Sentiment Analysis of Amazon Review

Roman Bellisari, Michael Delcid, Matthew Lepis

CPE 695: Applied Machine Learning

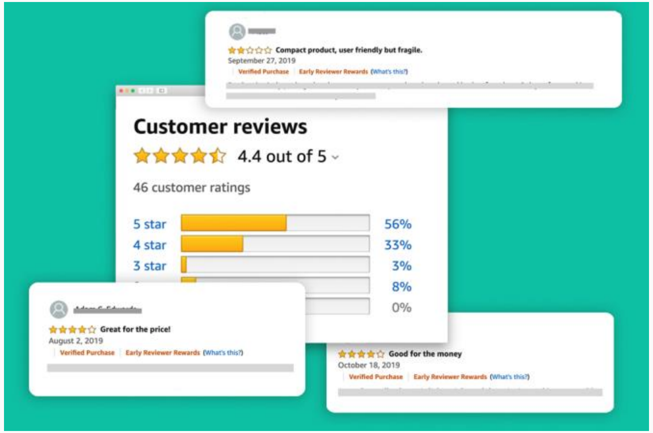
Project Scope

Objective

Determine the sentiment of Amazon user reviews

General Approach

Text Classification & Natural Language Processing (NLP)



[redacted]

☆☆☆☆☆ Compact product, user friendly but fragile.
September 27, 2019
Verified Purchase | Early Reviewer Rewards (What's this?)

Customer reviews

☆☆☆☆☆ 4.4 out of 5

46 customer ratings

5 star 56%
4 star 33%
3 star 3%
8%
0%

[redacted]

☆☆☆☆☆ Great for the price!
August 2, 2019
Verified Purchase | Early Reviewer Rewards (What's this?)

☆☆☆☆☆ Good for the money
October 18, 2019
Verified Purchase | Early Reviewer Rewards (What's this?)

The Dataset

Source

Originally: *SNAP* (Stanford Network Analysis Project)

~35 million samples from June 1995 to March 2013

Derived: Registry of Open Data on AWS

Reduced to 3.65 million samples; equally distributed between star ratings 1 through 5.

Dataset Features

1. Class index (star rating) [1 to 5]
2. Review title
3. Review text________

NLP Developments

- Transfer Learning
- Multilingual NLP
- Automating Customer Service

Known Solutions

In previous works, researchers identified the positive or negative semantic orientation of adjectives within a collection of text. The researchers achieved high precision in classifying this word class.

In another related project, researchers identified product features by analyzing customer opinions and ranking them by frequency. This type of NLP was used to isolate key features of a product.

In a paper released by NYU, researchers utilized CNN for text classification. The researchers achieved high accuracy without spending large amount of time on preprocessing data.

Our Solution

- We retrieve the semantic orientation for more than one class of word and accurately tokenize them to use in the sentence analysis of product reviews
- The team mines product features in reviews for the purpose of classifying them within a five star ranking system
- These ratings are then compared against the actual rating of the product and will be used to distinguish between real and fake reviews left by customers

ML Techniques Used

- Random Forest
- Ensemble Learning
- Multi Layer Perceptron (MLP) Artificial Neural Network (ANN)
- Overfitting Reduction
- Dimensionality Reduction: Principal Component Analysis
- Dropout Layers
- L2 Regularization
- Learning Rating Optimization
- Model Evaluation & Prediction Generation

