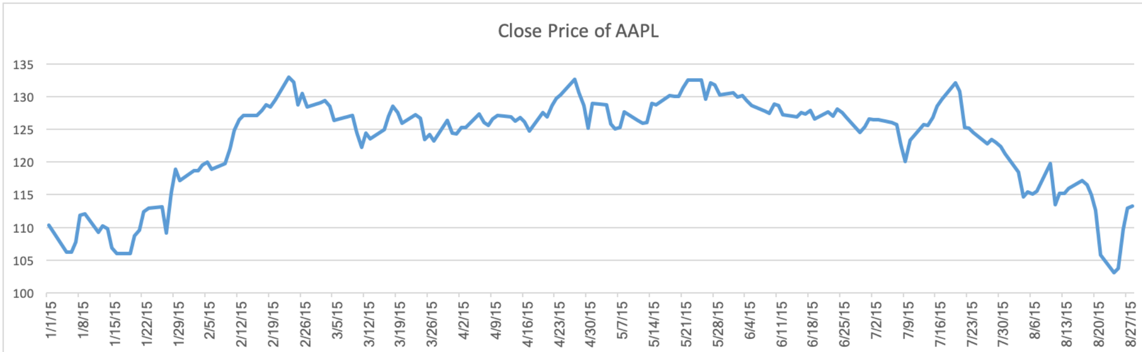

Figure 4.3. Close Price of AAPL 01/01/2015 - 08/31/2015

4.3 Data Preprocessing

The retrieved data contains missing values and is in undesired format. Some data preprocessing work are done before putting the dataset into use.

4.3.1 Sentiment for tweets

Suggested by existing works [5,8,9,11], social sentiment can be used as an effective tool in business and financial analysis. Various techniques are available including .... In this research, the sentiment of each tweet have been produced by SentiStrength – an English social web texts sentiment analysis (opinion mining) program. Considering SentiStrength is not specially designed for financial topic analysis, in Ruan's work, Loughran and McDonald's financial dictionary were additionally added into lexicon as an amendment to improve the analysis performance [21]. Among different kinds of outputs of SentiStrength, single scale: [-4, +4] results is used to represent different strength of emotion. Let S denote the sentiment for a tweet, negative and positive numbers represent negative and positive altitudes respectively, the larger the absolute S is, the more intense the emotion is.

In our experiment, we want to study the relation between the change of stock price and the change of a specific user's sentiment towards the stock by day. However, it is very likely that in a single day, an user may post several tweets about the same topic. To aggregate multiple posts, we treat each tweet equally, and use the average sentiment to represent a user's daily sentiment (denoted as S_{d_n}):

$$S_{d_n} = \frac{\sum_{k=1}^{c} S_{d_n t_k}}{c}, S_{d_n t_k} \in \{S \mid S \text{ is sentiment of a tweet } t_k \text{ on day } d_n\} \tag{4.1}$$

4.3.2 Interpolation of Datasets

From Figure 4.1, we can see only a small portion of users posted more than 100 tweets in 8 months, which means for most users, there are a few days no sentiments are given. To deal with this problem, we use linear interpolation to get values at positions in between 2 sentiment data points. The null values are represented by a straight line segment joining 2 closet non-null values. To get sentiment value for day d_k, we can use the following Equation 4.2:

$$S_{d_n} = S_{d_i} + (d_n - d_i)\frac{S_{d_j} - S_{d_i}}{d_j - d_i}, (i < j) \tag{4.2}$$

For AAPL dataset, as mentioned in data collecting section, there are only 173 trading data. To match the instance number of Twitter dataset, we need to set a close price for every single day. In our experiment, we simply use the last close price to fill up the following non-trading days, since in non-trading days the stock price remains the same.

4.3.3 Derivation of Stock Price

In the experiment, actually we want to use the change of stock price as the ground truth, since a user's sentiment is a reflex of his/her opinion about how does an issue will go. When people talk positively about a company, it means people have confidence in its products, profit or growth, and a bullish stock price will be an objective fact, vice versa. Here, we can use the derivation $Deri_{day}$ (Equation 4.3) to represent the change of AAPL's stock price $Price_{day}$. Set the first day's value to 0, since there is not previous day for it.

$$Deri_{day} = \begin{cases} 0, & \text{if } day = 1 \\ \dfrac{Price_{day} - Price_{day-1}}{Price_{day}}, & \text{otherwise} \end{cases} \tag{4.3}$$

4.3.4 Normalization of Datasets

Since the sentiment and financial dataset have different ranges, the sentiment and derivation of price need to be normalized. Common normalization methods include Range Normalization and Standard Score Normalization, we use the Range Normalization in our study. Let $\hat{r}$ denotes the range of data, which is calculated as $data_{max} - data_{min}$, each value is scaled by Equation 4.4:

$$data_{norm} = \frac{data_{orig} - data_{min}}{\hat{r}} = \frac{data_{orig} - data_{min}}{data_{max} - data_{min}} \tag{4.4}$$

Normalized derivation $Norm_deri$ is added to AAPL financial dataset, and normalized sentiment $Norm_senti$ is added to Twitter sentiment dataset. Figure 4.4 and Figure 4.5 visualize the differences before and after normalization.

