Twitter Sentiment About Apple and Google products

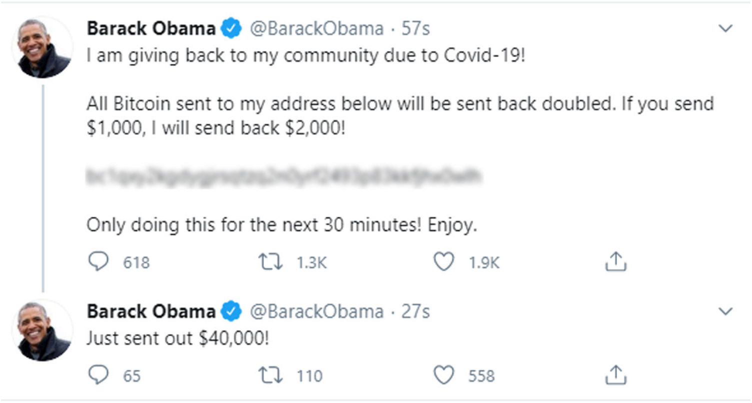
Prepared by: Kai Luo

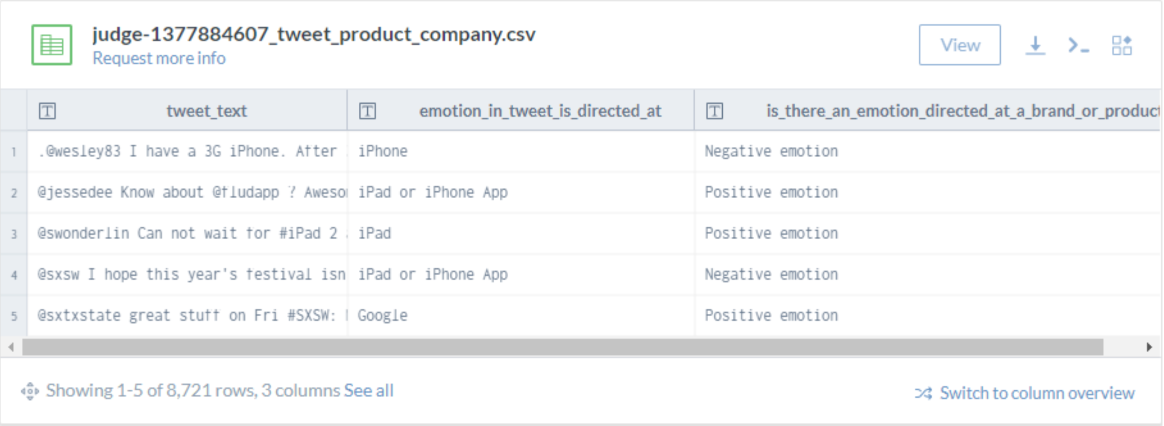
Introduction

- Background
- Exploratory Data Analysis
- Model
- Model Performance
- Recommendations & Future Work
- Summary

Background

- Judge Emotion About Brands and Products
- Contributors evaluated tweets about multiple brands and products. The crowd was asked if the tweet expressed positive, negative, or no emotion towards a brand and/or product. If some emotion was expressed they were also asked to say which brand or product was the target of that emotion.
  - Added: August 30, 2013 by Kent Cavender-Bares
  - Data Rows: 9093
- Source: https://www.crowdflower.com/data-for-everyone/




judge-1377884607_tweet_product_company.csv
Request more info
View

	tweet_text	emotion_in_tweet_is_directed_at	is_there_an_emotion_directed_at_a_brand_or_product
1	.@wesley83 I have a 3G iPhone. After	iPhone	Negative emotion
2	@jessedee Know about @fludapp ? Aweso	iPad or iPhone App	Positive emotion
3	@swonderlin Can not wait for #iPad 2	iPad	Positive emotion
4	@sxsw I hope this year's festival isn	iPad or iPhone App	Negative emotion
5	@sxtxstate great stuff on Fri #SXSW:	Google	Positive emotion

