# Amazon Pets Product Reviews Sentiment Analysis

Amanda (Man Kuei) Chen

M.S. Business Analytics

DATS 6312_11: Natural Language Processing

Professor: Amir Jafari, PhD

# Contents

- Problem understanding
- Data understanding
- Data preparation
- Models
- Evaluation
- Conclusion

# Problem understanding

The customer service team spends too much time on reading each product's review, which might delay the messages that could be more urgent, leading to more complaints.

Therefore, to mitigate this problem, I built a sentiment classifier to help the customer service team prioritize their tasks by categorizing the sentiment of product reviews. This sentiment classifier can detect whether the reviews are positive or negative.

# Data understanding

| | overall | reviewText | year_month |
|---|---|---|---|
| 0 | 1.0 | I was not happy with product would like to ret... | 2016-12 |
| 1 | 1.0 | This cd is scratched and it constantly skips. ... | 2016-12 |
| 2 | 4.0 | It works just fine and repeats when I'm not he... | 2016-12 |

**Data: Amazon pets product reviews**

**Size: 1.2GB**

**Year: 2013-10 ~ 2018-10**

**Records: 5,816,874 rows and 3 columns**

**Source:**

**https://cseweb.ucsd.edu/~jmcauley/pdfs/emnlp19a.pdf**

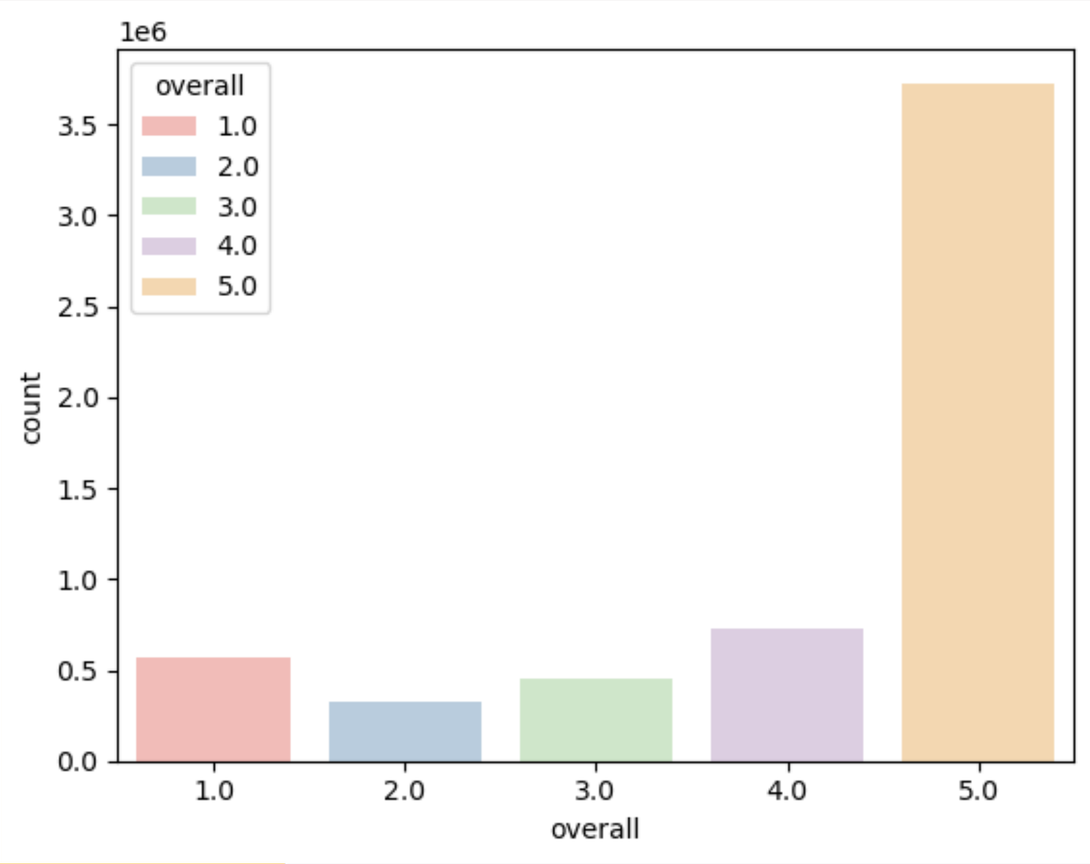
# Data preparation

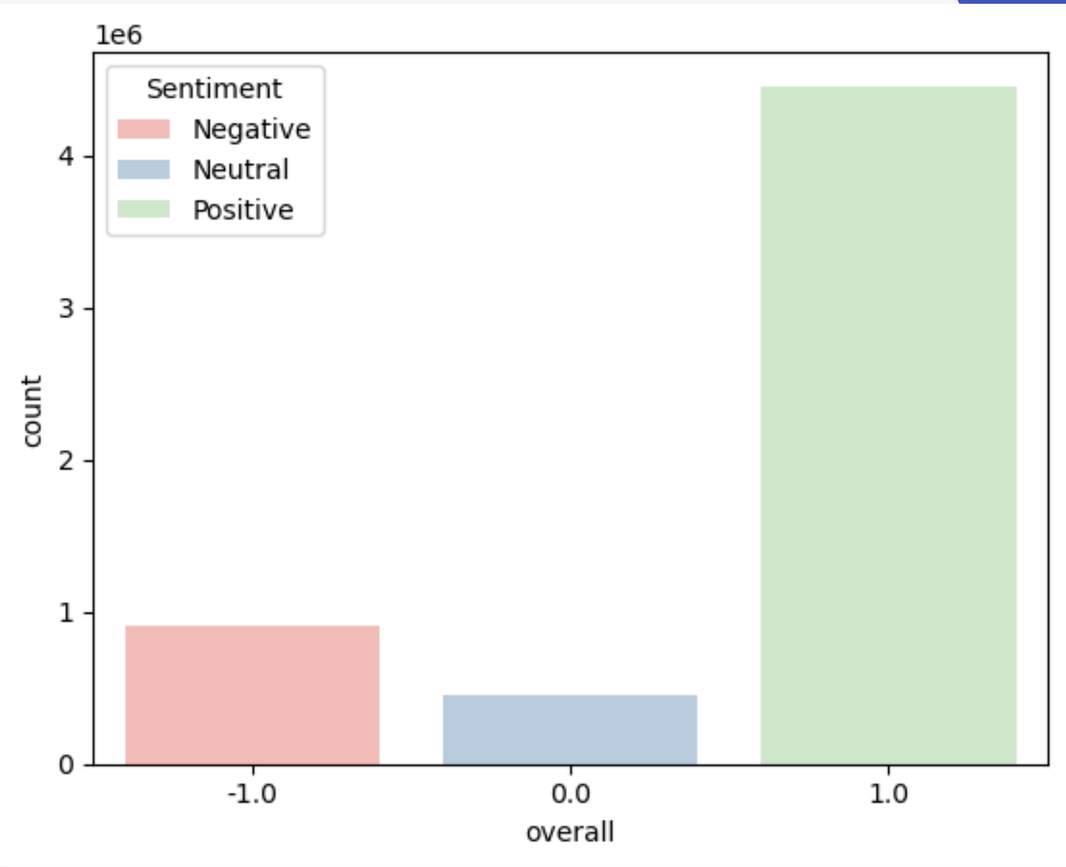
**This project encompasses two parts –**

**Part1. Exploratory Data Analysis, which helps us know more about this dataset and gain insights into what customers talk about in their reviews**