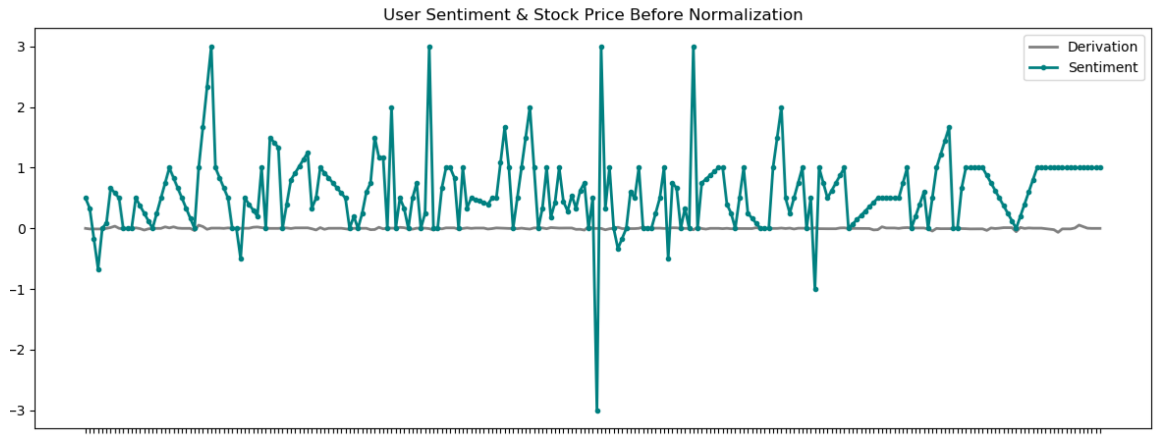

Figure 4.4. Example: User Sentiment & Stock Price Before Normalization

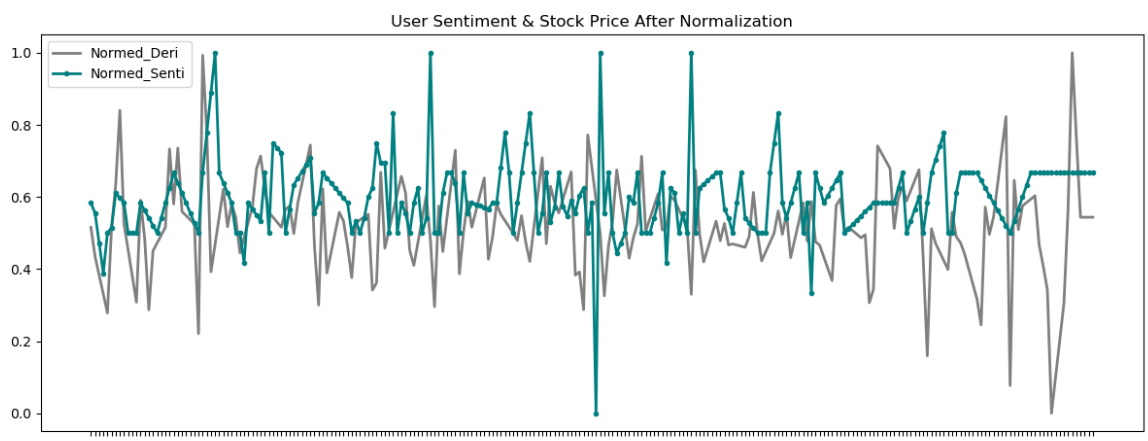

Figure 4.5. Example: User Sentiment & Stock Price After Normalization

5 EXPERIMENTS, RESULTS, AND EVALUATION

5.1 Experiments

To validate the iterative trust evaluation method, we use the Twitter Sentiment and Stock Price datasets as inputs. In the meanwhile, how different parameters will effect the performance is studied in the experiments.

5.1.1 Active User Filtration

We set a couple of thresholds for active accounts and tried to find out an appropriate one, The relation between number of active accounts and thresholds is displayed in Figure 5.1. We can see that if the threshold is lower than 100, there are more than 26 active users, and if the threshold is higher than 150, there will be less 10 active users. Therefore, we choose 100, 110, 120, 130, 140, and 150 as reasonable thresholds, and will examined respectively.

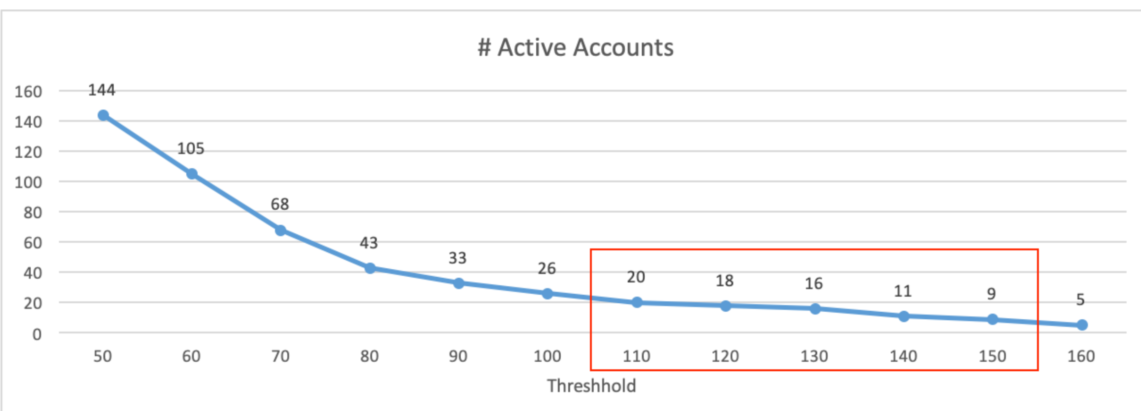

Figure 5.1. The relation between number of active accounts and thresholds

5.1.2 Data Spliting for Test

Considering the total length of our data is 243, and the trust score for a user is time sensitive, we set training size to 70 days and 30 days for test. We randomly set 10 start dates, and then for each start date, train the trust coefficients with the data of following 70 days iteratively (Figure 5.2). To verify the visualization aided denoising, we perform 2 rounds of tests: with and without segmentation.