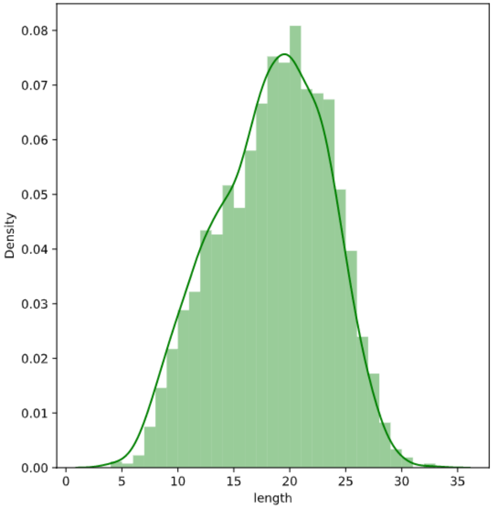
Showing 1-5 of 8,721 rows, 3 columns See all
Switch to column overview

Exploratory Data Analysis

- Both positive and negative sentiment tweets have mean word count of 18 and standard deviation of 5.
- Tweet word length have no correlation to sentiment.

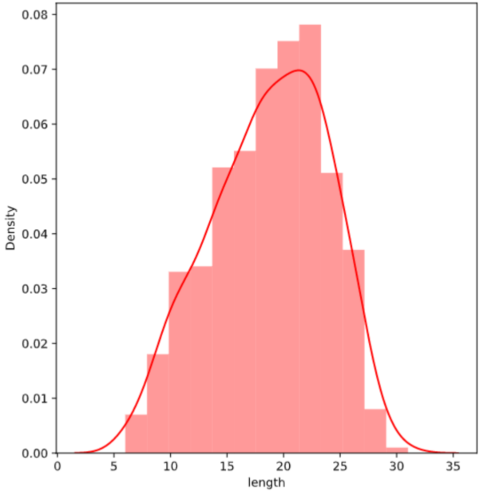
Distribution of text length for positive sentiment tweets.

	length
count	2672.0
mean	18.2
std	4.95
min	4.0
25%	15.0
50%	19.0
75%	22.0
max	33.0



Distribution of text length for negative sentiment tweets.

	length
count	519.0
mean	18.72
std	5.13
min	6.0
25%	15.0
50%	19.0
75%	23.0
max	31.0

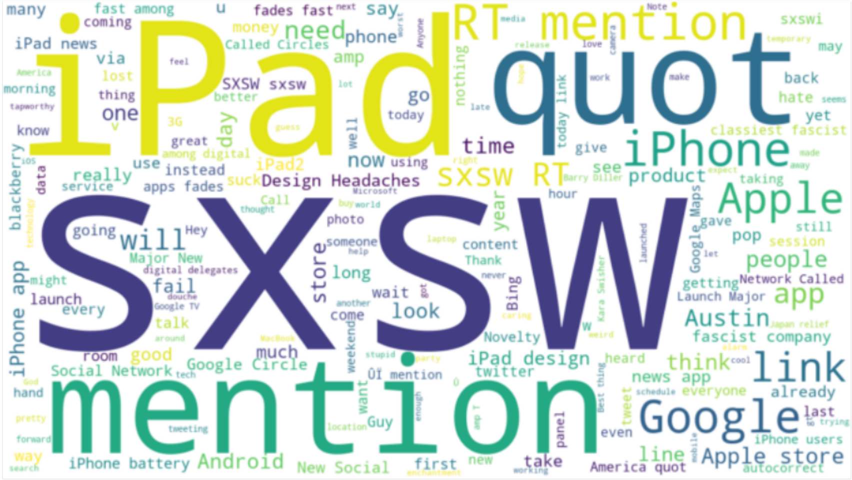


Exploratory Data Analysis

- Top words for both positive and negative tweets:
  - SXSW, mention, iPad, iPhone, Apple, Google
- Simple keyword search won't indicate positive or negative sentiment.