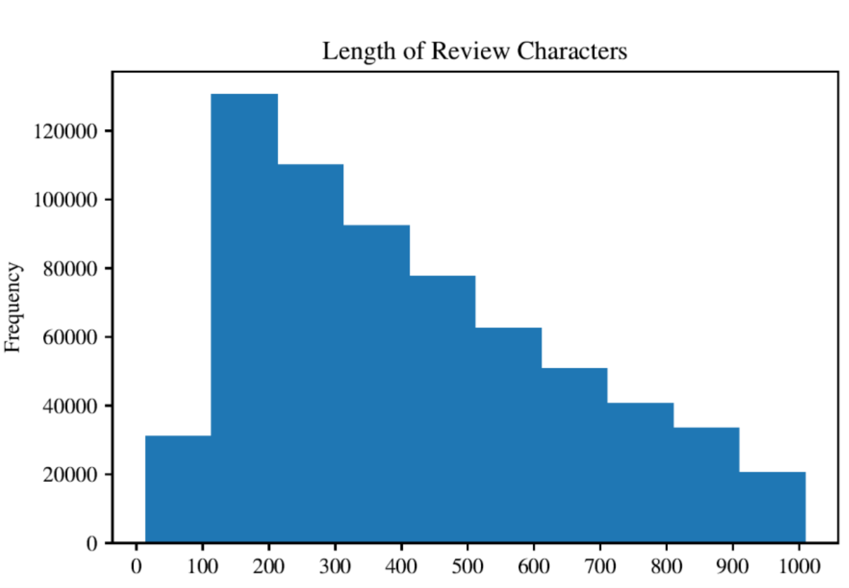
Technologies Used

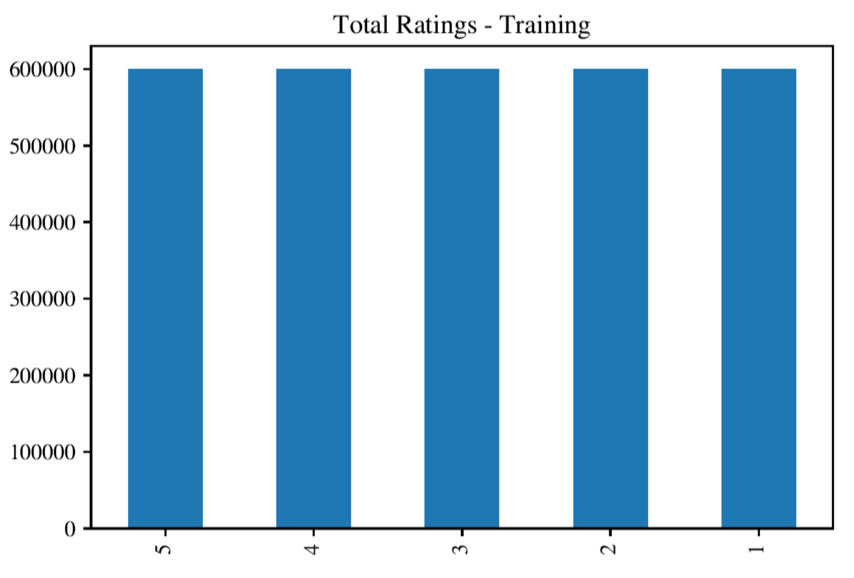
NumPy
matplotlib
spaCy
pandas
GENSIM
topic modelling for humans
scikit
learn
K
python
NLTK

NLP Processing

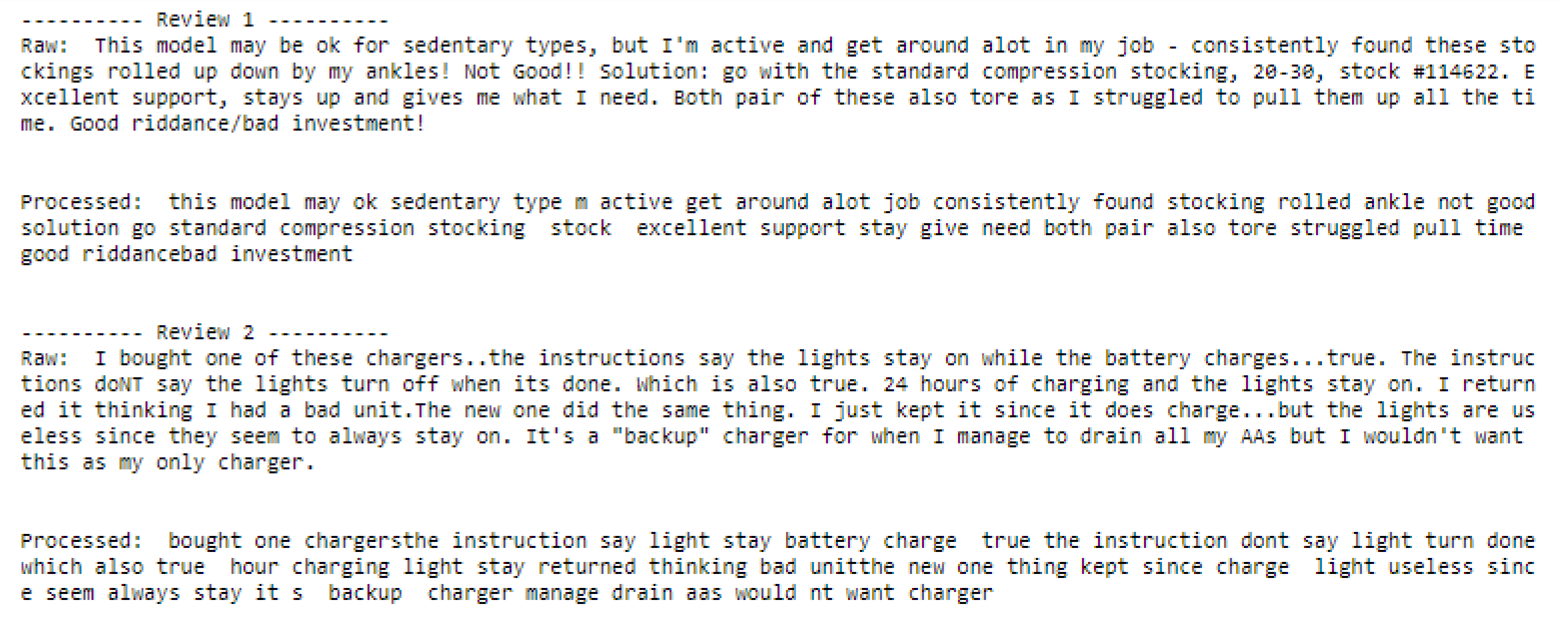
- Output Class Distribution Leveling
- Language Detection
- Space Removal
- Foreign Character Removal
- Stop Word Removal
- Punctuation Removal
- Case Transformation
- Tokenization
- Lemmatization
    - POS Processing
- Vectorization
    - Gensim vs Keras

