HW1

An Evaluation of Sentiment Classification Tools

Bing-Je Wu

IST 736 – Text Mining

Table of Contents

Introduction

Artificial Intelligence (AI) is the trending topic in this day. Many start to think of implementing the Artificial Intelligence applications to businesses. However, a group of people do not trust Artificial Intelligence and are still rejecting the existence of AI. This assignment is to focus on how people think of Artificial Intelligence, the feel toward AI, in social media such as Facebook and Twitter and evaluate the public sentiment. The following questions will be answered during this assignment:

1. What is the sampling strategy?
2. Would the result in a representative sample of public sentiment toward AI?
3. How to perform the system evaluation process?
4. Are those tools used in the assignment are suitable for the task?

Analysis and Models

Two sentiment analysis tools, NLTK's built-in sentiment analysis tools (NLTK) and Valence Aware Dictionary and sEntiment Reasoner (VADER), were applied and compared during the analysis.

About the Data

Data (https://github.com/bing020815/Syracuse-University/blob/master/IST%20736/HW1/data/AI.csv) were collected from twitter. A Twitter API was intended to be used during the data collection phase. However, due to the limitation of the length of the tweets for free version of Twitter API, a manual collecting approach was implemented instead. In order to get the public sentiment toward AI, a search term, 'Artificial Intelligence', was utilized to search tweets about Artificial Intelligence. The dataset contains 60 tweets with three columns, 'Item', 'Sentiment', and 'Tweet'. The label of each tweet has been assigned the sentiment toward AI based on the judgement of the dataset collector with consideration of content, tones, topic and word selection. 20 tweets were labeled as -1 (negative); another 20 tweets were labeled as 0 (neutral); and the last 20 tweets were labeled as 1 (positive).

Attribute	Description	Data Type
Item	The id number of tweets	Numeric
Sentiment	Dataset collector's sentiment toward AI with output values -1, 0, 1	Numeric
Tweet	The text of tweets with the term, 'Artificial Intelligence', in the content	String

Data Processing

The dataset was clean with some of trailing spaces which was accidently created during the data collection phase. A string method, 'strip()', has been implemented to remove all the possible leading or trailing spaces. A map function has been used to convert the numeric values, '-1', '0' and '1', to classes, 'neg', 'neu' and 'pos', on the 'Sentiment' column. Two extra columns, 'NLTK_score' and 'VADER_score', were added with the 'compound score' from the output of NLTK sentiment analysis tool and VADER sentiment tool. Two features columns, 'NLTK_sentiment' and 'VADER_sentiment', were added by re-engineering the score columns with the condition of threes (1. score is greater than 0.05 as 'pos'; 2. score between 0.05 and -0.05 as 'neu'; and 3. score is less than -0.05 as 'neg'). The last column, 'match', was added by comparing the result of two tools, NLTK sentiment analysis tool and VADER sentiment tool. If the outputs from both tools are the same, it should return '0'. Otherwise, it should be '1'. Some of stop words were removed for analysis when building plots.

Results

The dataset has three different sentiment results from three different sentiment analysis system for each of tweets. One is from NLTK sentiment tool; another is from VADER sentiment tool, and the other is from public sentiment toward AI. The distributions of the sentiment levels are quite different. For example, by looking at the figure 1, it is easy to see that the distributions of the sentiment levels from NLTK and VADER are slightly different. NLTK tends to have more neutral tweet than VADER. Both NLTK and VADER has the same amount number of positive tweets. However, public sentiment has evenly distributed classes in 'pos', 'neg', and 'neu'. The reason why the distributions are different could be that NLTK tool and VADER tool cannot analyze the sentiment of text as human beings. The labeling system used for public sentiment toward AI is specific designed for detecting if people feel positive or negative on approaching AI. On the other hand, NLTK and VADER evaluate each of tweets base on the polarity of words. Although some tweets mention the term, 'Artificial Intelligence', the content itself is not centering around the same topic. Thus, both tools cannot detect the tweets not showing the attitude about AI as what human can normally do.

Figure 1: Distribution plots for NLTK Sentiment Tool, VADER Sentiment Tool and Public Sentiment Toward AI.