**Part2. Sentiment Analysis using a baseline model and a fine tuned pre-trained model.**

# Part 1 -Exploratory Data Analysis



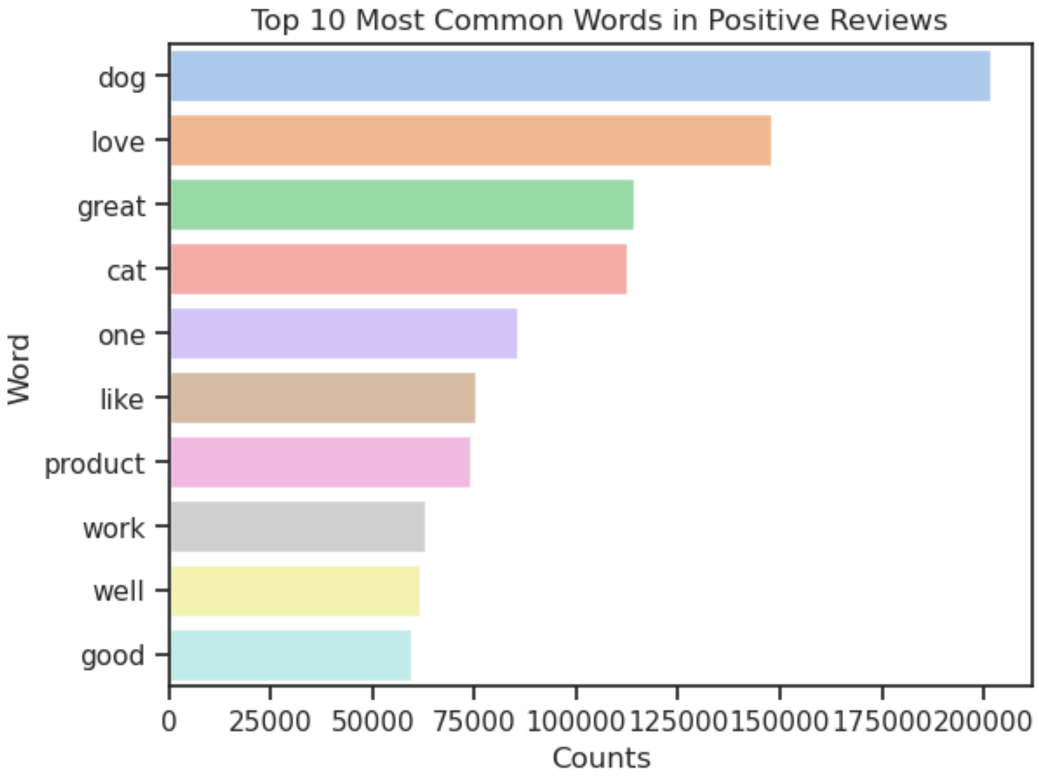



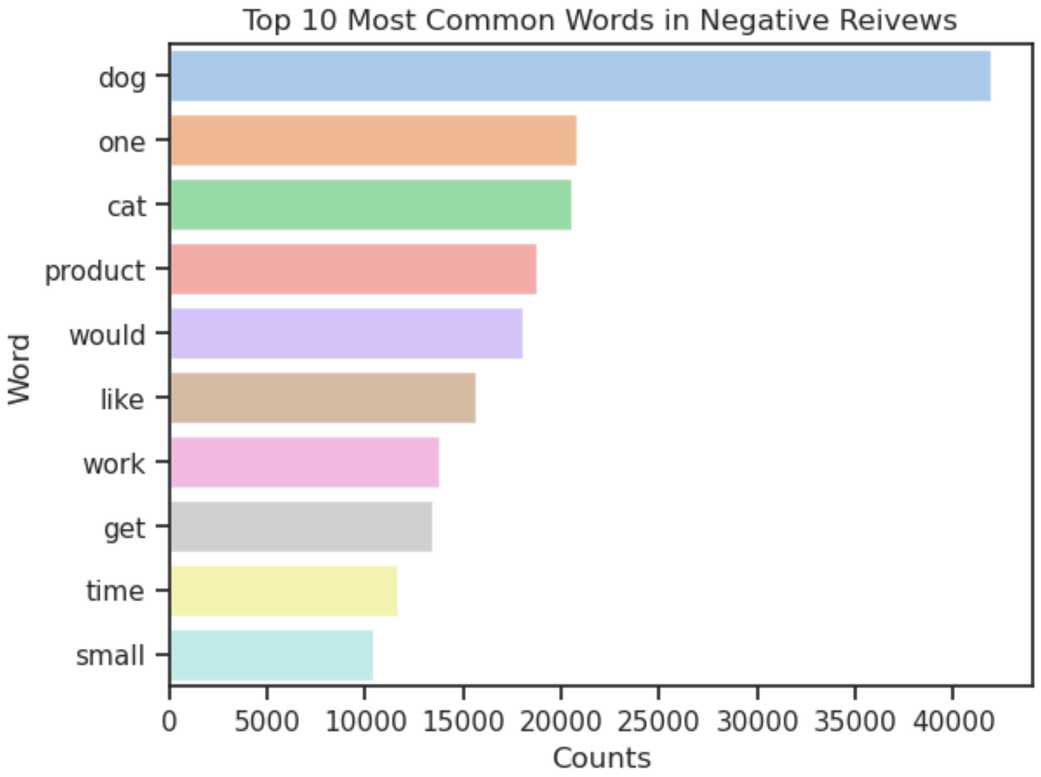
# Part 1 -Exploratory Data Analysis

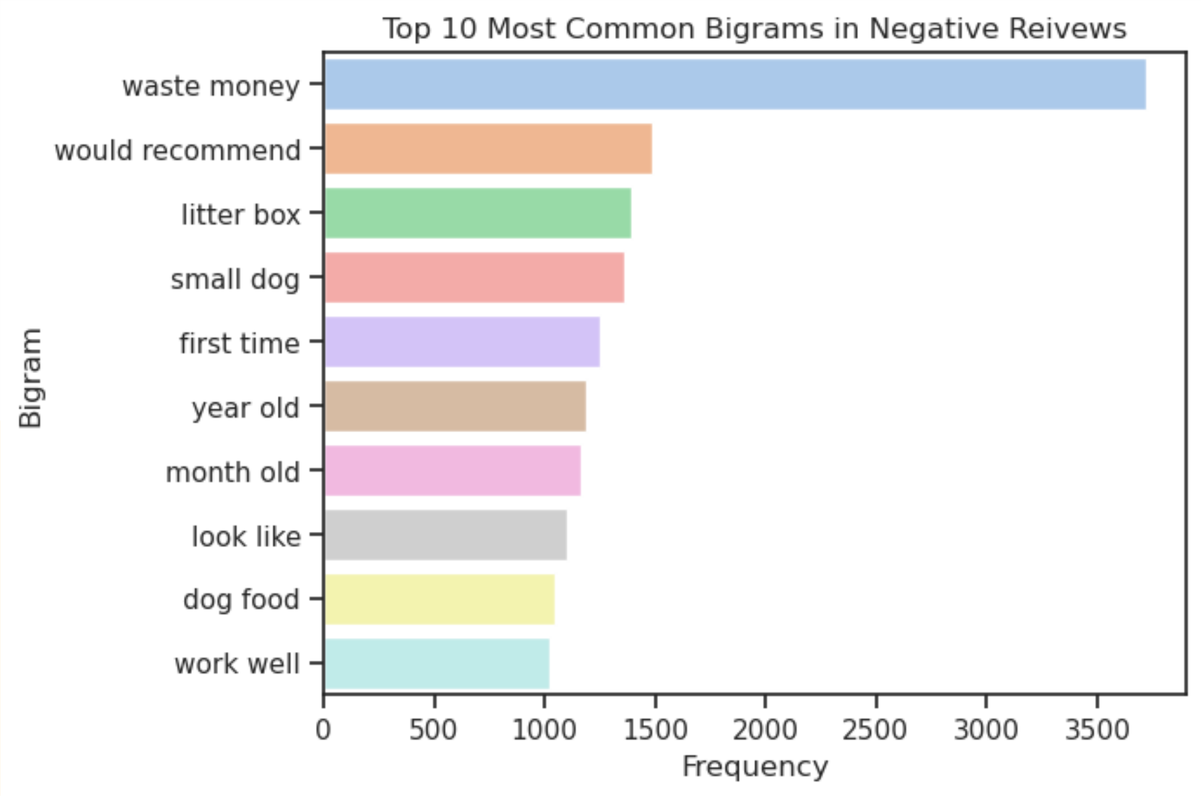
## Text preprocessing

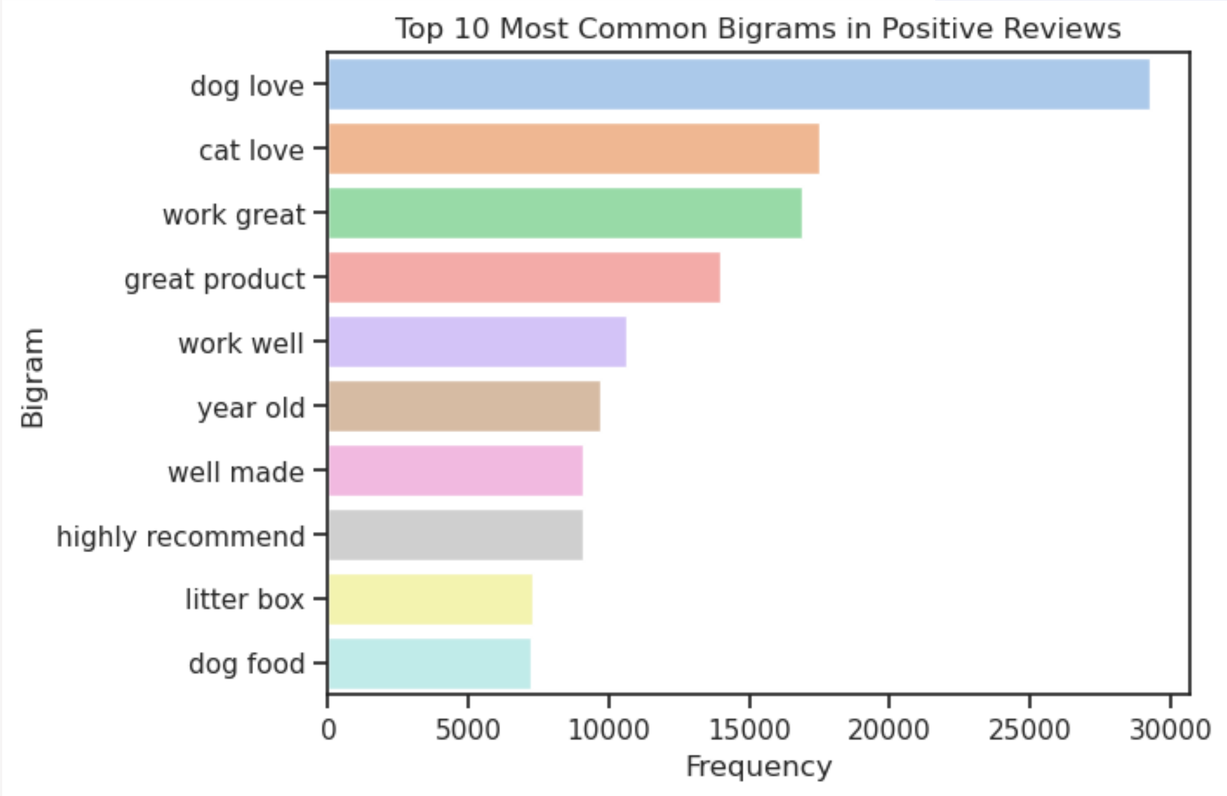
| | overall | reviewText | year_month | Text | Text1 |
|---|---|---|---|---|---|
| 0 | -1.0 | I was not happy with product would like to ret... | 2016-12 | i was not happy with product would like to ret... | happy product would like return work |
