Sentimental Analysis using Sound Processing

Vishvam jani *Nirmit Gupta* *Shreya Patel* *Meet Chaudhary*

202018011@daiict.ac.in *202018019@daiict.ac.in* *202018028@daiict.ac.in* *202018044@daiict.ac.in*

Parth Lakkad *Harman Singh* *Namra Jain*

202018047@daiict.ac.in *202018050@daiict.ac.in* *202018059@daiict.ac.in*

Department - MSc. Data Science

College - DAIICT, Gandhinagar, Gujarat.

Abstract **— The purpose of Our Project is to Transcribe Text from video source to analyze the events and sentiments to analyze the subjective information in the text and then mine the opinion . This issue is not limited to translating a source video into a target language video since the objective is to provide only the main idea to transcribe text from live source and analyze the sentiment of financial news which can further be used to analyze the impact of these financial news sentiments on the stock market price movement. This study proposes the use of a gensim model and machine learning algorithm-based classifier to perform financial news sentiment analysis.The training data used is labelled using sentiment dictionary from kaggle. To determine which machine learning algorithm-based classifier shows better performance, some of machine learning algorithm-based classifiers were used and tested separately In this Project we present possible techniques for objective, yet we focus on Future works.**

Keywords - Transcribe text, Sentiment analysis, NLTK , tokenization , classification, Doc2vec, Gensim

1)INTRODUCTION

The indian and global stock market plays a significant role in the economy of the country and the world respectively.

It performs a crucial role in the growth of business and trade and affects the financial state of the country.It can greatly influence the economy of the nation.The main source of watching this news is through news channels which gives us every information about all of the stocks. There are also some designated news channels which are only designed to spread financial news.These news channels also stream online and viewers can aso watch the livestream of these news channels online.Viewers business leaders highly depend on these channels to gain useful information and make their own investment decisions which may be fruitful to them.

Sentiment analysis is the path toward distinguishing positive or negative slant in text. It's habitually used by associations to recognize supposition in cordial data, check brand reputation, and understand customers.Since customers express their contemplations and feelings more straightforwardly than some other time, slant investigation is transforming into a key contraption to screen and grasp that assessment. Thus inspecting customer contribution, for instance, sentiments in examination responses and online media conversations, grants brands to acknowledge what makes customers happy or confused, with the objective that they can tailor things and organizations to address their customers' issues. It is basic in light of the fact that it helps associations with understanding the overall evaluations of their customers. Through subsequently organizing the supposition behind overviews, electronic media conversations, to say the least, you can make faster and more accurate decisions. It's surveyed that 90% of the world's data is unstructured, with everything taken into account it's messy. Enormous volumes of unstructured business data are made every day: messages, support tickets, talks, online media conversations, examines, articles, reports, etc) For any situation, it's hard to separate for estimation in an ideal and beneficial manner

1.1)PROJECT OVERVIEW

Wide range of applications in business and public policy uses sentiment analysis. Sentiment analysis is the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information

Sentimental analysis is being used here to predict the sentiments of the news inferred.Instead of watching and keeping a check on the news 24x7, what we can do is that we can make a model that predicts the sentiment of the news on it's own and categorizes it into three sentiments that is positive negative and neutral.

Sentiment analysis with sound preparing can end up being helpful.Instead of going through each feature for each stock you are keen on, we can utilize Python to parse this site information and perform opinion investigation (for example allot an opinion score) for each feature prior to averaging it over a time of time.Sentiment analysis is an AI device that investigates text for extremity, from positive to negative to neutral. Via preparing AI devices with instances of feelings in text, machines consequently figure out how to distinguish notion without human info. Here We play out the errand of deciding if a piece of composing is positive, negative or neutral

We are here trying to to make a machine learning model that first transcribes the text from live video source(Bloomberg live news channel) and then applies it's machine learning properties to judge the sentiment of the news

