Analysis of Amazon Cell Phone Reviews

1. Introduction:

1.1 Overview:

With the advent of Industry 4.0 various sectors of the industry and companies have started to use Deep learning in order to make things easier and increase efficiency and, at the end of the day, increase the profit margin. Analyzing and monitoring is one area in which AI can be easily be used in order to have remote monitoring and control.

1.2 Purpose:

The aim of this project is predicting the helpfulness of an online product review serves both the customer and online retailer. The customer is provided with the most helpful and relevant reviews about a product of interest without having to mine for good reviews.

2. Literature Survey:

2.1 Existing Problem:

90 percent of the consumers read online reviews before they decide to purchase

any Mobile phone from any e-commerce website .Online Mobile applications has revolutionised the way consumers purchase mobile phones online as these apps have all the information regarding any mobile phone at users finger tips. Amazon is one of the best mobile applications which is considered as a treasure trove of all mobile reviews, and their review system is accessible across all channels presenting reviews in an easy-to-use format. So,There should be a system which analyses thousands of reviews of unlocked mobile phones sold on Amazon.com to find insights with respect to reviews, ratings, price and their relationships

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 of the given review .A sample web application is integrated to the model built.

3. Theoretical Analysis:
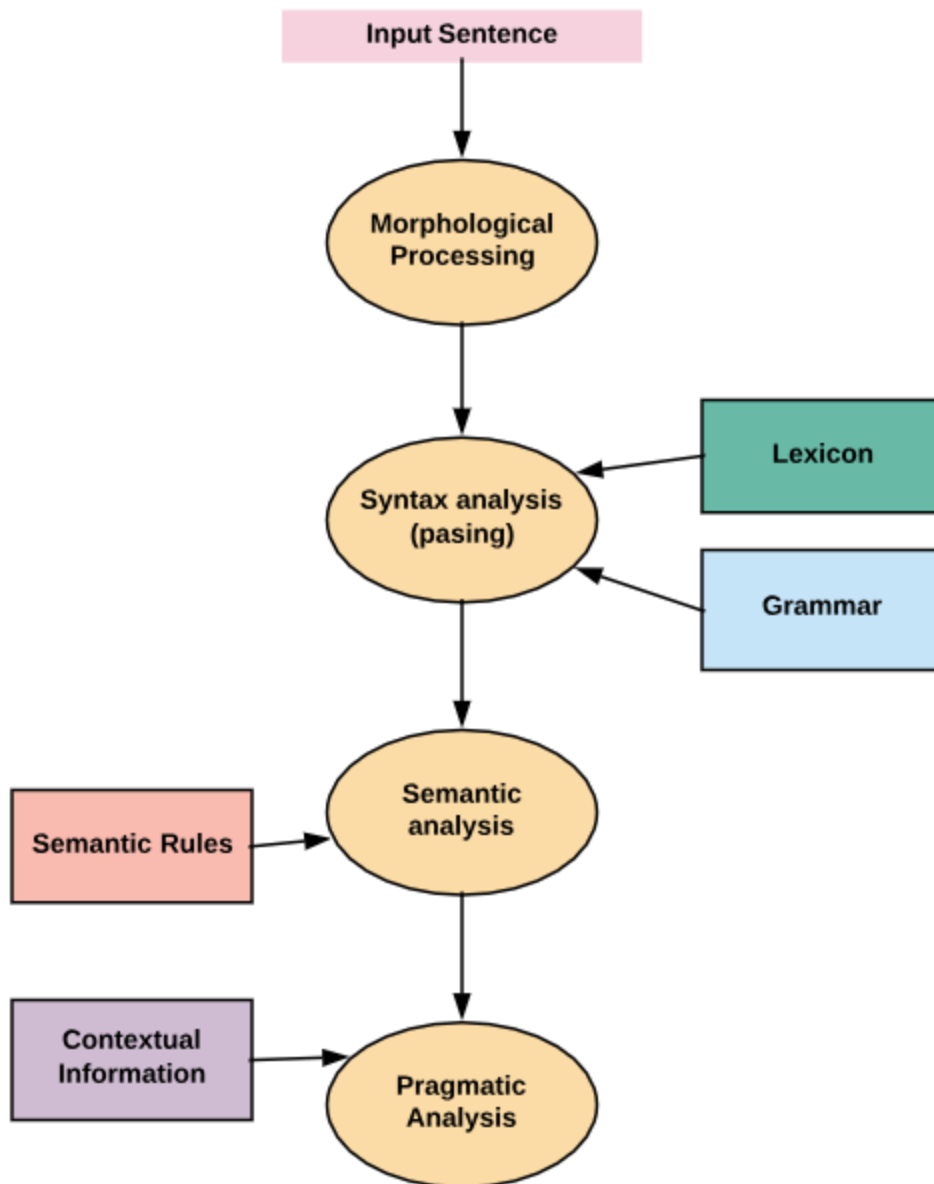
3.1 Block Diagram:
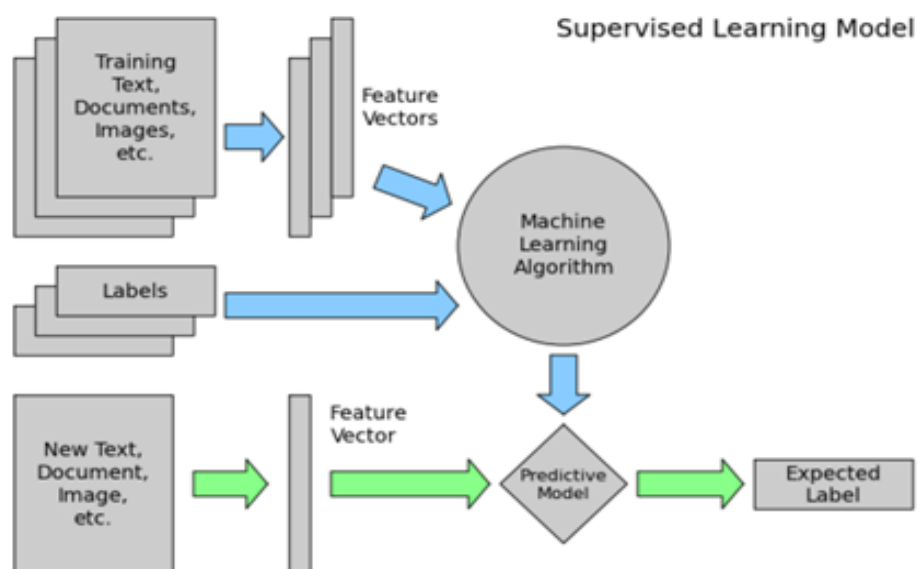
3.2 Hardware / Software Designing:

In this project I used Natural Language Processing for the sentimental analysis for the review with the anaconda service(Spyder as well as jupyter notebook). With help of NLP , we are able to recognize the reviews. After that I used a Artificial Neural network for this project, then will help of html we getting the input from the customers and with help of the flak web app framework we find the helpfulness of the review. If we want we can use the Nodered and the MIT services also.

4. Experimental Investigations:

The working depends on the accuracy of the visual recognition service and in turn on the training data. Also the system has to be monitored frequently at least in the initial stages of implementation to make sure there are no errors.NLP lie in a number of disciplines, viz. computer and information sciences, linguistics, mathematics, electrical and electronic engineering, artificial intelligence and robotics, psychology, etc. Applications of NLP include a number of fields of studies, such as machine translation, natural language text processing and summarization, user interfaces, multilingual and cross language information retrieval (CLIR), speech recognition, artificial intelligence and expert systems, and so on

5. Flowchart:

6. Result:

Jupyter NLP FINAL-PROJECT (autosaved)   Logout

File   Edit   View   Insert   Cell   Kernel   Widgets   Help   Not Connected Trusted | Python 3

```
                sentiment    False
                dtype: bool
```

In [12]:
```
import re #regular expressions to replace special characters
import nltk
nltk.download("stopwords") # for is then that or , is here where
from nltk.corpus import stopwords
from nltk.stem.porter import PorterStemmer

[nltk_data] Downloading package stopwords to C:\Users\Anantha
[nltk_data]     Priya\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
```

In [13]:
```
ps = PorterStemmer()
data = []
```

In [14]:
```
for i in range(0,67906):
    review = ds["title"][i]
    review = re.sub("[^a-zA-Z]", ' ',review)
    review = review.lower()
    review = review.split()
    review = [ps.stem(word) for word in review if not word in set(stopwords.words('english'))]
    review = ' '.join(review)
    data.append(review)
```

In [15]:
```
from sklearn.feature_extraction.text import CountVectorizer
import pickle
cv = CountVectorizer(max_features = 3000)
x = cv.fit_transform(data).toarray()
with open('Countvectorizer','wb') as file:
    pickle.dump(cv,file)
```