| 1 | -1.0 | This cd is scratched and it constantly skips. ... | 2016-12 | this cd is scratched and it constantly skips d... | cd scratched constantly skip disappointed |
| 2 | 1.0 | It works just fine and repeats when I'm not he... | 2016-12 | it works just fine and repeats when i m not he... | work fine repeat bird still talking stay quiet... |
| 3 | 0.0 | I purchased this cd for my Pocket Parrot. It h... | 2016-12 | i purchased this cd for my pocket parrot it ha... | purchased cd pocket parrot woman man speaking ... |
| 4 | 0.0 | Maybe it's just my Amazon parrot, but she's no... | 2016-10 | maybe it s just my amazon parrot but she s not... | maybe amazon parrot picking quickly well |
| 5 | 1.0 | I bought this to help me teach my blue quaker ... | 2016-09 | i bought this to help me teach my blue quaker ... | bought help teach blue quaker named booger tal... |
| 6 | 0.0 | Bird showed no interest in it.... | 2016-08 | bird showed no interest in it | bird showed interest |
| 7 | -1.0 | did not like at all | 2016-08 | did not like at all | like |
| 8 | -1.0 | Didn't do a thing for my African Grey! He hate... | 2016-08 | didn t do a thing for my african grey he hated... | thing african grey hated obnoxious |
| 9 | 1.0 | I like it - but the Cockatiel I am not so sure... | 2016-07 | i like it but the cockatiel i am not so sure h... | like cockatiel sure listens far speak |