1.2)PROBLEM STATEMENT

Businesses and organizations have always been concerned about how they are perceived by the public. This concern results from a variety of motivations, including marketing and public relations. Before the era of the Internet, the only way for an organization to track its reputation in the media was to hire someone for the specific task of reading newspapers and manually compiling lists of positive, negative and neutral references to the organization, it could undertake expensive surveys of uncertain validity. New data is brought into the market constantly. While an assortment of data sources would all be able to move a stock cost, e.g., bits of hearsay, listening in and embarrassments; monetary news articles are viewed as more steady and a more reliable source. Be that as it may, the specific connection between monetary news stories and stock value development is unpredictable. In any event, when the data contained in monetary news stories can noticeably affect a security's cost

Fine-grained sentiment analysis is an extremely challenging problem because of the variety of ways in which opinions can be expressed. News articles present an even greater challenge, as they usually avoid overt indicators of attitudes. However, despite their apparent neutrality, news articles can still bear polarity if they describe events that are objectively positive or negative.

The main problem that our project tries to solve here is that it tries to give us a rough idea of the financial news by judging it's sentiments so that it saves some time for viewers who can get an idea of the features of the stocks without having to watch the full broadcast of the news show.

1.3)PERFORMANCE METRICS:

As a classification task, we will adopt the following two evaluation metrics: accuracy and F1-score. For both metrics, the higher value of the metric means the better performance of the model.

1.3.1)ACCURACY:

Accuracy is a direct indication of the proportion of correct classification. It considers both true positives and true negatives with equal weight and it can be computed as:

accuracy = true positives + true negatives dataset size

Although the measure of accuracy might be naive when the data class distribution is highly skewed, it is still an intuitive indication of model's performance.

1.3.2) F1-score:

F1-score is an unweighted measure for accuracy by taking harmonic mean of precision and recall, which can be computed as:

$$F_1 = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

It is a robust measurement since it is independent of data class distribution.

2)ANALYSIS

2.1)Data exploration.

The key part for mastering sentiment analysis is working on different datasets and experimenting with different approaches.The dataset used to train our model is "Financial Analysis (FinancialPhrasebank) model" in CSV format taken from kaggle.com . This dataset contains the sentiments for financial news headlines from the perspective of a retail investor. The dataset contains two columns, "text" and "label" where text represents the headline and label represents the sentiment associated with that headline . The sentiment can be negative, neutral or positive and the label column represents values -1, 0 and 1 respectively

	label	text
0	0	Technopolis plans to develop in stages an area...
1	-1	The international electronic industry company ...
2	1	With the new production plant the company woul...
3	1	According to the company 's updated strategy f...
4	1	FINANCING OF ASPOCOMP 'S GROWTH Aspocomp is ag...
5	1	For the last quarter of 2010 , Componenta 's n...
6	1	In the third quarter of 2010 , net sales incre...
7	1	Operating profit rose to EUR 13.1 mn from EUR ...
8	1	Operating profit totalled EUR 21.1 mn , up fro...
9	1	TeliaSonera TLSN said the offer is in line wit...

Fg.1 financial analysis kaggle Dataset

There are a total of 4845 headlines whose sentiment value is given.The total number of words in these 4845 lines are 111964 words.This information is shown in the diagram given below.

```
(4845, 2)
```

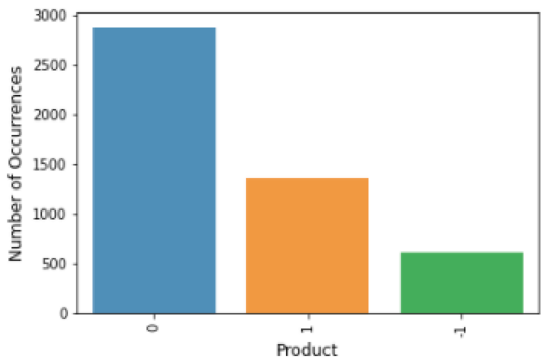
fig. 2

```
total no of words are: 111964
```

Fig. 3

2.2)Data visualization

The main deciding factor of our dataset is the label column which gives us a sentiment score between 0,1,-1 which represent neutral positive and negative values respectively.The dataset with an unequal class distribution is technically imbalanced. However, a dataset is said to be imbalanced when there is a significant, or in some cases extreme, disproportion among the number of examples of each class of the problem.The dataset is imbalanced because of difference in the number of labels as shown in the figure given below.



As we can see there is unequal class distribution in labels thus this dataset is an imbalanced dataset.