Results


Length of Review Characters
Frequency
0
20000
40000
60000
80000
100000
120000
0
100
200
300
400
500
600
700
800
900
1000


Total Ratings - Training
0
100000
200000
300000
400000
500000
600000
5
4
3
2
1

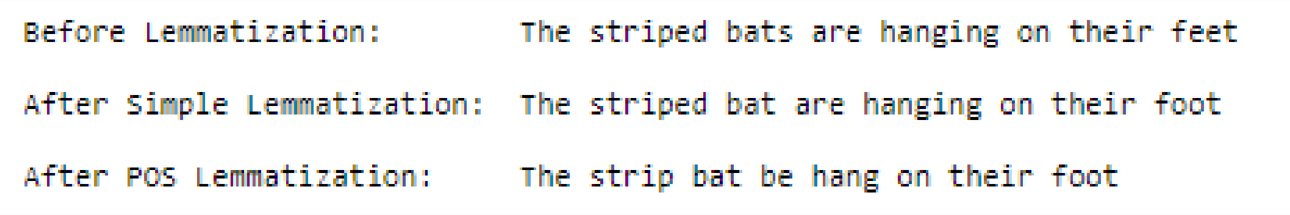
NLP Results



```
---------- Review 1 ----------
Raw:  This model may be ok for sedentary types, but I'm active and get around alot in my job - consistently found these sto
ckings rolled up down by my ankles! Not Good!! Solution: go with the standard compression stocking, 20-30, stock #114622. E
xcellent support, stays up and gives me what I need. Both pair of these also tore as I struggled to pull them up all the ti
me. Good riddance/bad investment!


Processed:  this model may ok sedentary type m active get around alot job consistently found stocking rolled ankle not good
solution go standard compression stocking  stock  excellent support stay give need both pair also tore struggled pull time
good riddancebad investment


---------- Review 2 ----------
Raw:  I bought one of these chargers..the instructions say the lights stay on while the battery charges...true. The instruc
tions doNT say the lights turn off when its done. Which is also true. 24 hours of charging and the lights stay on. I return
ed it thinking I had a bad unit.The new one did the same thing. I just kept it since it does charge...but the lights are us
eless since they seem to always stay on. It's a "backup" charger for when I manage to drain all my AAs but I wouldn't want
this as my only charger.


Processed:  bought one chargersthe instruction say light stay battery charge  true the instruction dont say light turn done
which also true  hour charging light stay returned thinking bad unitthe new one thing kept since charge  light useless sinc
e seem always stay it s  backup  charger manage drain aas would nt want charger
```