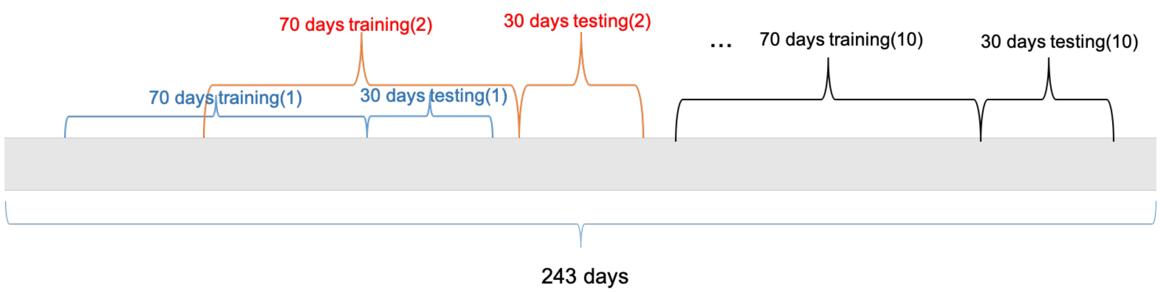

Figure 5.2. The configuration of experiments data for training and testing

Take *threshold* : 150 as an example, there are 9 active accounts. After applying the iterative method, we can observe that the overall correlation between user sentiments and the change of stock price keeps increasing until it converge. We run 10 tests in total for this 9-user group, and recorded the begin day, number of iterations until converge, original correlation, and the final trained correlation. The converge threshold ϵ is set to 0.01, and the detailed statistics are displayed in Table 5.1. In the 10 tests, it takes average 8.3 iterations to converge, the average correlation raise from, originally, 0.124 to 0.350 with the trust scores as weight.

We also did the same experiments for other thresholds: 100, 110, 120, 130, and 140. The average iteration number, original, and trained correlation are recorded in Table 5.2. From the data, we can see that when information sources are treated as equal, the average original correlation can be as low as 0.131, after we iteratively evaluate the trust score for each user, the final average correlation can reach 0.291. If we plot the threshold and the final correlation after training (Figure 5.3), we can

Table 5.1.
Training Data & Result with $Threshold : 150$

No.Test	BeginDay	No.Iter	OriginalCorr	TrainedCorr
01	85	7	0.049	0.371
02	56	8	0.180	0.389
03	21	11	0.150	0.348
04	65	8	0.082	0.317
05	3	10	0.162	0.261
06	30	9	0.171	0.329
07	123	7	0.210	0.378
08	80	7	0.081	0.409
09	67	8	0.095	0.314
10	77	8	0.057	0.377
AVG	–	8.3	0.124	0.350

find that except as we raise the active account threshold, the corresponding trained correlation tend to increase as well.

5.2 Model Evaluation

After the training work has been done, all the users have been assigned with trust coefficients. In the next step, we are going to test the correlation before and after applying the trained user trust scores as weight. For model without segmentation, the trust CoEs are unique within each user, with updated sentiment dataset S', test dataset S_{test}, ground truth T_{test}, the testing procedures are described as following Algorithm 3:

Take the $threshold : 150$ as an example again, the original, trained, predicted correlation, and improvement are recorded in Table 5.3, after weighted by trust score, the

Table 5.2.
Average Training Results with Different Thresholds

Threshold	AVG Iterations	AVG Original Corr	AVG Final Corr
100	11.2	0.154	0.269
110	10.5	0.119	0.226
120	10.2	0.121	0.278
130	9.8	0.135	0.306
140	10.4	0.131	0.317
150	8.3	0.124	0.350
AVG	10.1	0.131	0.291