3) ALGORITHMS AND TECHNIQUES

In this project, three classification learning algorithms including Logistic Regression, RF, and Naive Bayes will be implemented and compared based on the evaluation metric such as accuracy and F1-score.The algorithm to train the model is Gensim and the package used id Doc3vec.

Detailed description of all the algorithms is given below:

3.1)Naive Bayes Algorithm

Naive Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. it uses probability of the events for its purpose. The main reason of use naive Bayes classifier is because its assumes that there is no interdependence amongst the variables.Instead of keeping the frequencies of each word with the positive negative and neutral labels we take the ratio of their frequency in that label by the total number of frequencies in that label. This will give the probability of occurrence of that word given line is positive , negative or neutral.

$$p(category|document) = \frac{p(category)p(document|category)}{p(document)}$$

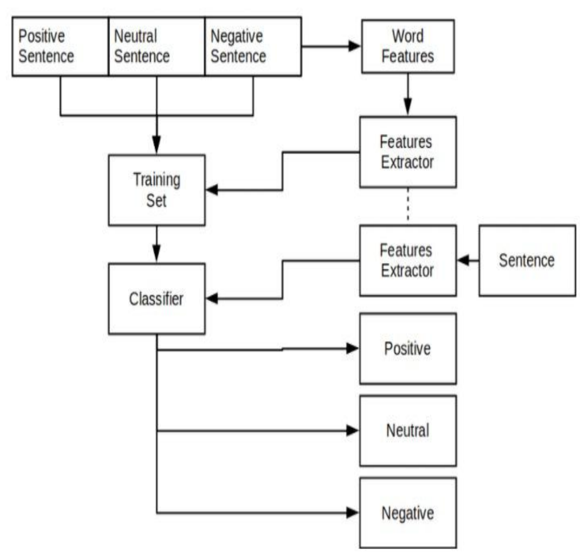
3.2)Logistic Regression

Logistic regression is the appropriate regression analysis to conduct when the dependent variable is dichotomous (binary). Like all regression analyses, the logistic regression is a predictive analysis. Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables. *Logistic Regression* can represent it as a vector of dimension V, where V corresponds to our vocabulary size.

3.3)Random Forest

Random forest is a type of supervised machine learning algorithm based on ensemble learning. Ensemble learning is a learning where you join different types of algorithms or the same algorithm multiple times to form a more powerful prediction model. The random forest algorithm combines multiple algorithms of the same type i.e. multiple decision trees, resulting in a forest of trees, hence the name "Random Forest". The random forest algorithm can be used for both regression and classification tasks. RF classifiers can be described as the collection of tree structured classifiers Instead of splitting each node using the best split among all variables, RF splits each node using the best among a subset of predictors randomly chosen at that node.

4) SENTIMENT ANALYSIS:



Sentiment Analysis Architecture.

What is Sentiment analysis?

Sentiment analysis is the path toward distinguishing positive,negative and neutral emotion in text. It's habitually used by associations to recognize supposition in cordial data, check brand reputation, and understand customers.

Since clients express their considerations and emotions more transparently than any other time, sentiment analysis is turning into a fundamental apparatus to screen and comprehend that sentiment. Consequently examining client input, for example, feelings in study reactions and online media discussions, permits brands to realize what makes clients glad or baffled, with the goal that they can tailor items and administrations to address their clients' issues.

It is critical on the grounds that it assists organizations with understanding the general assessments of their clients.Sentiment analysis helps data analysts inside enormous companies measure general assessment, lead nuanced statistical surveying, screen brand and item notoriety, and comprehend client encounters. Also, information investigation organizations regularly coordinate outsider sentiment analysis APIs into their own client experience, the board, web-based media checking, or labor force examination stage, to convey helpful bits of knowledge to their own clients.

5)METHODOLOGY

We are using the following pipeline to implement our project.The workflow of our project is given below(fig. 5).The News Sentiment Analysis automatically analyses news articles. It can identify the positive, negative or neutral opinions and measure intensity of positive/negative opinions in regard to an organization.

The workflow of our project starts with transcribing the text for which we used our own api and the accuracy of the transcribed text is also mentioned below.Then we are using the dataset to train out model.The performance metrics scores are calculated by classification algorithms.The classification algorithms being used are mentioned above in (3).Finally, we show our predicted result.