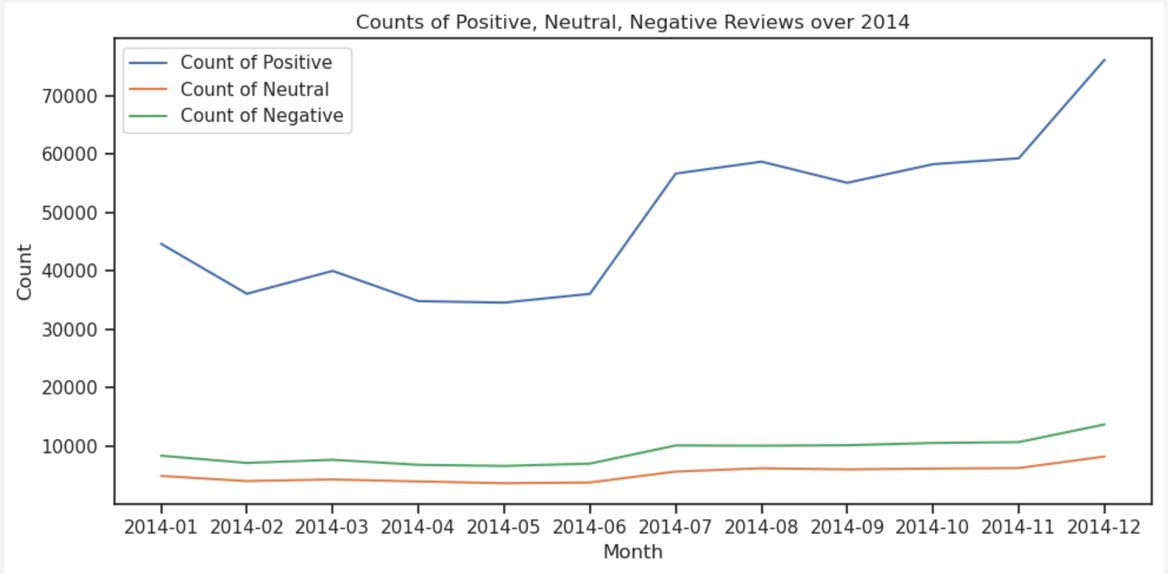
# Part 1 -Exploratory Data Analysis

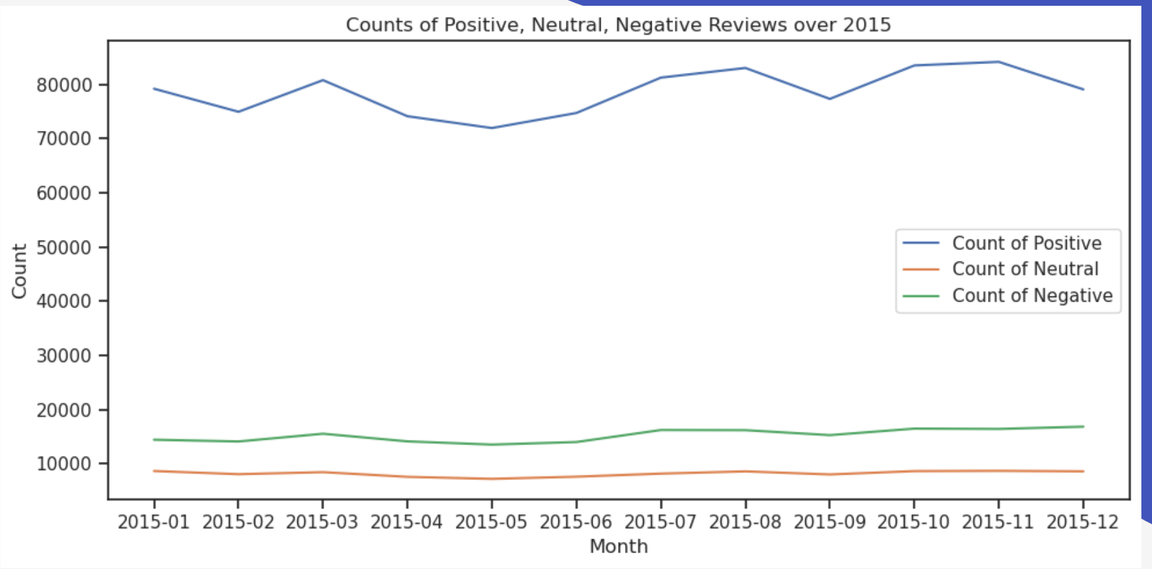


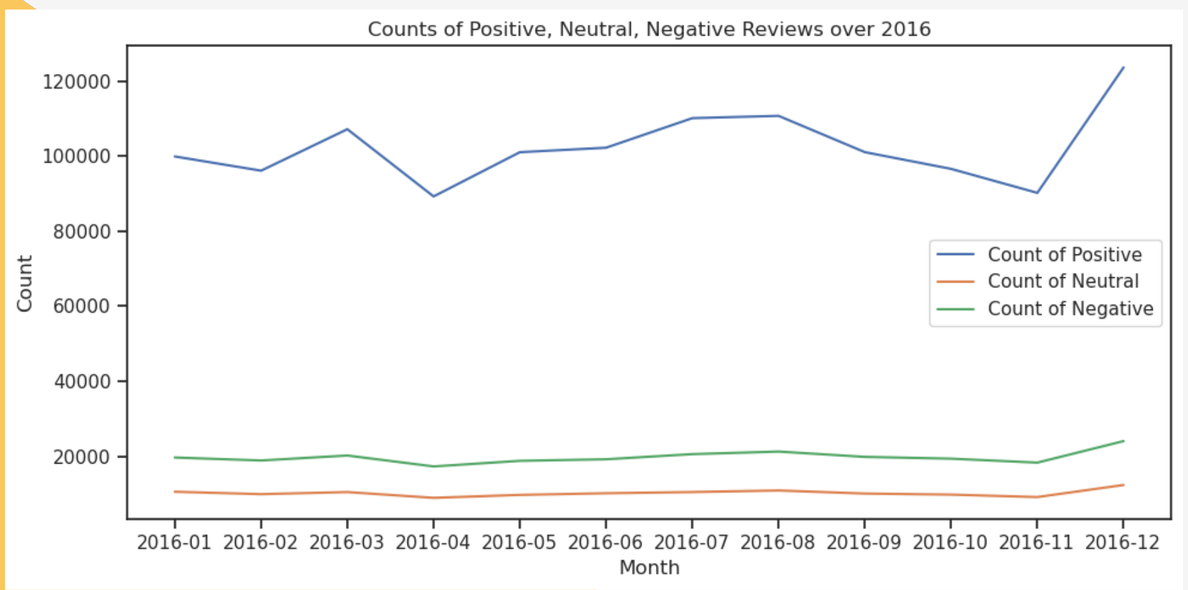



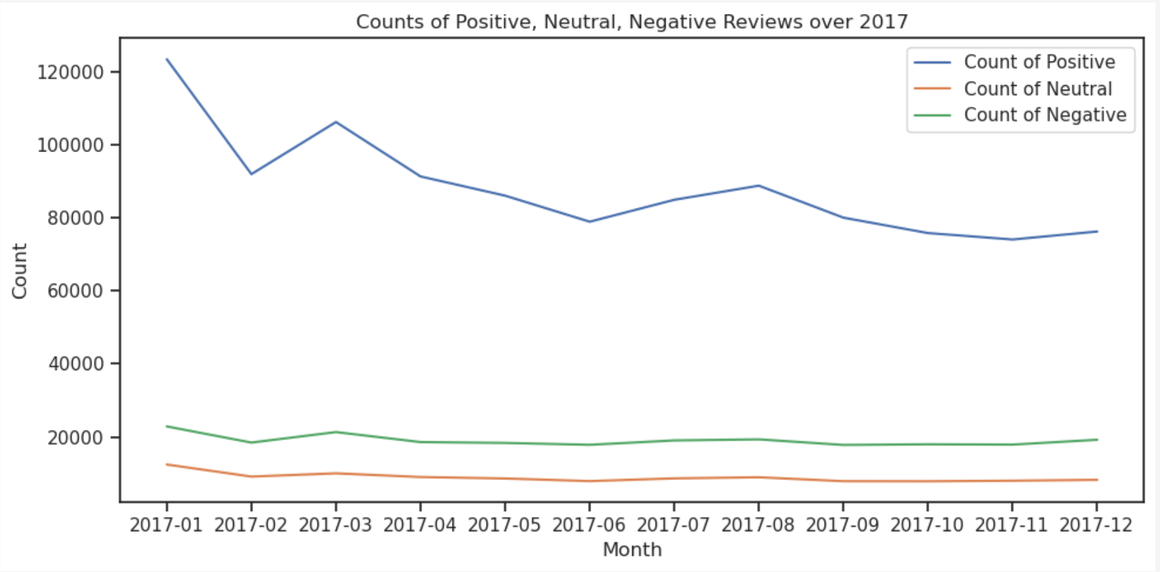