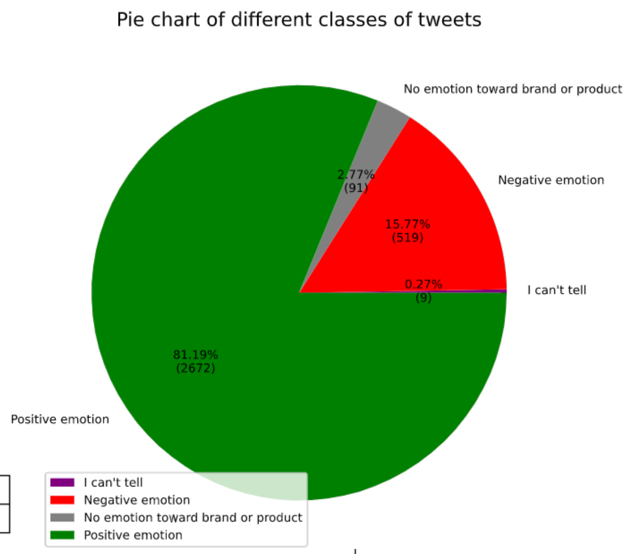
Most common words in positive sentiment tweets.

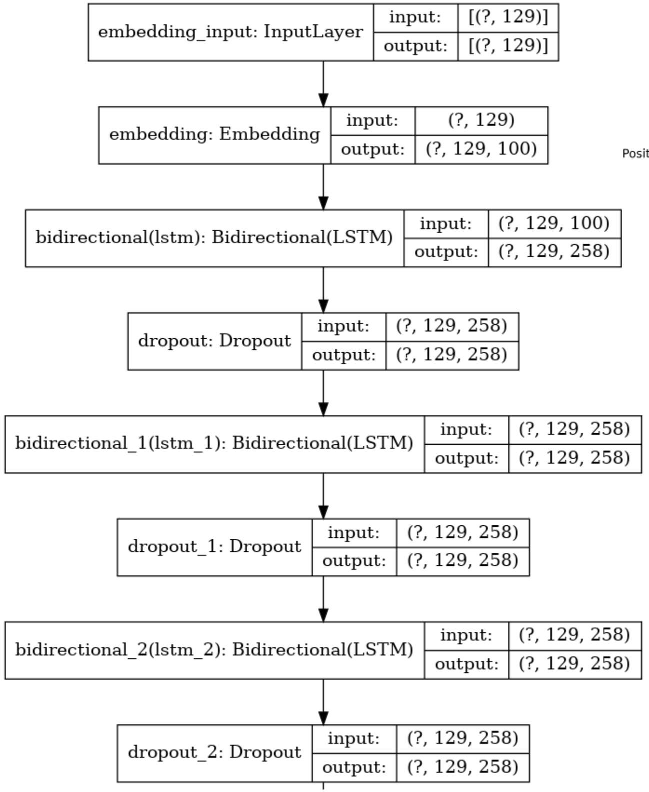
Most common words in negative sentiment tweets.

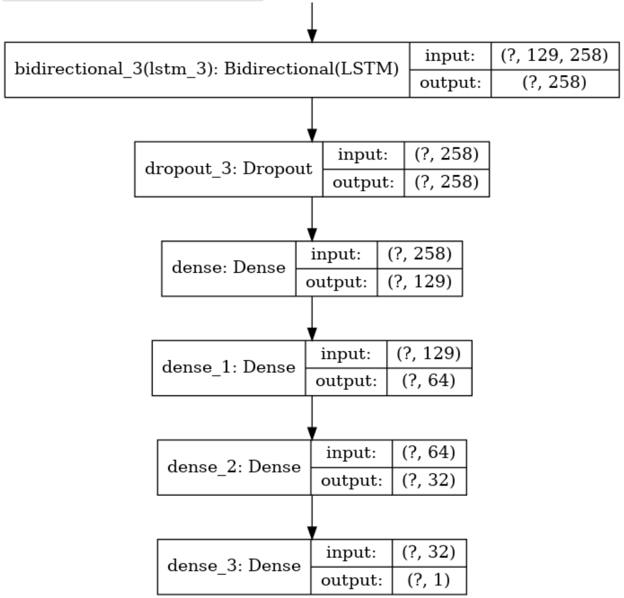


Best Model

- Word count and keyword search don’t identify sentiment, therefore we will try to fit a deep learning model.
- There are a lot more positive emotions then the rest.
  - In order to fully utilize all our available data, we will make the target into only two categories: positive and non-positive.
  - Then undersample based on the number of non-positive data.
- Many variation of the word_embedding-LSTM-dropout-dense layer models were explored, this is the one with the best performance.