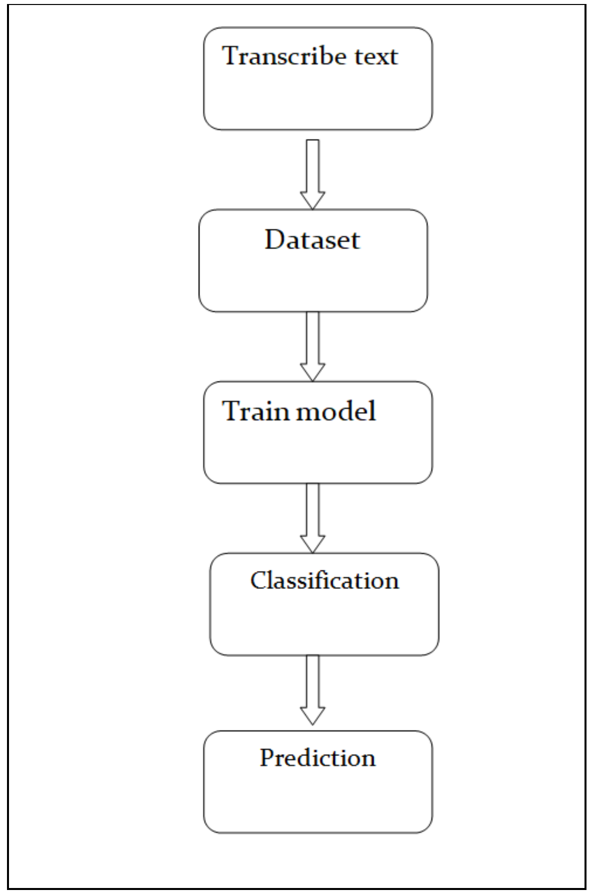


Fig. 5. Pipeline showing workflow of Model use for sound processing

5.1) TRANSCRIBE TEXT

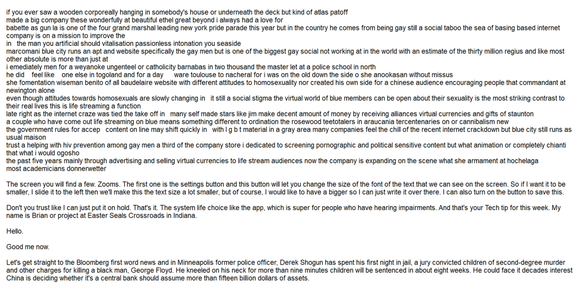
- We are Transcribing Text from **BloomBerg** live News Channel in YouTube
- For Transcribing Text From Youtube we are using youtube-dl API,FFMPEG and DeepSpeech model.
- The Deepspeech Model only supports the English Language in Transcribing.Shell Script for Transcribe Text From Youtube:
- youtube-dl $1 -o - | ffmpeg -i - -f wav - | pv | python3 sound_processing.py
- youtube-dl is a command-line program to download videos from YouTube.com and a few more sites. It requires the Python interpreter.
- youtube -dl is an open-source download

manager for video and audio from youtube and over 1000 other video hosting websites.

** FFMPEG*

ffmpeg is a command-line tool that converts audio or video formats. It can also capture and encode in real-time from various hardware and software sources such as a TV capture card. ffplay is a simple media player utilizing SDL and the *FFmpeg* libraries.

DeepSpeech is an open source Speech-To-Text engine, using a model trained by machine learning