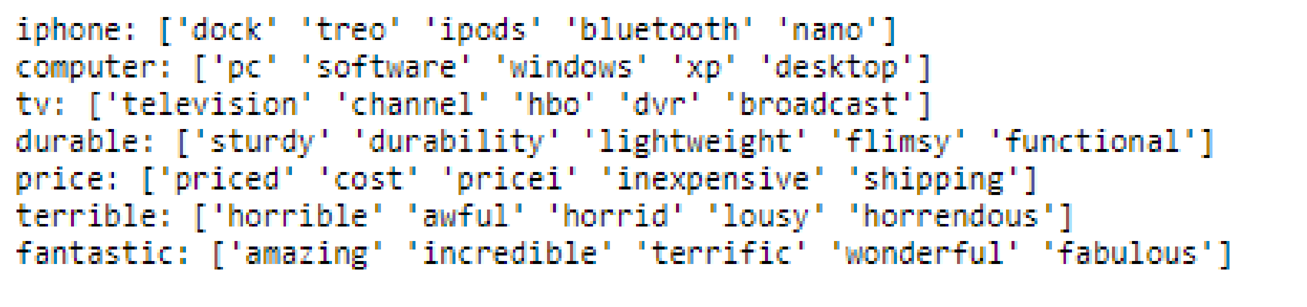
NLP Results

Lemmatization



```
Before Lemmatization:         The striped bats are hanging on their feet

After Simple Lemmatization:   The striped bat are hanging on their foot

After POS Lemmatization:      The strip bat be hang on their foot
```

Gensim
Word2Vec.most_similar



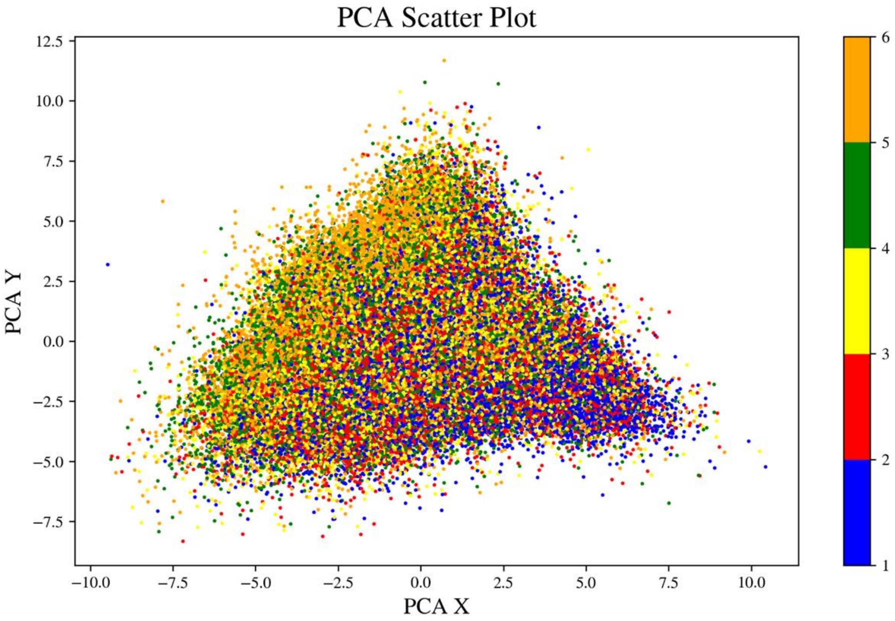
```
iphone: ['dock' 'treo' 'ipods' 'bluetooth' 'nano']
computer: ['pc' 'software' 'windows' 'xp' 'desktop']
tv: ['television' 'channel' 'hbo' 'dvr' 'broadcast']
durable: ['sturdy' 'durability' 'lightweight' 'flimsy' 'functional']
price: ['priced' 'cost' 'pricei' 'inexpensive' 'shipping']
terrible: ['horrible' 'awful' 'horrid' 'lousy' 'horrendous']
fantastic: ['amazing' 'incredible' 'terrific' 'wonderful' 'fabulous']
```

Results

Input Vector Shape:
(300000, 100)