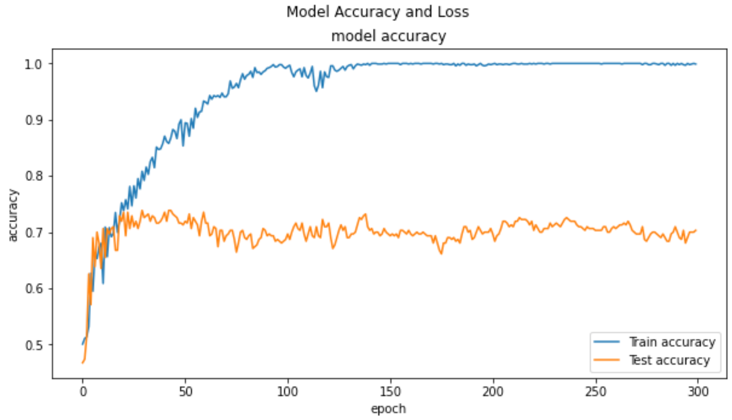


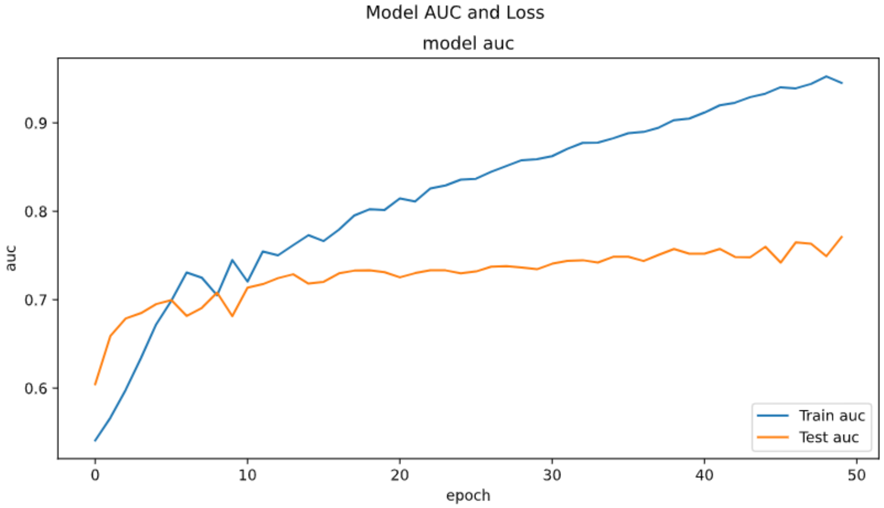


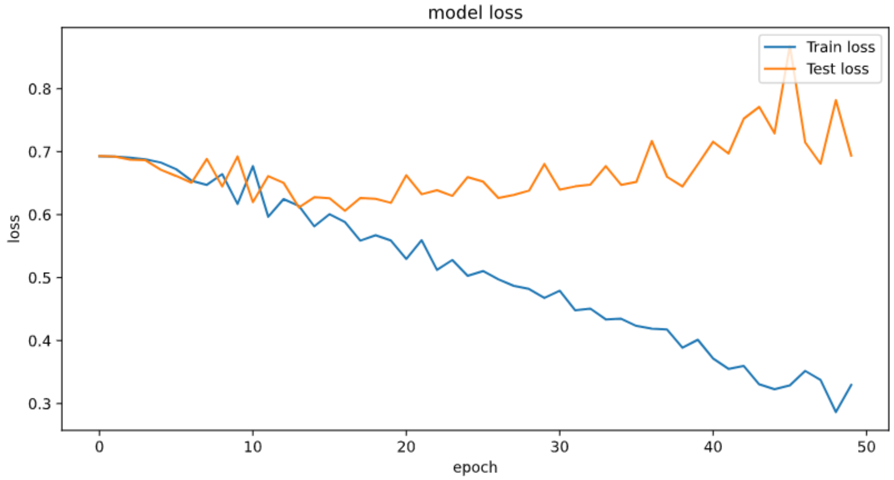
Best Model Performance

- Average precision is 0.71.
- Average recall is 0.70.
- AUC score is around 0.70.
- Further training doesn't improve performance.
- This model can be used to predict how customers like the product!

