



# Sentiment Analysis and Feature Extraction on Phone Product Reviews

Shuyu Ding, Jiatuan Luo, Qihang  
Zhang



# Motivation/Project Goals

- Build a deep learning based sentiment classifier
- Extract product features for popular phone brands
- Provide potential business insights to manufacturers

# Data

- **Amazon Reviews: Unlocked Mobile Phone from Kaggle**
  - 413,840 data entries, keep 413,778 reviews
  - Six Features: Product Name, **Brand Name**, **Rating**, **Reviews**, Price, Review Votes
  - New column: **Label**
- **Properties of the data**
  - The vocabulary size is 60,689 after data cleaning
  - The average number of sentences per review is around 2.958
  - The average number of tokens per sentence is 13.92

# Approaches

- **Baselines: K-Nearest Neighbors(KNN), Random Forest, Naive Bayes**
  - Train a Word2Vec model for word embedding
  - Create a bag of K-Means centroids for each review and save as the designed matrix
  - Perform hyperparameter tuning and cross-validation with the training data
  - Evaluate model performances with the validation data
- **Proposed model: LSTM(Long Short Term Memory) Networks**
  - Tokenize the most frequent 50,000 word with Keras word tokenizer
  - Truncate and pad the input sequences with the same length
  - Build a LSTM network with one embedding layer and two LSTM layers

# Approaches

- **Feature Extraction**

- Perform POS tagging technique on each word
- Filter out features that are either considered as noun or noun phrase
- Find the adjectives within a certain distance to the feature
- Calculate the corresponding sentiment score

# Experiments and Evaluations

- **Experiment setup**

- Training set 80%, validation set 20%
- All experiments run on **CPU**
- **Baselines:** Grid-search cross validation method with 3 folds
- **Proposed Method:** Manually hyper parameters tuning

Hyperparameter	Tuning Range
Embedding Vector Length	[32,128]
Batch Size	[256,1024]
Units Per Layer	[32,128]
Drop Out	[0.1-0.4]
Recurrent Drop Out	[0.1-0.4]
Number of Epoches	[5-10]

- **Evaluation Metrics:**

- F1 score, Recall, Precision and Accuracy

- **Feature Extraction**

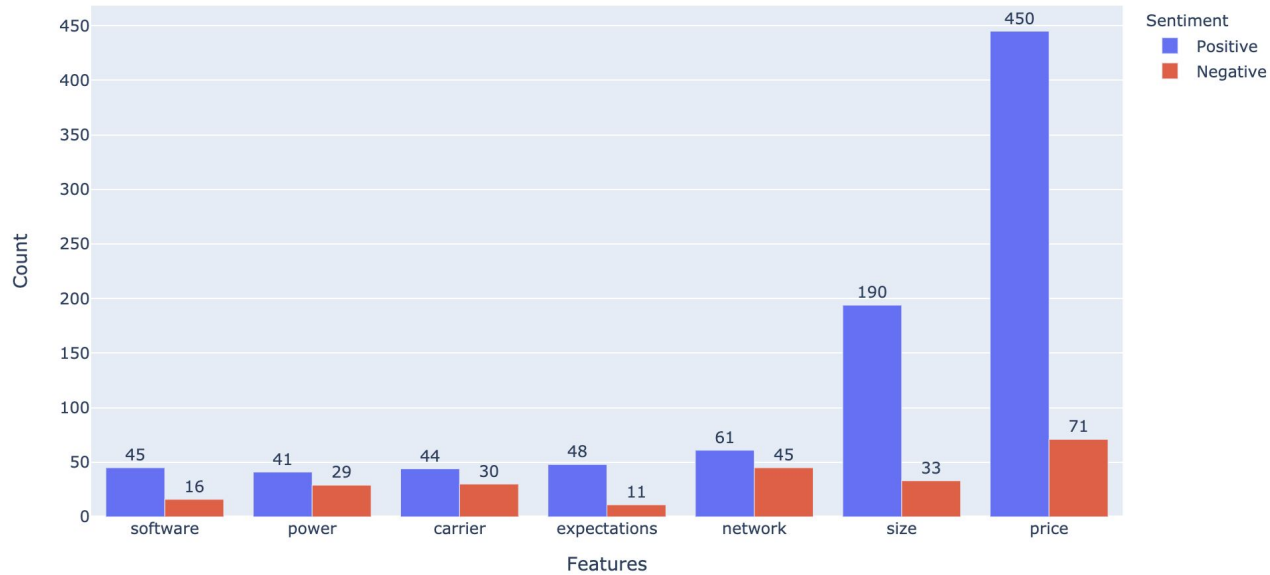
- Cluster the reviews by phone brand
- Run the algorithm on top 7 brands