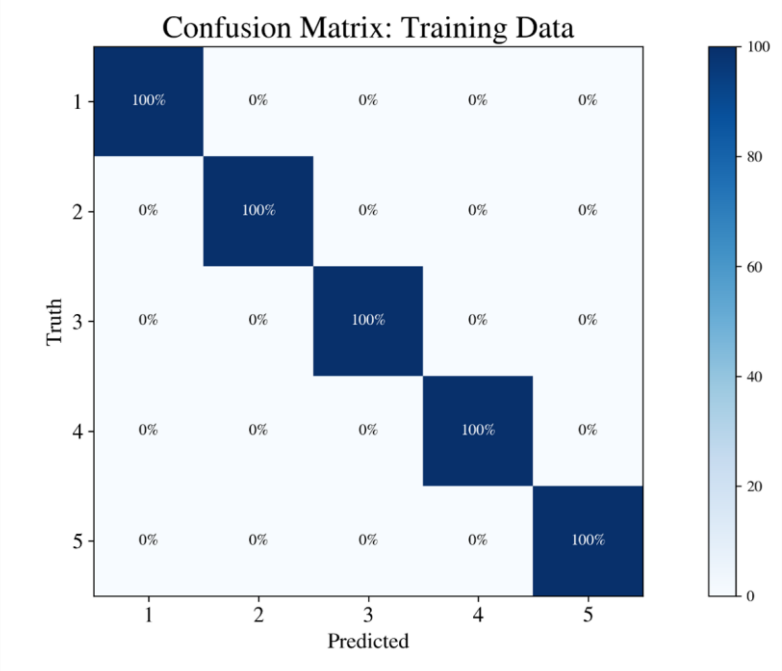
PCA Vector Shape:
(300000, 2)

Not Linearly
Separable



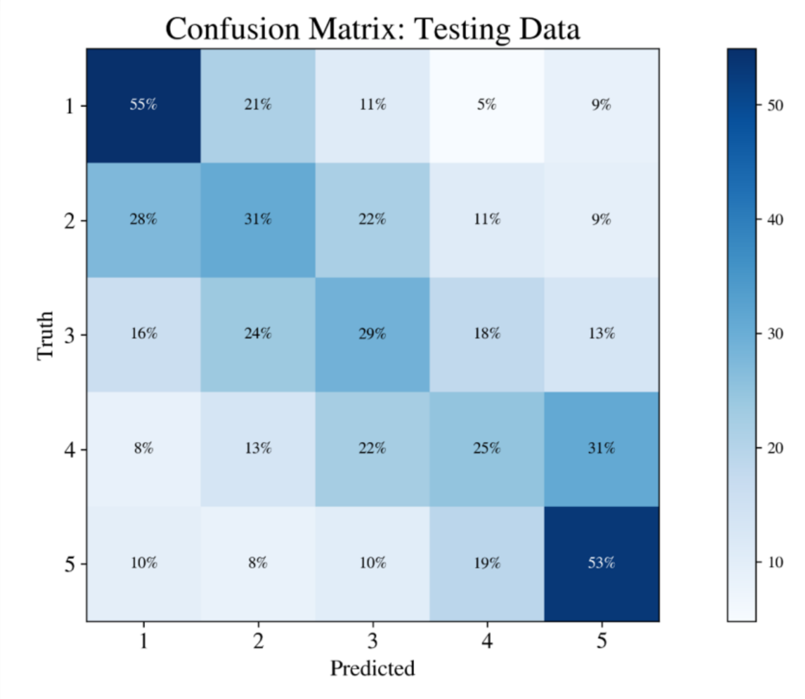
Random Forest

Testing Set Accuracy: 41.27%



Confusion Matrix: Training Data

Truth \ Predicted	1	2	3	4	5
1	100%	0%	0%	0%	0%
2	0%	100%	0%	0%	0%
3	0%	0%	100%	0%	0%
4	0%	0%	0%	100%	0%
5	0%	0%	0%	0%	100%