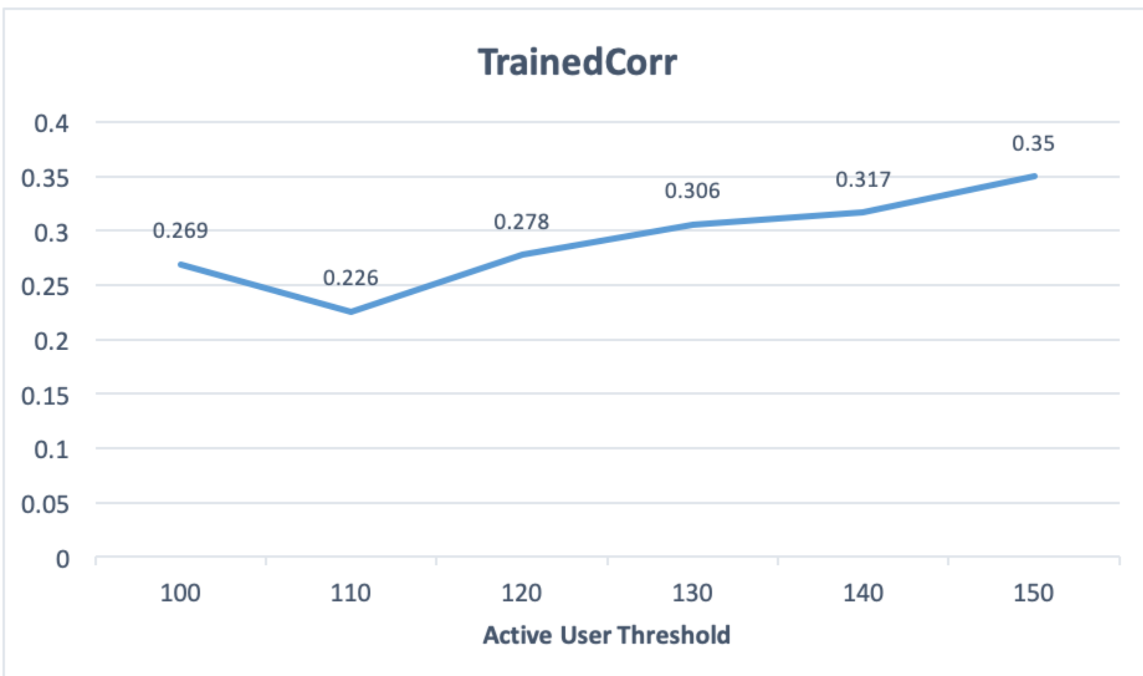

Figure 5.3. The relation between thresholds and final correlations after training

Algorithm 3 Iterative Trust Evaluation Model Testing Algorithm

procedure TEST(S', S_{test}, T_{test})
  $overallSenti \leftarrow [0, ..., 0]$ ▷ length of dataset
  for $user$ in S' **do**
    $CoE \leftarrow S'[user][TrustCoe]$
    $overallSenti \leftarrow overallSenti + S_{test}[user][Senti] * CoE$
  $predCorr \leftarrow Corr(overallSenti, T_{test})$
  return $predCorr$ ▷ return the predicted correlation

average predicted correlation can reach 0.193, the variance σ^2 of predicted correlation is 0.017. One thing we need to notice is that the model may fail to improve the correlation, see Figure 5.4. It also demonstrates that although the predicted correlation outperforms the original one, it falls far behind the trained correlation.

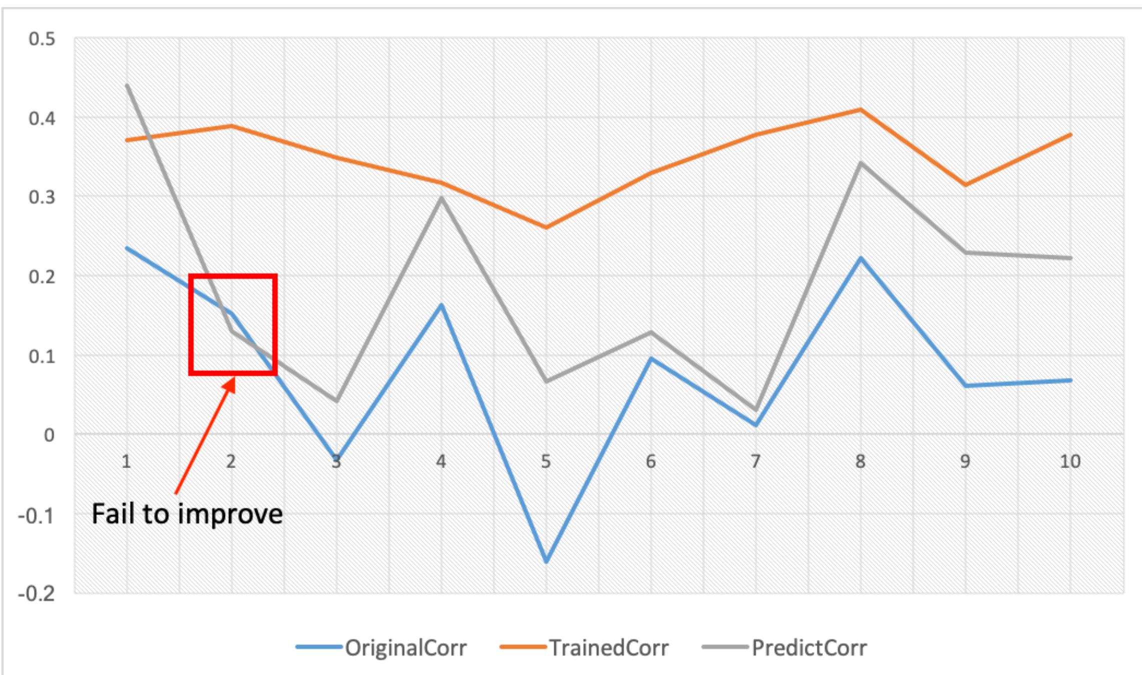

Figure 5.4. The Comparison between Original, Trained, and Predicted Correlations

Table 5.3.
Prediction Result with Threshold: 150

No. Test	OriginalCorr	TrainedCorr	PredictedCorr	Improvement%
01	0.235	0.371	0.439	86.81
02	0.152	0.389	0.130	-14.97
03	-0.032	0.348	0.042	231.25
04	0.163	0.317	0.297	82.20
05	-0.161	0.261	0.067	141.61
06	0.096	0.329	0.128	33.33
07	0.012	0.378	0.031	158.33
08	0.222	0.409	0.342	54.05
09	0.061	0.314	0.229	275.41
10	0.068	0.377	0.222	226.47
AVG	0.082	0.350	0.193	144.30

5.3 Iterative Method with Visualization Aided Denoising

As demonstrated in Figure 3.2, an user's performance may be unstable due to all kinds of reasons. To lower the effect caused by noise, we can have a more objective evaluation by dividing user's sentiments into several segments. Again, take *threshold* : 150 as an example, we still run 10 tests with random start dates. The 9 active users' sentiments and AAPL price from day 55 to 155 in one of the tests are visualized in Figure 5.5 for demonstration. For each user, we can make arbitrary number of segments wherever we think there is anomaly. The segment points are recorded in the system and used when we calculate PCC for each segment in each user.