In [16]: `y = ds.iloc[:,3:4].values`

In [17]: `y`

Out[17]:
```
array([[0],
       [0],
```

Type here to search   ∧ ⚹ ◧ ◁» ENG 14:27 08-07-2020

Jupyter NLP FINAL-PROJECT (autosaved)   Logout

File   Edit   View   Insert   Cell   Kernel   Widgets   Help   Not Connected Trusted | Python 3

```
Epoch 14/20
54388/54388 [==============================] - 893s 16ms/step - loss: 0.1056 - accuracy: 0.9566
Epoch 15/20
54388/54388 [==============================] - 961s 18ms/step - loss: 0.1056 - accuracy: 0.9565
Epoch 16/20
54388/54388 [==============================] - 895s 16ms/step - loss: 0.1075 - accuracy: 0.9558
Epoch 17/20
54388/54388 [==============================] - 890s 16ms/step - loss: 0.1062 - accuracy: 0.9563
Epoch 18/20
54388/54388 [==============================] - 892s 16ms/step - loss: 0.1054 - accuracy: 0.9569
Epoch 19/20
54388/54388 [==============================] - 892s 16ms/step - loss: 0.1023 - accuracy: 0.9575
Epoch 20/20
54388/54388 [==============================] - 893s 16ms/step - loss: 0.1021 - accuracy: 0.9572
```

Out[35]: `<keras.callbacks.callbacks.History at 0x267f1d04d88>`

In [36]: `model.save("nlp.h5")`

In [24]:
```
text = "hello, I have this phone and used it until I decided to buy a flip phone. I have had NO problems with the battery or ne
text = re.sub("[^a-zA-Z]", ' ',text)
text = text.lower()
text = text.split()
text = [ps.stem(word) for word in text if not word in set(stopwords.words('english'))]
text = ' '.join(text)
```

In [60]: `y_ps = model.predict(cv.transform([text]))`

In [61]: `y_ps`

Out[61]: `array([[0.00312615]], dtype=float32)`

In [62]: `y_ps>0.50.9572`

Out[62]: `array([[False]])`

In [63]: `y_ps<0.5`

Type here to search   ∧ ⚹ ◧ ◁» ENG 14:27 08-07-2020

Welcome to Rating from Reviews

Write a Review to see if its a positive one 🙂 or a negative one 🙁!

Enter your review:

Show Rating

I like this phone.
This is a positive review.

7. Advantages & Disadvantages:

Advantages:

- Reduces the waste of human resource for such busy work. The personnel can be assigned some other task.
- There is virtually no chance for error once the system is well trained.
- Good for "expert" users- can quickly access commands .
- With [natural language processing](), computers can infer and analyze human language with more meaning than simply digesting and outputting programmed responses. In this sense, NLP offers greater use and interaction with humans.

Disadvantages:

- There is obvious loss of jobs.

- Commands have to be typed precisely. If there is a spelling error the command will fail.
- Difficult to recognise all the different ways of saying things.
- Difficulty in dealing with homonyms.
- The big disadvantage of natural language processing is that you have to have a separate version of the product for each language and, where appropriate, for each dialect of a language.

8. Applications:

Sentiment Analysis – It is widely used in social media analytics and web monitoring which allow knowing the insights of the customers concerning particular products or services. It can be advantageous for any company to know about the thinking of the customers about a product so that they can know about the scope of improvement and how to achieve robustness. Natural language processing can not solely handle this task; it requires integration with highly computational methods such as Machine learning and deep learning to do the back end computation and Big data analytics to digest the data at an enormous scale.

9. Conclusion:

Thus the project can be used in amazon for analysing the review for the all types of products without the need of any human help.

10. Future Scope:

Exponential growth of the digital data is one of the major factors that is driving the growth of the cloud natural language processing market. Data/information has emerged as one of the most important assets of the organizations. Companies are collecting, analyzing, and reporting vast volume of data for extracting meaningful insights to get competitive edge. Growing adoption of the big data technologies such as Hadoop and IoT among organizations along with popularity of cloud deployment is one of the major factors that are fueling the growth of the digital data. Digital data is estimated to grow

at a rate of 40% for next 10 years and by the year 2020 data generation is expected to reach 44 zettabytes per year. As the data increases, it leads to the requirement of an effective analytics solution to process the information, therefore, driving the cloud NLP market growth

11. Bibliography:

- Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*,

- Ion Androutsopoulos, Graeme Ritchie, and Peter Thanisch. Natural language interfaces to databases — an introduction. Journal of Natural Language Engineering.

- K. Bretonnel Cohen and Lawrence Hunter. Natural language processing and systems biology. In Werner Dubitzky and Francisco Azuaje, editors, *Artificial Intelligence Methods and Tools for Systems Biology*, page ??? Springer Verlag.

- Mary Dalrymple, V. Gupta, John Lamping, and V. Saraswat. Relating resource-based semantics to categorial semantics. In Mary Dalrymple, editor, *Semantics and syntax in Lexical Functional Grammar: the resource logic approach*, pages 261–280. MIT Press, Cambridge, MA, 1999.