	precision	recall	f1-score	support
0	0.76	0.65	0.70	165
1	0.66	0.76	0.71	145
accuracy			0.70	310
macro avg	0.71	0.71	0.70	310
weighted avg	0.71	0.70	0.70	310

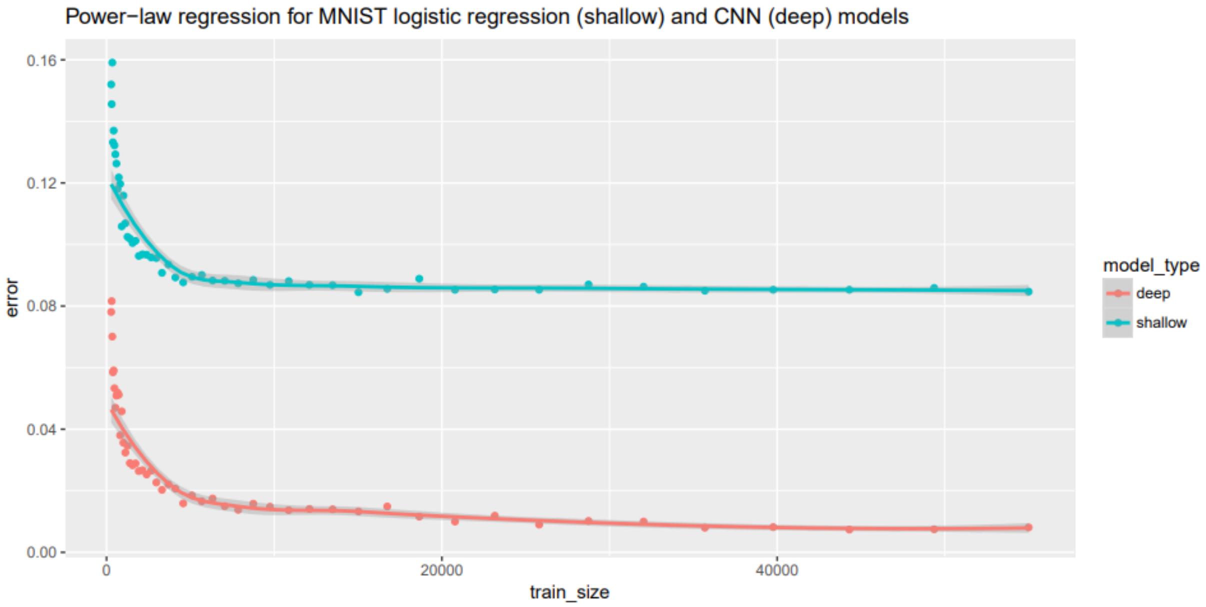






Impact of Data Size

- Modified National Institute of Standards and Technology database Hand Written Digit
- Plot source: https://web.science.mq.edu.au/~mjohnson/papers/Johnson17Power-talk.pdf
- For model training, quantity of data have large impact on error (1-accuracy) for low training data size.



Recommendations and Future Works

1) Gather more data.
   - Sometimes more data improve model performance, but too much may increase training complexity by too much!
2) Build a continuous learning model.
   - In case there are new terms and slogans: language changes over time.
3) Pick an acceptable classification threshold value from the AUC ROC curve
   - Pick an acceptable risk.

Summary

- Analyzed tweets about product sentiments.
- Word count and keywords don’t predict sentiments, so a word-embedding model was created.
- Model has precision of 0.71, and recall of 0.70.
- Training the model for more epoch won’t improve the performance.
- That’s why, recommendation:
  - Gather more data.
  - Build a continuous learning model.
  - Pick an acceptable classification threshold value from the AUC ROC curve

Thank you for you attention!

Any questions?