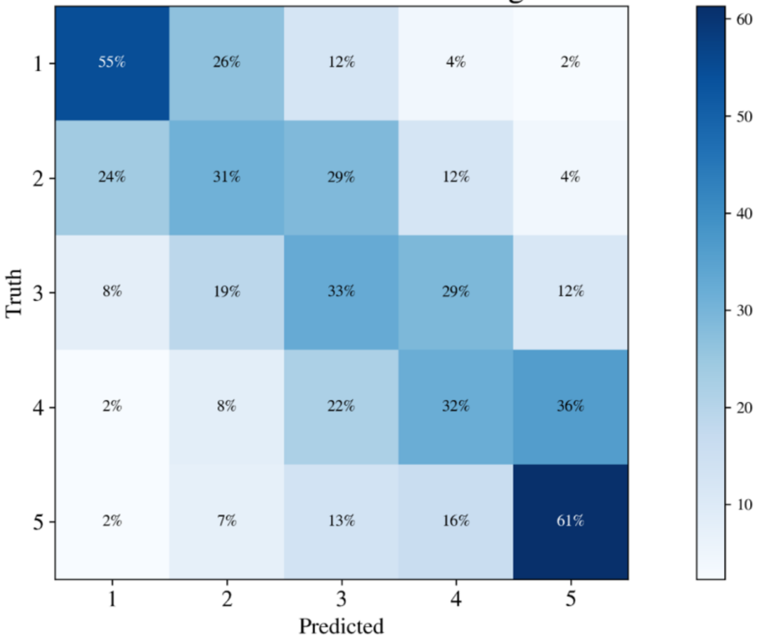


Confusion Matrix: Testing Data

Truth \ Predicted	1	2	3	4	5
1	55%	21%	11%	5%	9%
2	28%	31%	22%	11%	9%
3	16%	24%	29%	18%	13%
4	8%	13%	22%	25%	31%
5	10%	8%	10%	19%	53%

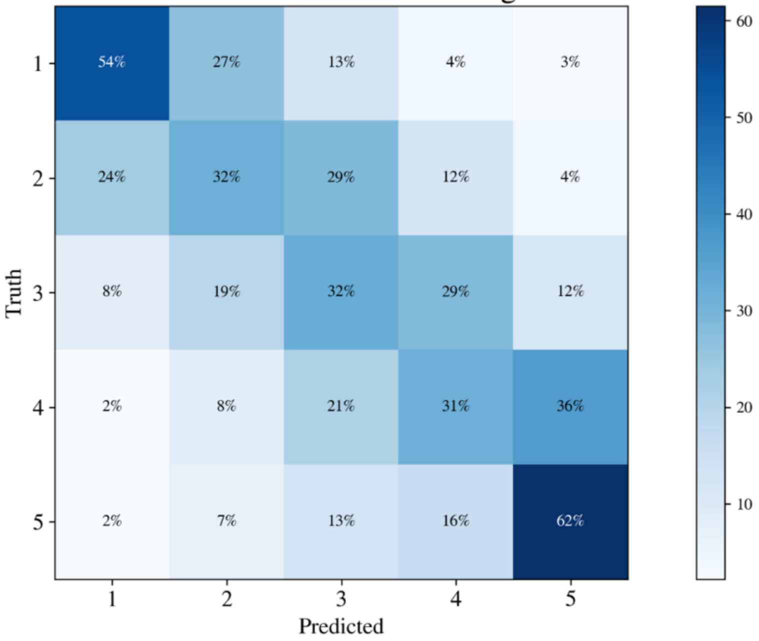
MLP

Testing Set Accuracy: 42.02%



MLP Confusion Matrix: Training Data

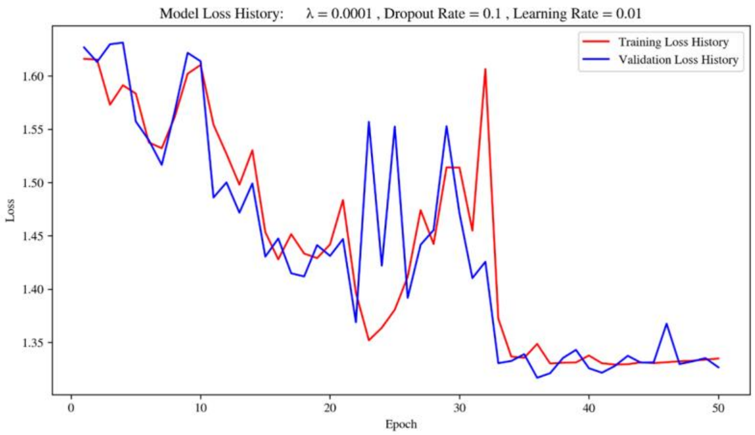
Truth \ Predicted	1	2	3	4	5
1	55%	26%	12%	4%	2%
2	24%	31%	29%	12%	4%
3	8%	19%	33%	29%	12%
4	2%	8%	22%	32%	36%
5	2%	7%	13%	16%	61%

