

Documentation

OF

Analysis of Amazon Cell Phone Reviews Using Natural Language Processing

Done By:
Aisha Mustafa Ali Hatem
Debabratta Mishra

1-Introduction:

1.1 Overview:

Predict sentiment from Amazon reviews using Natural Language Processing

1.2 Purpose:

To analyse thousands of reviews of unlocked mobile phones sold on Amazon.com to find insights with respect to reviews, ratings, price and their relationships.

2.Literature Survey:

2.1 Existing problem:

Mobile phones have revolutionized the way we purchase products online, making all the information available at our fingertips. As the access to information becomes easier, more and more consumers will seek product information from other consumers apart from the information provided by the seller. Reviews and ratings submitted by consumers are examples of such of type of information and they have already become an integral part of customer's buying-decision process. The review and ratings platform provided by eCommerce players creates transparent system for consumers to take informed decision and feel confident about it.

2.2 Proposed solution:

This project aims at building a model to predict the helpfulness of the review and the rating based on the review text. Corpus-based and knowledge-based methods can be used to determine the semantic similarity of review text. We will be using Natural language processing to analyse the sentiment (positive or a negative) of the given review . A sample web application is integrated to the model built.

3. Theoretical Analysis:

Theoretical Analysis

3.1 Block diagram:

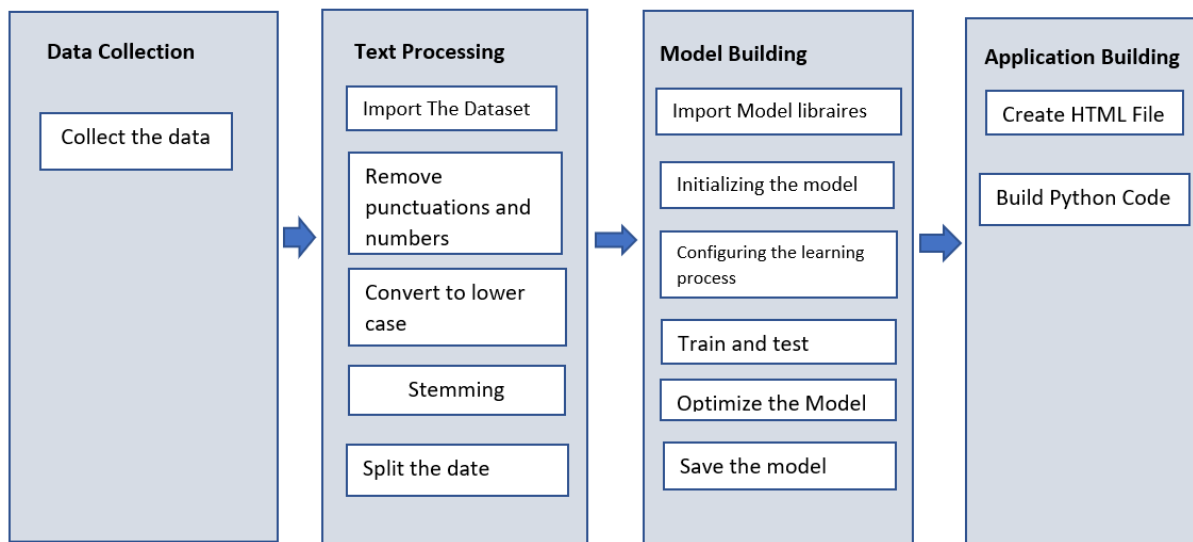


Figure 1. Block diagram of tasks performed

The Block diagram shows the tasks that were performed on the data set

tical Analysis

3.2 Hardware / Software designing:

3.2.1 Software Designing:

- i. Operating system:
.MacOS, Windows or Linux
- ii. Anaconda
- iii. Python 3.7 or below
- iv. Natural Language Toolkit
- v. Python Libraries:
.Pandas
.Numpy
.Keras
.Sklearn
- vi. Jupyter Notebook
- vii. FlaskML
- viii. HTML
- ix. CSS

3.2.2 Hardware Designing:

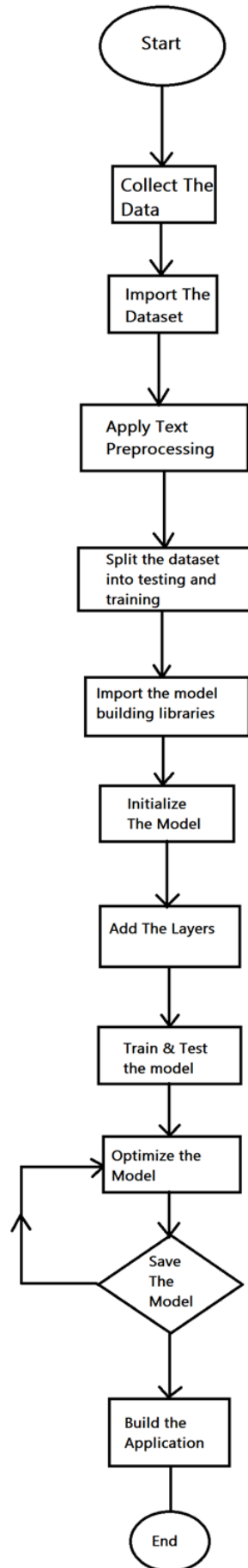
- System architecture: Windows- 64-bit x86, 32-bit x86; MacOS- 64-bit x86; Linux- 64-bit x86

- minimum of 8GB RAM
- minimum of 7th generation (Intel Core i7 processor)
- GPU

4. Experimental Investigations:

- A. The data has to be cleansed properly and spaces have to be removed for the model to work perfectly.
- B. The dependant variable (ratings) has to be converted into a binary format.