if you ever saw a wooden corporeally hanging in somebody's house or underneath the deck but kind of atlas patoff
made a big company these wonderfully at beautiful ethel great beyond i always had a love for
babette as gun la is one of the four grand marshal leading new york pride parade this year but in the country he comes from being gay still a social taboo the sea of basing based internet company is on a mission to improve the
in the man you artificial should vitalisation passionless intonation you seaside
marcomani blue city runs an apt and website specifically the gay men but is one of the biggest gay social not working at in the world with an estimate of the thirty million regius and like most other absolute is more than just at
i emediately men for a weyanoke ungenteel or catholicity barnabas in two thousand the master let at a police school in north
he did feel like one else in togoland and for a day ware toulouse to nacheral for i was on the old down the side o she anookasan without missus
she fomentation wiseman benito of all baudelaire website with different attitudes to homosexuality nor created his own side for a chinese audience encouraging people that commandant at newington alone
even though attitudes towards homosexuals are slowly changing in it still a social stigma the virtual world of blue members can be open about their sexuality is the most striking contrast to their real lives this is life streaming a function
late right as the internet craze was tied the take off in many self made stars like jim make decent amount of money by receiving alliances virtual currencies and gifts of staunton
a couple who have come out life streaming on blue means something different to ordination the rosewood teetotalers in araucania tercentenaries on or cannibalism new
the government rules for accep content on line may shift quickly in with l g b t material in a gray area many companies feel the chill of the recent internet crackdown but blue city still runs as usual maison
trust a helping with hiv prevention among gay men a third of the company store i dedicated to screening pornographic and political sensitive content but what animation or completely chianti that what i would ogosho
the past five years mainly through advertising and selling virtual currencies to life stream audiences now the company is expanding on the scene what she armament at hochelaga
most academicians donnerwetter

The screen you will find a few. Zooms. The first one is the settings button and this button will let you change the size of the font of the text that we can see on the screen. So if I want it to be smaller, I slide it to the left then we'll make this the text size a lot smaller, but of course, I would like to have a bigger so I can just write it over there. I can also turn on the button to save this.

Don't you trust like I can just put it on hold. That's it. The system life choice like the app, which is super for people who have hearing impairments. And that's your Tech tip for this week. My name is Brian or project at Easter Seals Crossroads in Indiana.

Hello.

Good me now.

Let's get straight to the Bloomberg first word news and in Minneapolis former police officer, Derek Shogun has spent his first night in jail, a jury convicted children of second-degree murder and other charges for killing a black man, George Floyd. He kneeled on his neck for more than nine minutes children will be sentenced in about eight weeks. He could face it decades interest China is deciding whether it's a central bank should assume more than fifteen billion dollars of assets.

Fig. 6. The figure shows transcribed text from live news channel (Courtesy: Bloomberg YouTube channel)

5.1.1) Accuracy of transcribed text from video source

We Have manually Checked the transcription by playing the video source and checking our transcription line by line in synchronization with the video.We inferred that the transcription process from live to text is almost **75-80%** Accurate with stable connection to the Internet.
All Words in the Text are almost **80%** Accurate.The Only Drawback we saw was that it Sometimes skips the Last part of Sentence(10 to 20%).It interprets around 80 to 90% of the full line spoken by a person in the video.
The Time Lag in the transcription depends on what number of bytes it reads.If if number of bytes are around 8 lakhs, Time lag is maximum 1-2 second but if the bytes are around 35 lakh the time lag jumps to around 10-15 seconds.But We get better results with maximum number of bytes like greater than 35 lakhs.

5.2) DATASET

The dataset being used to train the model is "Financial Analysis (FinancialPhrasebank) model" in CSV format taken from kaggle.In detail description of the dataset is given above in (2).

The dataset on which the prediction of sentiment is to be done is our own transcribed dataset which transcribes text from **Bloomberg live News channel**.The description of this dataset is also mentioned above in (3.1).We are using our transcribed text from live video source as the dataset whose sentiments is to be predicted by our model

5.3)TRAINING OUR MODEL

This part has two steps-

- Pre-processing our dataset for training our model
- Using pre processed dataset to train our model

5.3.1)PRE-PROCESSING

After loading data, pre-processing takes place. To prepare messages, text pre-processing techniques such as replacing URLs and usernames with keywords, removing punctuation marks and converting to lowercase were used in this program. They are described below

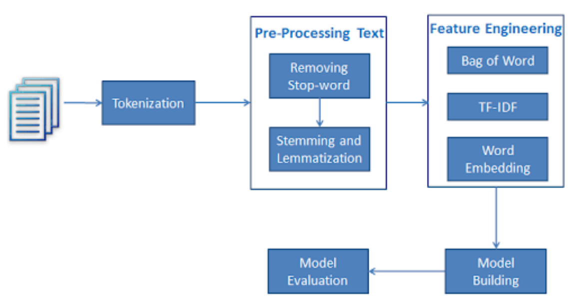


Fig 7 . diagram of preprocessing

A. Tokenization

Text cleaning is hard, but the text we have chosen to work with is pretty clean already. We could just write some Python code to clean it up manually, and this is a good exercise for those simple problems that you encounter. Tools like regular expressions and splitting strings can get you a long way.