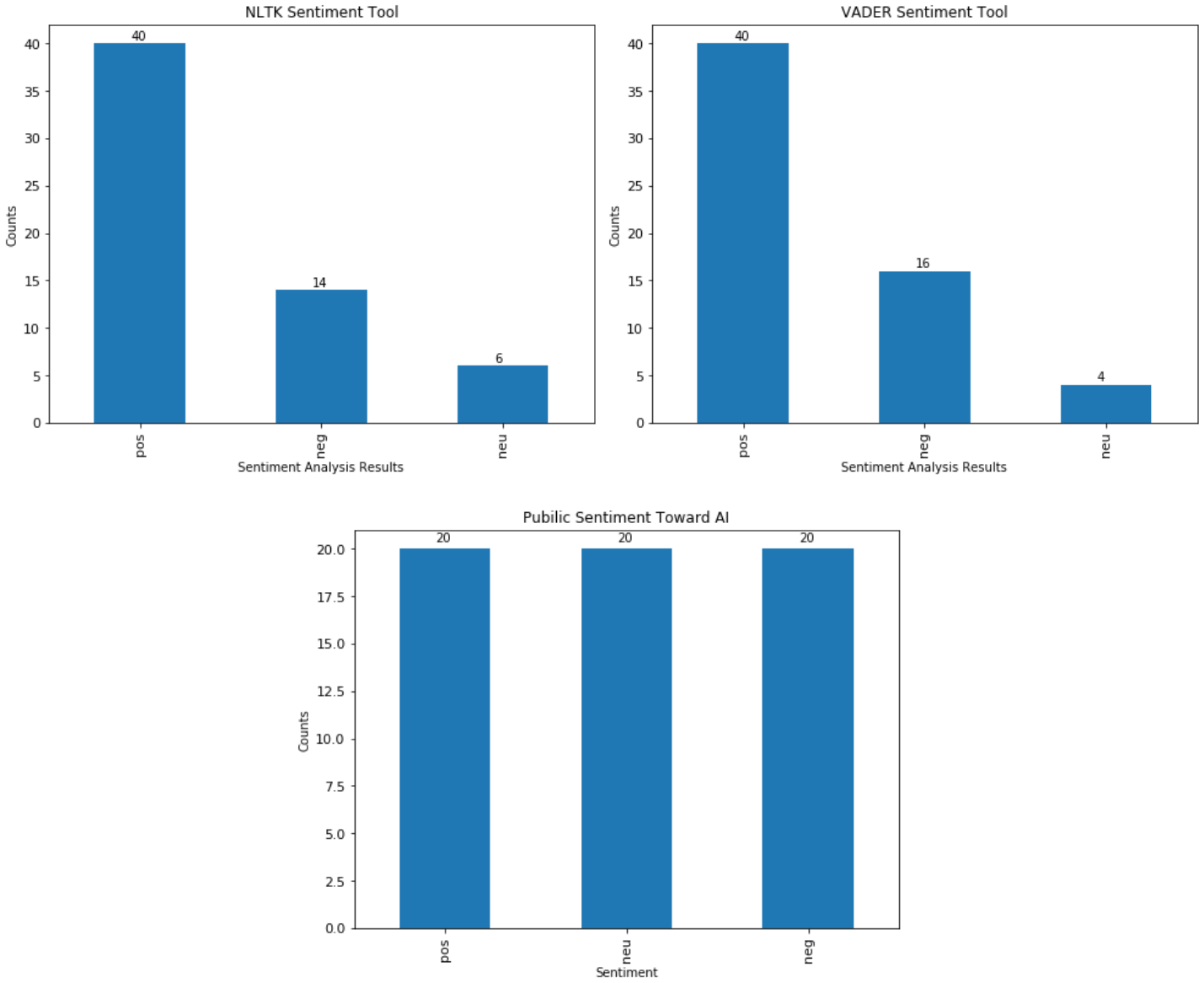


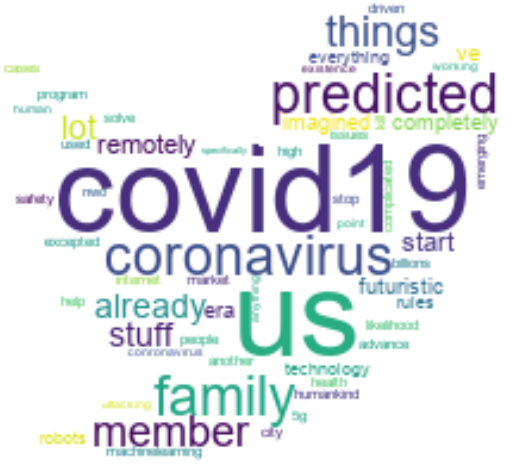
Figure 2: NLTK_pos word cloud and word counts

	Counts
human	7
think	6
world	6
will	6
time	5
make	5
really	4
machine	4
people	3
need	3

Figure 3: NLTK_neg word cloud and word counts

	Counts
will	6
need	3
us	3
now	3
countries	3
going	3
hands	2
teams	2
current	2
fucked	2