Top 10 Most Common Bigrams in Negative Reivews
waste money
would recommend
litter box
small dog
first time
year old
month old
look like
dog food
work well
Bigram
0
500
1000
1500
2000
2500
3000
3500
Frequency

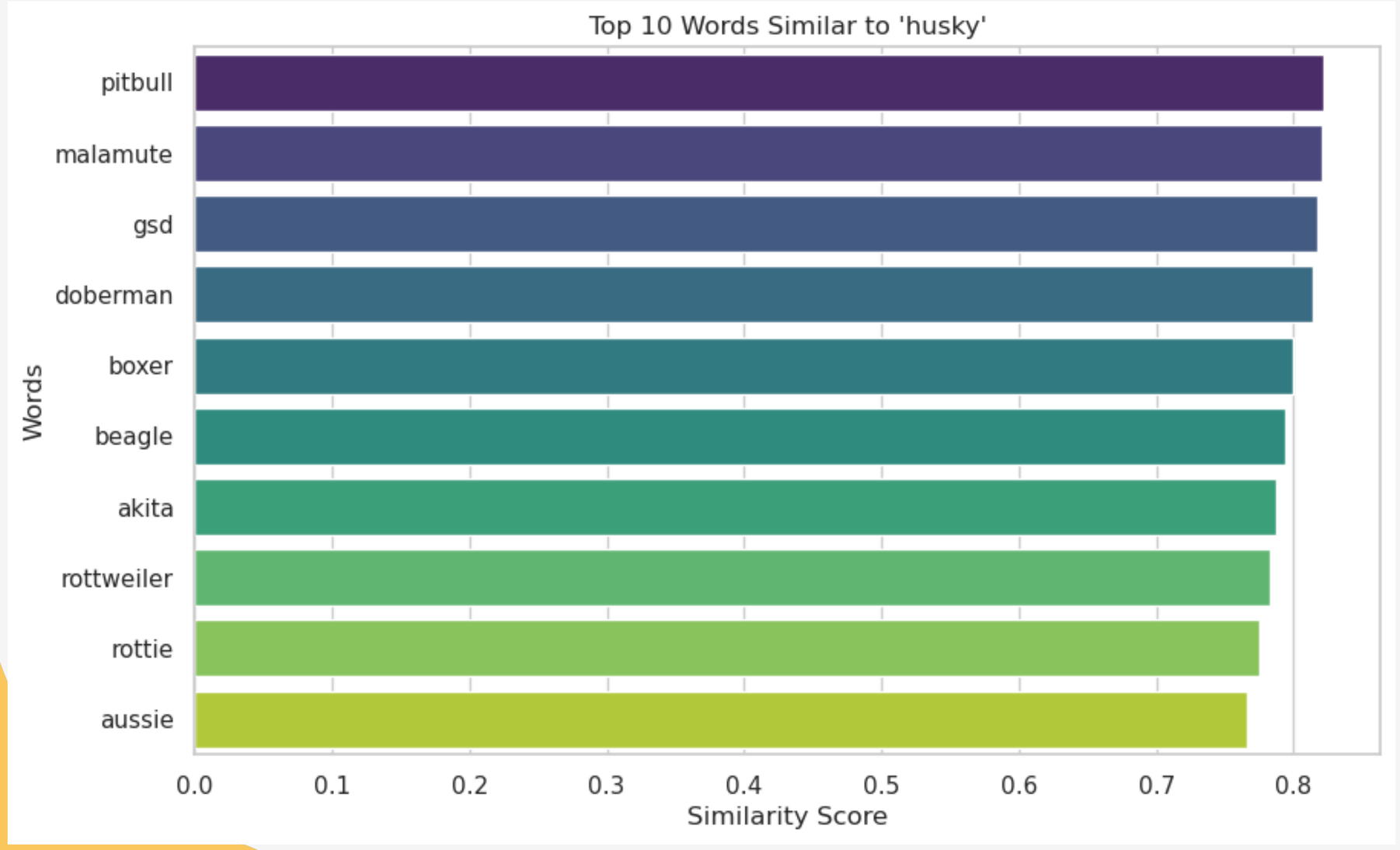

Top 10 Most Common Bigrams in Positive Reviews
dog love
cat love
work great
great product
work well
year old
well made
highly recommend
litter box
dog food
Bigram
0
5000
10000
15000
20000
25000
30000
Frequency