Table 5.4 recorded the training and prediction results for 10 tests. The average trained correlation is improved by 166.0% from 0.147 to 0.391. Compared with unseg-

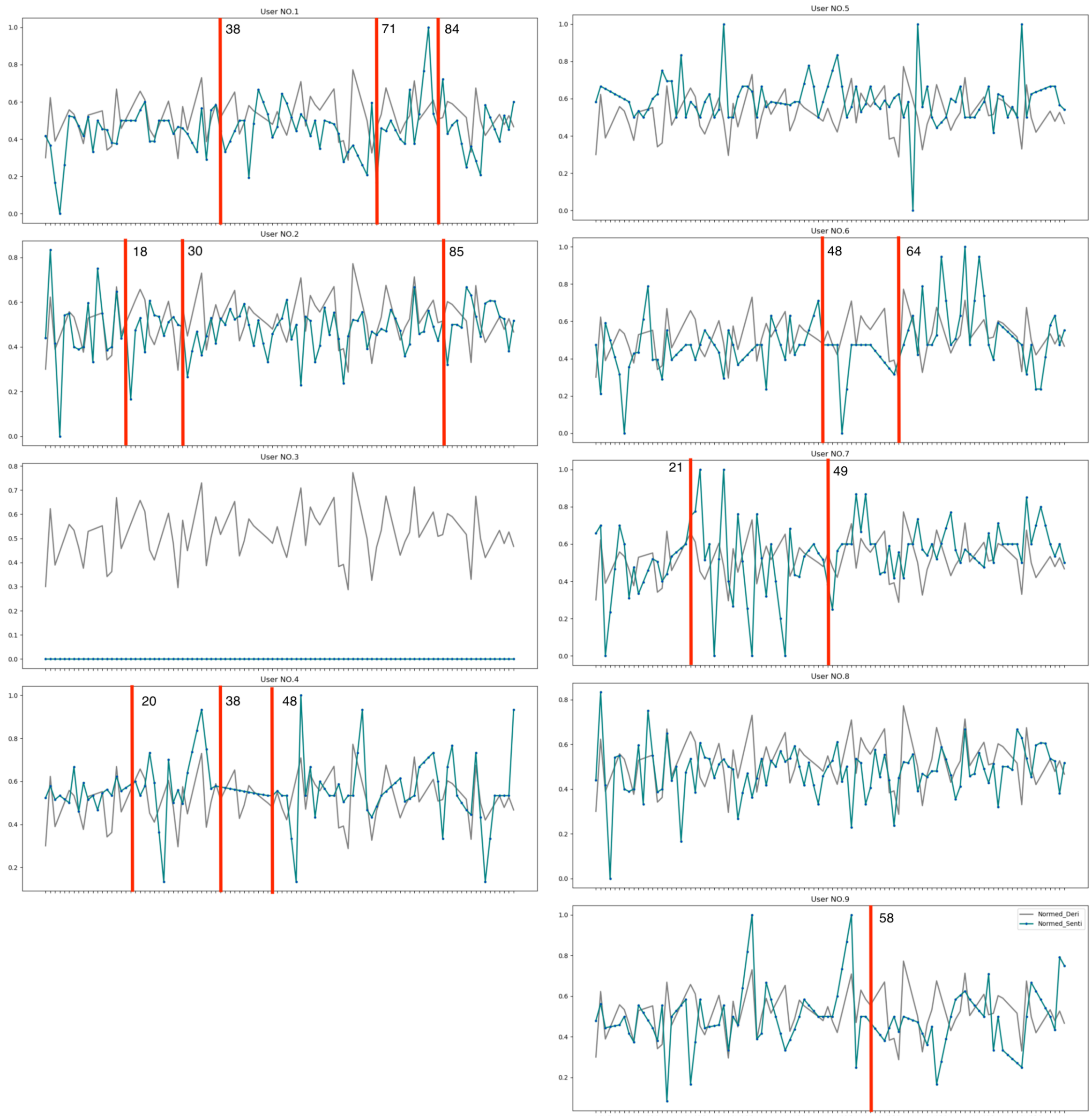

Figure 5.5. Segmented User Sentiment & AAPL Stock Price Derivation

mented sentiments, the average trained correlation raised by 0.1, which means with segmentation, the overall weighted sentiment can better correlate with the AAPL stock price derivation. After training, we predicted the correlation in the following 30 days with each set of trust score for the 10 test using Algorithm 4. With segmentation, we care more about the last trust CoE for each user, which is used in prediction. The prediction results are listed in the last column of Table 5.4, compared with unsegmented method, the average predicted correlation is improved by 0.106 to 0.310, and the variance σ^2 is reduced from 0.017 to 0.004. The comparison between original, trained, and predicted correlation is plotted in Figure 5.6.