# Results

Model	Accuracy	Precision (weighted)	Recall (weighted)	F1 (weighted)
Random Forest	0.815	0.838	0.815	0.822
KNN	0.862	0.846	0.852	0.847
Naive Bayes	0.674	0.628	0.674	0.610
LSTM	<b>0.877</b>	<b>0.895</b>	<b>0.852</b>	<b>0.874</b>

# Results

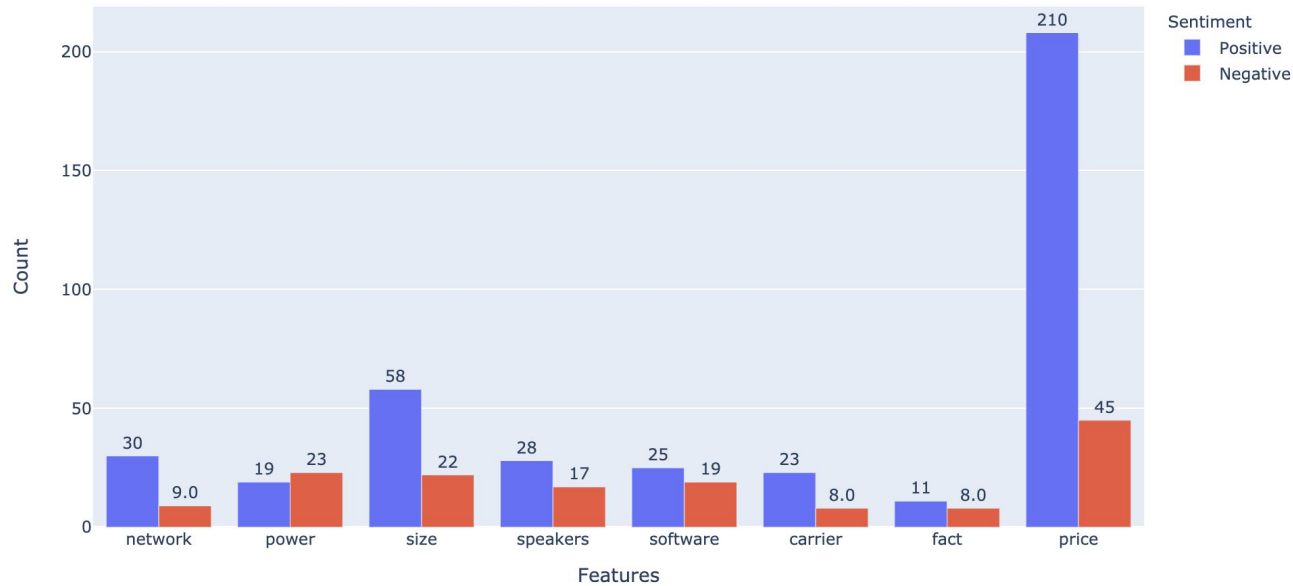
Samsung: Positive/Negative Sentiment by Features





# Results

LG: Positive/Negative Sentiment by Features



# Conclusion/Future work

- Evaluate feature extraction results
- Modify model structure to further improve performance
- Apply to different datasets

Thank You!