Figure 4: NLTK_neu word cloud and word counts



	Counts
coronavirus	2
member	2
things	2
predicted	2
covid19	2
us	2
family	2
program	1
remotely	1
robots	1

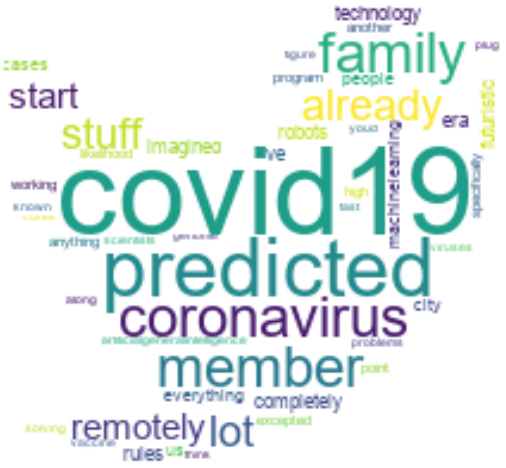
Figure 5: VADER_pos word cloud and word counts

	Counts
human	7
will	6
think	6
make	5
time	5
really	4
world	4
things	3
day	3
data	3

Figure 6: VADER_neg word cloud and word counts

	Counts
will	6
countries	3
going	3
us	3
now	3
need	3
world	3
teams	2
hands	2
jobs	2

Figure 7: VADER_neu word cloud and word counts



	Counts
member	2
predicted	2
covid19	2
family	2
coronavirus	2
point	1
working	1
problems	1
program	1
remotely	1

Figure 8: Public_toward_AI _pos word cloud and word counts

	Counts
will	5
world	5
machine	4
human	3
technology	3
day	3
countries	3
need	3
back	3
covid19	2

Figure 9: Public_toward_AI_neg word cloud and word counts

	Counts
will	4
us	4
people	3
don	3
part	3
think	3
teams	2
company	2
virus	2
feel	2

Figure 10: Public_toward_AI _neu word cloud and word counts

	Counts
member	2
predicted	2
covid19	2
family	2
coronavirus	2
point	1
working	1
problems	1
program	1
remotely	1

Figure 10: Not matching tweets

Sentiment	Tweets	NLTK_score	NLTK_sentiment	VADER_score	VADER_sentiment	match
1	Life is eternal in HELL, Mind Controlling itself or Artificial Intelligence the Machine that always wins the Game through Artificial means in an artificial world. The replications have no chance to beat the Master of the World/Satan/AI.	0.068800	pos	-0.070300	neg	0.000000
0	5G is about the internet of things and artificial intelligence... and billions $ in an emerging market driven by those things. It's not about us or our health/safety. #nwo	0.000000	neu	0.476700	pos	0.000000
1	Artificial Intelligence is used to advance and solve most complicated issues for and against humankind. So, can artificial intelligence help stop and kill the conronavirus attacking human existence?	0.016400	neu	-0.051600	neg	0.000000

To evaluate each of systems, other techniques, such as word cloud and words counts, were used for the analysis. Figure 2, Figure 3 and Figure 4 show the word cloud and words count for NLTK system.

Conclusion

The fashion-MINST dataset was used for the eight algorithms. Among the 10 classes, 'T-shirt/top', 'Trouser', 'Pullover',

Reference

[1] Tensorflow.org. Web Link ; [2] Udemy – TensorFlow 2.0. Web Link

[3] MNIST database, from Wikipedia. Web Link ; [4] SGDClassifier, from scikit-learn. Web Link

[5] LogisticRegression, from scikit-learn. Web Link1 ; [6] SVC, from scikit-learn. Web Link1 , Web Link2

[7] Naive Bayes, from scikit-learn. Web Link ; [8] KNeighborsClassifier, from scikit-learn. Web Link

[9] RandomForestClassifier, from scikit-learn. Web Link ; [10] Keras, from TensorFlow. Web Link

[11] Stochastic gradient descent, from Wikipedia. Web Link ; [12] Precision and recall, from Wikipedia. Web Link

[13] Adam: A Method for Stochastic Optimization. Web Link ; [14] 2D Visualization of a Convolutional Neural Network. Web Link

[15] Naive Bayes classifier, from Wikipedia. Web Link ; [16] Machine Learning Course with Python. Web Link

[17] Understanding Convolutions for Deep Learning. Web Link ; [18] Machine Learning & Deep Learning Fundamentals Web Link

[19] About Convolutional Layer and Convolution Kernel. Web Link ; [20] Neural networks Web Link