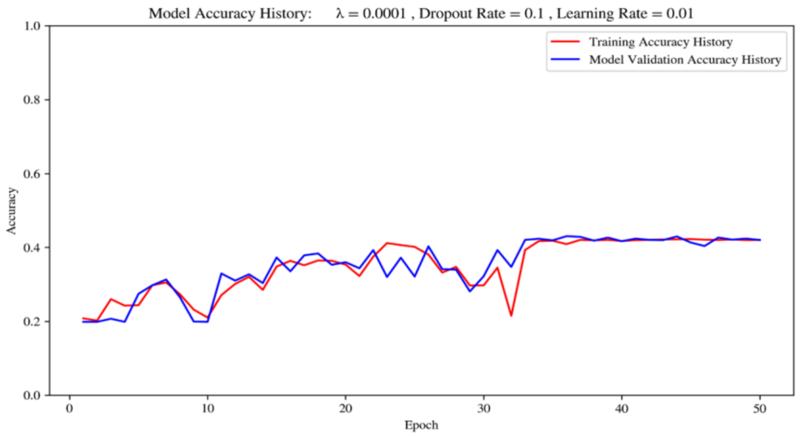


MLP Confusion Matrix: Testing Data

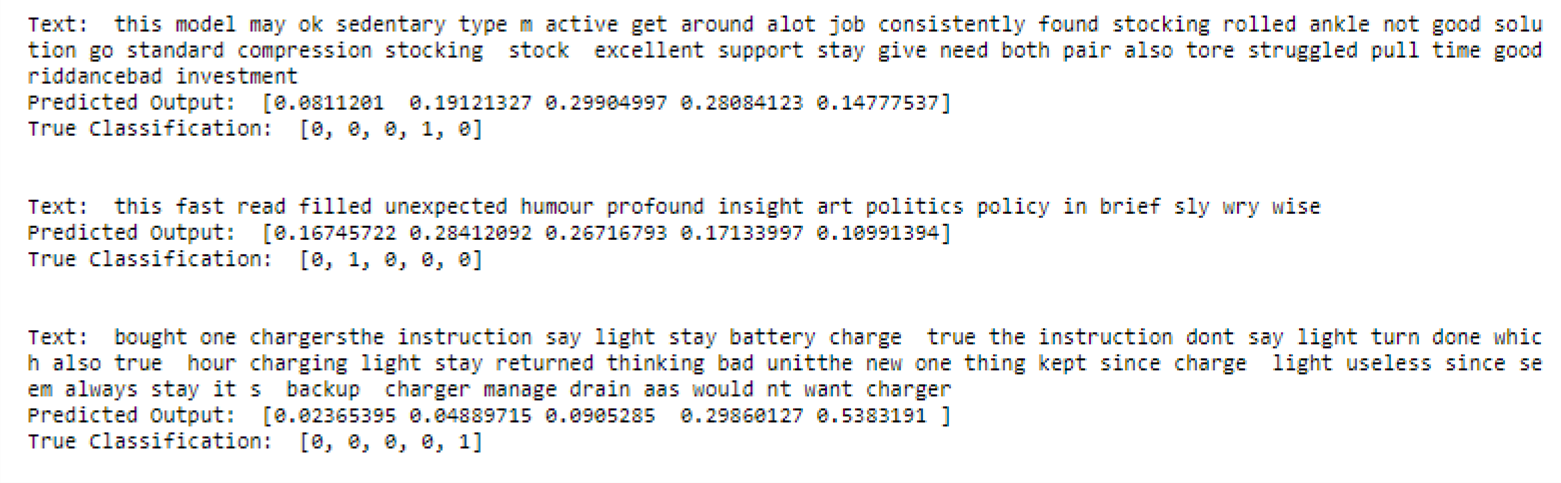
Truth \ Predicted	1	2	3	4	5
1	54%	27%	13%	4%	3%
2	24%	32%	29%	12%	4%
3	8%	19%	32%	29%	12%
4	2%	8%	21%	31%	36%
5	2%	7%	13%	16%	62%

MLP


Model Loss History: λ = 0.0001 , Dropout Rate = 0.1 , Learning Rate = 0.01
Training Loss History
Validation Loss History
Loss
1.60
1.55
1.50
1.45
1.40
1.35
Epoch
0
10
20
30
40
50


Model Accuracy History: λ = 0.0001 , Dropout Rate = 0.1 , Learning Rate = 0.01
Training Accuracy History
Model Validation Accuracy History
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Epoch
0
10
20
30
40
50

Results



N

```
Text:  this model may ok sedentary type m active get around alot job consistently found stocking rolled ankle not good solu
tion go standard compression stocking  stock  excellent support stay give need both pair also tore struggled pull time good
riddancebad investment
Predicted Output:  [0.0811201  0.19121327 0.29904997 0.28084123 0.14777537]
True Classification:  [0, 0, 0, 1, 0]
```