Algorithm 4 Visualization Aided Iterative Trust Evaluation Model Test

```
procedure ITERTRUST(S', S_test, T_test, Seg)
    overallSenti ← [0, ..., 0]                          ▷ length of dataset
    for user in S' do
        CoE ← S'[user][TrustCoe][−1]
        overallSenti ← overallSenti + S_test[user] ∗ CoE
    predCorr ← Corr(overallSenti, T_test)
    return predCorr                          ▷ return the predicted correlation
```

5.4 Comparison with Other Trust Index

The last experiment we did is a comparison of the effectiveness between our proposed model with using social accounts' follower numbers as trust score. Intuitively, the number of an social account's followers can reflect the trustworthiness of that account. The more reliable information that account provides, the more SNS users would follow it. Take the *threshold* : 150 example again, among these 9 accounts, account 3 has the largest number of follower: 5596, while account 1 has fewest followers: 97. Therefore, if the number of followers are used as trust indicator directly, we can infer that user 3 will overpower all other users and dominate the overall sentiment. Thus, it is more reasonable to take the logarithm of follower numbers before put into use.(See Figure 5.7 for more detailed distribution)

To find out whether follower number is a good trust index, we calculated the correlation with follower numbers (logarithm) as trust CoEs. The output shows

Table 5.4.
Segmented Training Data & Result with $Threshold : 150$

No.Test	BeginDay	No.Iter	OriginalCorr	TrainedCorr	PredictedCorr
01	65	9	0.216	0.366	0.324
02	26	8	0.225	0.438	0.352
03	89	11	0.136	0.464	0.328
04	12	9	0.093	0.381	0.311
05	32	10	0.162	0.424	0.216
06	3	9	0.121	0.461	0.324
07	112	8	0.165	0.446	0.376
08	81	7	0.081	0.387	0.190
09	53	8	0.065	0.349	0.283
10	42	8	0.153	0.378	0.291
AVG	–	8.7	0.142	0.410	0.310

(Table 5.5), with follower numbers (base e and base 2 logarithm) as trust CoEs, the overall correlation is higher than original correlation. But, from Figure 5.8, we can tell that with follower number, either with base e or base 2 logarithm, the predicted CoEs are highly correlated with original correlation, while with trust score form our model, the prediction outperforms the other two especially when the original correlation is low. Also, the variance of our model prediction is 0.009, far lower than the other two: 0.051 and 0.052 respectively.

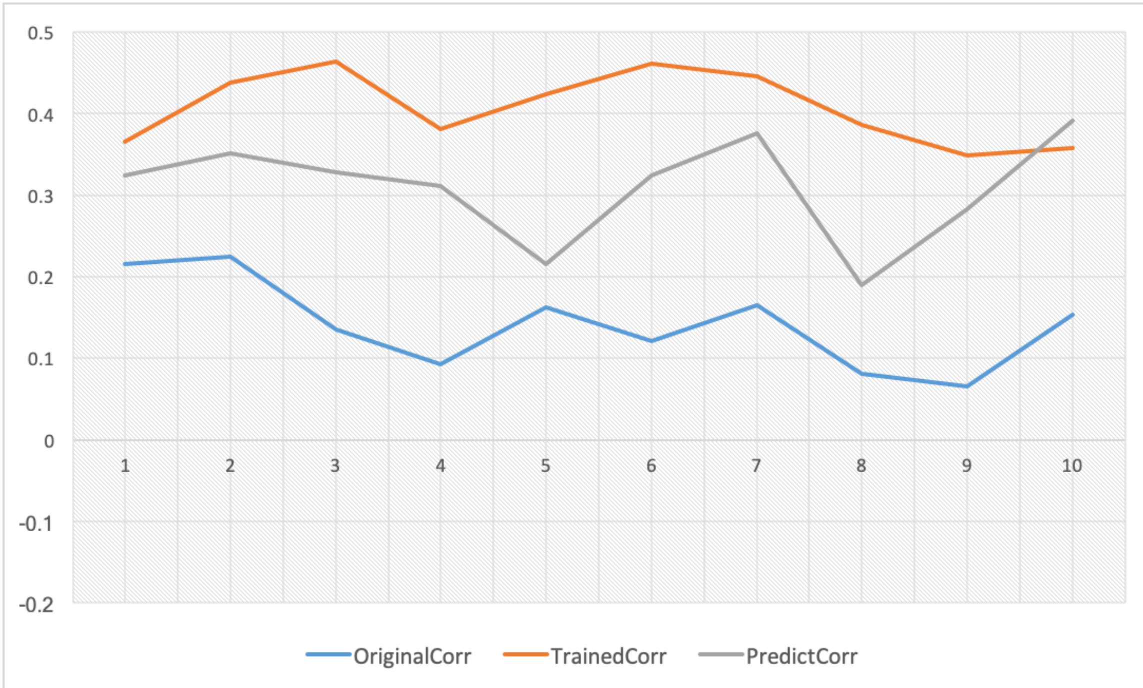

Figure 5.6. The Comparison between Original, Trained, and Predicted Correlations with Segmentation

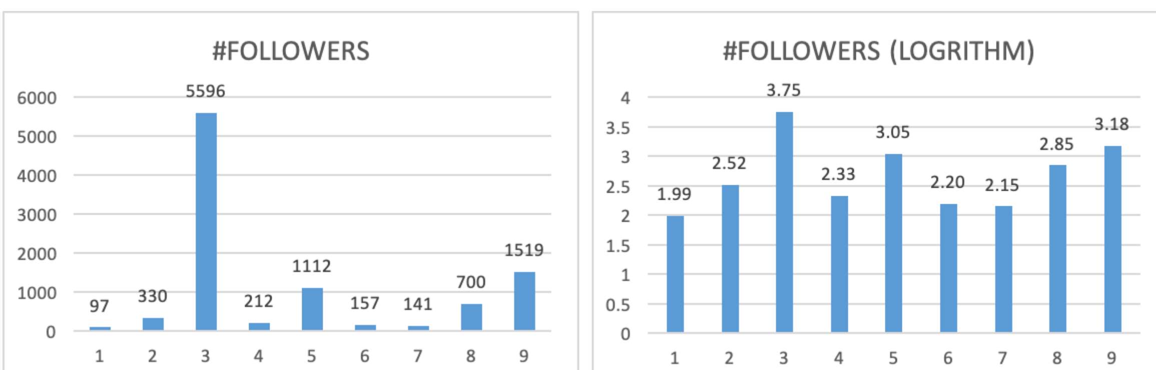

Figure 5.7. Distribution of Number of Followers with *threshold* : 150

Table 5.5.
Comparison between Iterative Evaluation and Follower Numbers as Trust Scores $Threshold : 150$