5.Flowchart:



6.

6.Result:

The Model of analyzing Amazon Cell Phone Reviews is Working with an accuracy percentage of 94.26%

7.Advantages & Disadvantages:

7.1 Advantages:

- Users can provide any review and get a direct response within seconds.
- NLP system provides answers to the questions in natural language
- NLP system offers exact answers to the questions, no unnecessary or unwanted information
- The accuracy of the answers increases with the amount of relevant information provided in the question.
- Allows you to perform more language-based data compares to a human being without fatigue and in an unbiased and consistent way.
- Structuring a highly unstructured data source

7.2 Disadvantages:

- Complex Query Language- the system may not be able to provide the correct answer it the Review that is poorly worded or ambiguous.
- The system is built for a single and specific task only; it is unable to adapt to new domains and problems because of limited functions.
- NLP system doesn't have a user interface which lacks features that allow users to further interact with the system

8. Applications:

Natural Language Understanding is particularly difficult for machines when it comes to opinions, given that humans often use sarcasm and irony. Sentiment analysis, however, is able to recognize subtle nuances in emotions and opinions – and determine how positive or negative they are.

When you analyze sentiment in real-time, you can monitor mentions on social media (and handle negative comments before they escalate), gauge customer reactions to your

latest marketing campaign or product launch, and get an overall sense of how customers feel about your company.

You can also perform sentiment analysis periodically, and understand what customers like and dislike about *specific aspects* of your business – maybe they love your new feature, but are disappointed about your customer service. Those insights can help you make smarter decisions, as they show you exactly what things to improve.

9.Conculsion:

In this project we tried to show the basic way of classifying reviews into positive or negative category using Natural Langugae Tool Kit.

10. Future Scope:

We could further improve our Analyzer by trying to extract more features from the Amazon Reviews, or trying another classifier all together.

11.Bibliography:

1-<https://www.guru99.com/nlp-tutorial.html>

2-<https://monkeylearn.com/blog/natural-language-processing-applications/>

12.Appendix:

12.1 Source Code:

Python Code:

```
1 #importing load_model for loading h5 file
2 from keras.models import load_model
```

```

3 #importing numpy
4 import numpy as np
5 from flask import Flask, request, jsonify, render_template
6 #importing pickle for load bag of word model
7 import pickle
8 #loading phone.h5 file using load_model
9 model=load_model('phone.h5')
10
11 app = Flask(__name__)
12
13 with open('count_vec.pkl','rb') as file:
14     cv=pickle.load(file)
15
16 @app.route('/')
17 def home():
18     return render_template('index.html')
19
20 @app.route('/y_predict',methods=["GET","POST"])
21 def y_predict():
22     '''
23     For rendering results on HTML GUI
24     '''
25     #Giving review
26     if request.method == 'POST':
27         inp = request.form.get("Review")
28
29         x=cv.transform([inp])
30         y=model.predict(x)
31         if(y>0.5):
32             y='Positive review'
33         else:
34             y='Negative review'
35
36         return render_template('index.html',prediction_text = 'Analysis
of Amazon Cell Phone Reviews '+y)
37
38
39 if __name__ == "__main__":
40     app.run(debug=True)

```

a

HTML Code:

```

1 <!DOCTYPE html>
2 <html >
3 <!--From https://codepen.io/frytyler/pen/EGdtg-->
4 <head>
5     <meta charset="UTF-8">

```



```

6   <title>ML API</title>
7   <link href='https://fonts.googleapis.com/css?family=Pacifico'
    rel='stylesheet' type='text/css'>
8   <link href='https://fonts.googleapis.com/css?family=Arimo'
    rel='stylesheet' type='text/css'>
9   <link href='https://fonts.googleapis.com/css?family=Hind:300'
    rel='stylesheet' type='text/css'>
10  <link
    href='https://fonts.googleapis.com/css?family=Open+Sans+Condensed:30
    0' rel='stylesheet' type='text/css'>
11  <link rel="stylesheet" href="{{ url_for('static',
    filename='css/style.css') }}">
12
13  <style>
14  .login{
15  top: 20%;
16  }
17  </style>
18  </head>
19
20  <body>
21  <div class="login">
22  <h1>Analysis of Amazon Cell Phone Reviews</h1>
23
24      <!-- Main Input For Receiving Query to our ML -->
25      <form action="{{ url_for('y_predict') }}"method="post">
26  <input type="text" name="Review" placeholder="Review"
    required="required" />
27
28      <button type="submit" class="btn btn-primary btn-block btn-
    large">Predict</button>
29
30  </form>
31
32  <br>
33  <br>
34  {{ prediction_text }}
35
36  </div>
37
38
39 </body>
40 </html>
41

```

12.2 Screenshots:

Analysis of Amazon Cell Phone Reviews

This is a good phone

Predict

Analysis of Amazon Cell Phone Reviews
Positive review

Analysis of Amazon Cell Phone Reviews

horrible

Predict

Analysis of Amazon Cell Phone Reviews
Negative review