Y

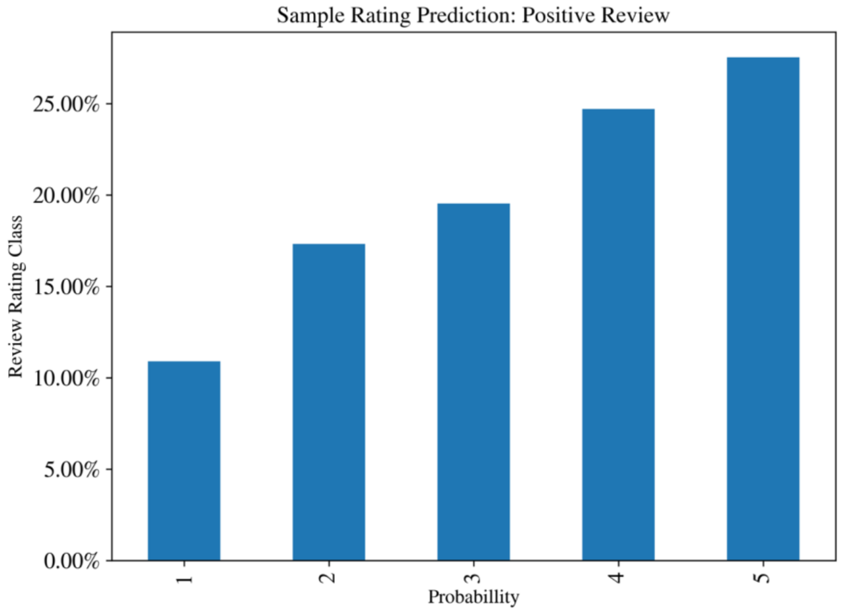
```
Text:  this fast read filled unexpected humour profound insight art politics policy in brief sly wry wise
Predicted Output:  [0.16745722 0.28412092 0.26716793 0.17133997 0.10991394]
True Classification:  [0, 1, 0, 0, 0]
```

Y

```
Text:  bought one chargersthe instruction say light stay battery charge  true the instruction dont say light turn done whic
h also true  hour charging light stay returned thinking bad unitthe new one thing kept since charge  light useless since se
em always stay it s  backup  charger manage drain aas would nt want charger
Predicted Output:  [0.02365395 0.04889715 0.0905285  0.29860127 0.5383191 ]
True Classification:  [0, 0, 0, 0, 1]
```

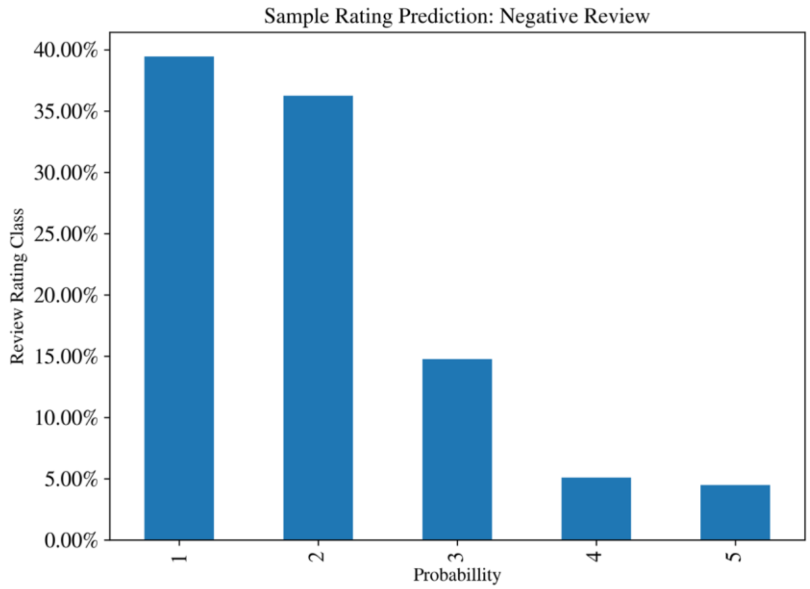
Results

Positive Review: 'I absolutely love this product, it is everything that I expected. The quality is unbeatable for the price. This product is amazing. I could not recommend this product more! ’



Results

Negative Review: ‘The package arrived damaged and the tv was fully broken. Do not buy this product, it is cheap and not made well. The tv is not very bright and makes a loud noise when you turn it on. The screen arrived broken and three days late.'



Future Considerations

- Use the Review Title Data Feature
- More Layers for a Deeper MLP
- Different Network Architecture (RNN, LSTM embedding layer)
- More time for training
    - More computing power
- Different Ensemble Techniques
- Bagging/Boosting