No.Test	StartDay	OriginalCorr	PredCorr	PredCorr(Base e)	PredCorr(Base 2)
01	65	0.093	0.386	0.250	0.283
02	35	0.387	0.473	0.458	0.464
03	79	-0.205	0.226	-0.170	-0.149
04	135	0.163	0.324	0.108	-0.091
05	28	0.140	0.348	0.308	0.359
06	50	0.468	0.413	0.482	0.471
07	102	0.094	0.237	0.060	0.049
08	10	0.013	0.304	0.081	0.103
09	38	0.429	0.446	0.469	0.468
10	94	0.021	0.195	-0.022	-0.031
AVG	-	0.160	0.335	0.202	0.218

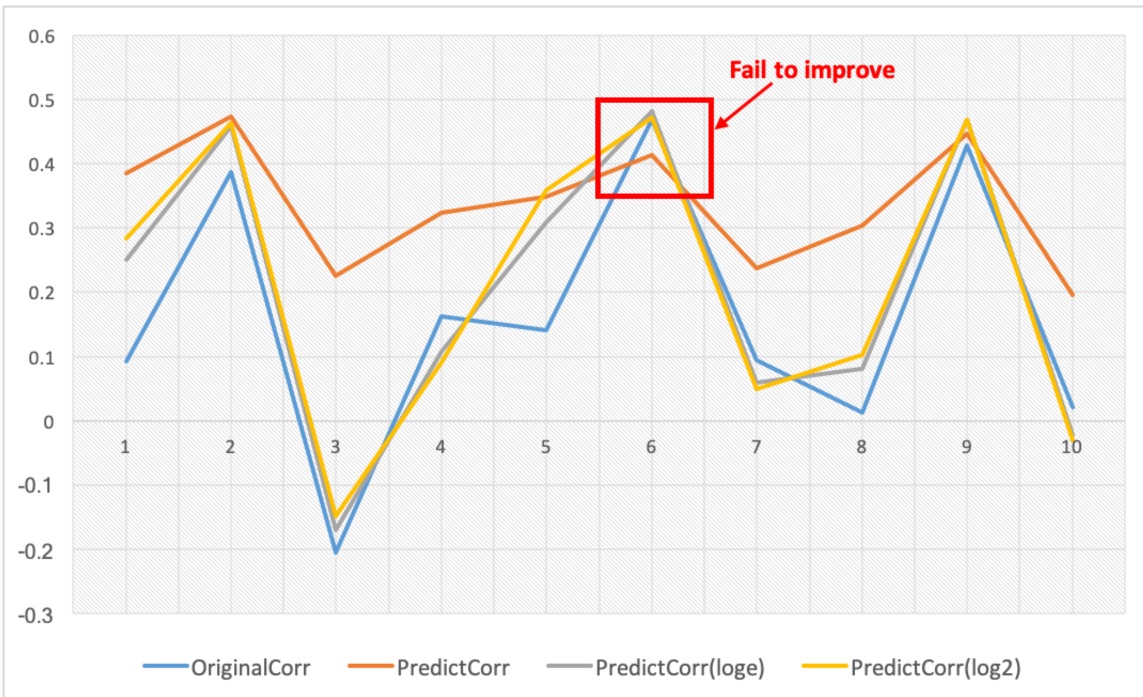

Figure 5.8. Comparison between Iterative Evaluation and Follower Numbers as Trust Scores

6 CONCLUSION

6.1 Summary

This research presented an iterative model for social network trust evaluation aided by data visualization segmentation. The experiments utilized a dataset of tweets about Apples stock price retrieved from Twitter and AAPL stock price as the ground truth. In this paper, we examined the method in two steps: (1) the basic iterative evaluation method; (2) iterative method with segmentation.

In the first step, the basic iterative method has been proved can find out trustworthy information sources with a set of trust score by comparing the original and predicted correlations between user sentiments and change of AAPL price. From the experiment results, we can conclude that with the trust score generated by basic model as weight, the overall correlation is increase to some extent, which means with the trust score, information seekers can tend to give more attention to those who has higher trust scores.

In the second step, we added data visualization segmentation into the model aiming at lower the bias caused by noisy data. From the test results, we found that, with segmentation, the average predicted correlation can be further improved compared with the basic model. What is more, the variance of the improved correlation is much lower than the basic model's result. This indicates the model with segmentation is more stable than the basic one, and the average performance is better as well.

6.2 Discussion

During the experiments, we have a couple of findings and observations which people might be interested in.

1. It is very important to choose a suitable converge threshold ϵ. In the basic model experiments, we set ϵ to 0.01 and 0.001, and find that, for both of them, although the correlations after training are close, the prediction results diverge. For ϵ: 0.001, the model fails to improve. Therefore, we can conclude that if the converge threshold is too small, the model will overfit, and cannot improve the overall correlation.