Counts of Positive, Neutral, Negative Reviews over 2014
Count of Positive
Count of Neutral
Count of Negative
Count
Month
2014-01 2014-02 2014-03 2014-04 2014-05 2014-06 2014-07 2014-08 2014-09 2014-10 2014-11 2014-12


Counts of Positive, Neutral, Negative Reviews over 2015
Count of Positive
Count of Neutral
Count of Negative
Count
Month
2015-01 2015-02 2015-03 2015-04 2015-05 2015-06 2015-07 2015-08 2015-09 2015-10 2015-11 2015-12


Counts of Positive, Neutral, Negative Reviews over 2016
Count of Positive
Count of Neutral
Count of Negative
Count
Month
2016-01 2016-02 2016-03 2016-04 2016-05 2016-06 2016-07 2016-08 2016-09 2016-10 2016-11 2016-12


Counts of Positive, Neutral, Negative Reviews over 2017
Count of Positive
Count of Neutral
Count of Negative
Count
Month
2017-01 2017-02 2017-03 2017-04 2017-05 2017-06 2017-07 2017-08 2017-09 2017-10 2017-11 2017-12


Top 10 Words Similar to 'husky'
pitbull
malamute
gsd
doberman
boxer
beagle
akita
rottweiler
rottie
aussie
Words
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Similarity Score

# Part 2. Sentiment Analysis

## Models

| Logistic Regression Model | DistlBERT |
| --- | --- |
| 1. The output can be interpreted as probabilities, making it easier to understand the impact of each variable on the outcome.<br>2. Lower risk of overfitting compared to more complex models. | DistilBERT is a small, fast, cheap and light Transformer model trained by distilling BERT base. It has 40% less parameters than bert-base-uncased, runs 60% faster while preserving over 95% of BERT's performances as measured on the GLUE language understanding benchmark. |

# Models - Baseline

## Logistic regression model

| | F1-Score | Accuracy |
|---|---|---|
| 0 | 0.71 | |
| 1 | 0.95 | |
| | | 0.92 |

Table1. Imbalanced classes

| | F1-Score | Accuracy |
|---|---|---|
| 0 | 0.92 | |
| 1 | 0.92 | |
| | | 0.92 |

Table2. Balanced classes

# Models - DistilBERT

**After having a 0.92 accuracy as my benchmark, I fine tuned with my data with distilBERT, configuring the batch size 32, epoch 8, learning rate at 3e-5 , and optimizer Adam. I experimented from batch size 16 to batch size 32 and epoch from 2 to 8.**

```
Epoch 1/8
45133/45133 [==============================]- 1651s 36ms/step- loss: 0.2074- accuracy: 0.9146
Epoch 2/8
45133/45133 [==============================]- 1627s 36ms/step- loss: 0.1746- accuracy: 0.9303
Epoch 3/8
45133/45133 [==============================]- 1627s 36ms/step- loss: 0.1560- accuracy: 0.9390
Epoch 4/8
45133/45133 [==============================]- 1627s 36ms/step- loss: 0.1404- accuracy: 0.9463
Epoch 5/8
45133/45133 [==============================]- 1627s 36ms/step- loss: 0.1259- accuracy: 0.9528
Epoch 6/8
45133/45133 [==============================]- 1627s 36ms/step- loss: 0.1131- accuracy: 0.9581
Epoch 7/8
45133/45133 [==============================]- 1627s 36ms/step- loss: 0.1019- accuracy: 0.9630
Epoch 8/8
45133/45133 [==============================]- 1626s 36ms/step- loss: 0.0929- accuracy: 0.9666
```

```
11284/11284 [==============================]- 164s 14ms/step- loss: 0.0577- accuracy: 0.9819
loss: 0.05769692733883858
accuracy: 0.9819391965866089
```

# Evaluation

## Baseline

| | F1-Score | Accuracy |
|---|---|---|
| 0 | 0.92 | |
| 1 | 0.92 | |
| | | 0.92 |

Table2. Balanced classes

# DistilBERT

11284/11284 [==============================]- 164s 14ms/step- loss: 0.0577- accuracy: 0.9819
loss: 0.05769692733883858
accuracy: 0.9819391965866089

# Conclusion

**Even though the logistic model is a baseline model, the accuracy score is already good by industrial standards.**

**However, the fine-tuned distilBERT has so much better performance despite of its disadvantages in terms of time consumption and high cost. If the cost of training is the priority concern, then the business can choose a logistic model. Otherwise, fine-tuning a pre-trained model gives a better model.**