2. From the basic model, we found a dilemma with threshold for finding active social accounts and including more information sources. As the threshold moves up, we can get fewer and fewer user accounts for investigation. But, higher threshold means denser data points, which leads to more accurate prediction.

3. In this study, finding segmentation points is a subjective process, and unreasonable segments would be counterproductive. During the experiments, we find it is tricky pick good segmentation points. Short-term segments are not encouraged especially at the end of training data.

4. Segmentation should be done before the training, if not, model may not work. We did experiments to see if adding segments after first round of convergence can further improve the correlation. Actually, after first convergence, the trust CoEs for user accounts have changed from 1 to very different values. Although different correlations are calculated for each segments, updating part of the CoEs does not have significant impacts on overall results.

6.3 Future Work

Many different, tests, and experiments have been left for the future due to lack of time (i.e. the experiments with visualization segmentation are very time consuming, requiring user manually picking segment points for each account we investigate in every test). Future work concerns deeper analysis of particular mechanisms, new proposals to try different methods, or simply curiosity. This thesis has been mainly focused on the use of trust evaluation on Twitter sentiments, leaving other kind of dataset outside the scope of the thesis. The following ideas could be tested:

1. Datasets on different topics should be tested by this model. We did experiments only for AAPL stock price. It is possible that the model we proposed here does not apply to other topics. Therefore, testing our method on various datasets need to be done in future. Furthermore, the active information sources are deeply limited by the frequency of user activities. So, in the future, we will consider collecting more data or think of how to use available data more effectively and efficiently.

2. For the visualization part, we use static image for segmentation. To maximize the usage of visual analytics, an interactive system can be built to allow decision makers to combine their creativity and background knowledge with the dynamic graphic presentation of massive data. Utilizing analysis capabilities of todays computer, with advanced visual interfaces, user may directly interact with massive data, and allow them to make well-informed decisions in complex situations.

REFERENCES

[1] danah m. boyd and Nicole B. Ellison. Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 10 2007.

[2] Kristina Lerman and Rumi Ghosh. Information contagion: An empirical study of the spread of news on digg and twitter social networks. In *Fourth International AAAI Conference on Weblogs and Social Media*, 2010.

[3] Andrea H Tapia, Kartikeya Bajpai, Bernard J Jansen, John Yen, and Lee Giles. Seeking the trustworthy tweet: Can microblogged data fit the information needs of disaster response and humanitarian relief organizations. In *Proceedings of the 8th International ISCRAM Conference*, pages 1–10. ISCRAM Lisbon, Portugal, 2011.

[4] Andrew J. Flanagin Miriam J. Metzger. Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59:210–220, 12 2013.

[5] Xia Liu, Alvin C. Burns, and Yingjian Hou. An investigation of brand-related user-generated content on twitter. *Journal of Advertising*, 46(2):236–247, 2017.

[6] Juntao Lai, Tao Cheng, and Guy Lansley. Improved targeted outdoor advertising based on geotagged social media data. *Annals of GIS*, 23(4):237–250, 2017.

[7] Amir Karami, London S. Bennett, and Xiaoyun He. Mining public opinion about economic issues: Twitter and the u.s. presidential election. *Int. J. Strateg. Decis. Sci.*, 9(1):18–28, January 2018.

[8] J Naren. News analytics and dual sentiment analysis for stock market prediction, 12 2017.

[9] Gyz Gidófalvi. Using news articles to predict stock price movements, 2001.

[10] Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *Journal of computational science*, 2(1):1–8, 2011.

[11] Huina Mao, Scott Counts, and Johan Bollen. Predicting financial markets: Comparing survey, news, twitter and search engine data. *arXiv preprint arXiv:1112.1051*, 2011.

[12] Paul Gronke and Timothy E. Cook. Disdaining the media: The american public's changing attitudes toward the news. *Political Communication*, 24(3):259–281, 2007.

[13] Jason Turcotte, Chance York, Jacob Irving, Rosanne M. Scholl, and Raymond J. Pingree. News Recommendations from Social Media Opinion Leaders: Effects on Media Trust and Information Seeking. *Journal of Computer-Mediated Communication*, 20(5):520–535, 06 2015.

[14] Ruohan Li and Ayoung Suh. Factors influencing information credibility on social media platforms: Evidence from facebook pages. *Procedia computer science*, 72:314–328, 2015.

[15] Richard E Petty and John T Cacioppo. The elaboration likelihood model of persuasion. In *Communication and persuasion*, pages 1–24. Springer, 1986.

[16] J. J. Thomas and K. A. Cook. A visual analytics agenda. *IEEE Computer Graphics and Applications*, 26(01):10–13, jan 2006.

[17] Michael Friendly. A Brief History of Data Visualization. In *Handbook of Data Visualization*. Springer, Berlin, Heidelberg, 2008.

[18] W. Cui, Y. Wu, S. Liu, F. Wei, M. X. Zhou, and H. Qu. Context preserving dynamic word cloud visualization. In *2010 IEEE Pacific Visualization Symposium (PacificVis)*, pages 121–128, March 2010.

[19] W. Dou and S. Liu. Topic- and time-oriented visual text analysis. *IEEE Computer Graphics and Applications*, 36(4):8–13, July 2016.

[20] Yiju Hou and Noam Ben-Avi. Reddit multi-dimensional sentiment visualization.

[21] Yefeng Ruan, Arjan Durresi, and Lina Alfantoukh. Using twitter trust network for stock market analysis. *Knowledge-Based Systems*, 145:207 – 218, 2018.