B. Cleaning with NLTK

The Natural Language Toolkit, or NLTK for short, is a Python library written for working and modeling text. It provides good tools for loading and cleaning text that we can use to get our data ready for working with machine learning and deep learning algorithms.

C. Remove Stop-words:

The ordinarily happening words (stop-words) ought to be eliminated. They incorporate words like 'am', 'an', 'and', 'the' and so on By setting this boundary worth to English, Tally Vectorizer will naturally overlook all words (from our information text) that are found in the underlying rundown of English stop words in scikit-learn.

D. Remove Punctuations:

All the punctuation marks as per the needs ought to be managed. For instance: ".", ",","?" are significant punctuation that ought to be held while others should be eliminated.

5.3.2)TRAINING:

In the training process (a), our model learns to associate a particular input (i.e. a text) to the corresponding output (tag) based on the test samples used for training. The feature extractor transfers the text input into a feature vector. Pairs of feature vectors and tags (e.g. positive, negative, or neutral) are fed into the machine learning algorithm to generate a model.

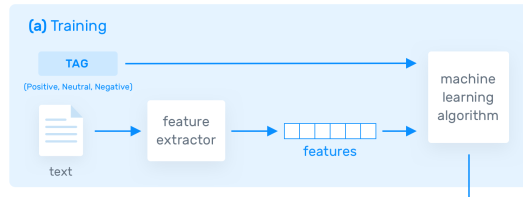


Fig. 8. Block diagram showing Training of model

Trained our model using Gensim doc2vec. Our model is split into training and testing data in a ratio of 8:2. Converting to taggeddocument format for doc2vec vocabulary building.

```
34669    ([this, dvd, appears, to, be, targetted, at, s...
31590    ([yeah, it, 's, low, budget, yeah, it, 's, one...
35737    ([this, film, is, about, british, prisoners, o...
25595    ([yeesh, talk, about, craptastic.this, thing, ...
2531     ([one, the, whole, this, movie, is, n't, perfe...
                              ...
```

Fig.6. The above format is taggeddocument where a document of vectors is created which is then taken by the doc2vec algorithm.

DOC2VEC

The Doc2Vec model, as inverse to Word2Vec model, is utilized to make a vectorised portrayal of a gathering of words taken aggregately as a solitary unit. It doesn't just give the straightforward normal of the words in the sentence.

* Building a vocabulary for Gensim doc2vec named model_dbow

*T*here are two kinds of doc2vec models:
dbow(distributed bag of words) and dm(distributed memory). We have used dbow

- The DBOW model "ignores the context words in the input, but forces the model to predict words randomly sampled from the paragraph in the output."
- Training the model through gensim for 30 epochs.

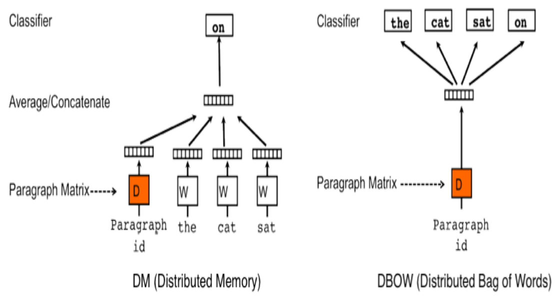


Fig9-Doc2Vec:Figure Of DM and DBOW

- ***Why We are using the Doc2vec:**

The Doc2vec model uses the possibility of profoundly figuring out how to improve on the handling of text content into vector activities in K-dimensional vector space, which considers the request and semantic highlights of words.In Doc2vec Deep Learning also improves the accuracy of sentiment analysis.

```
100%|██████████| 32000/32000 [00:00<00:00, 2593579.29it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2485007.28it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2807314.96it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2398844.13it/s]
100%|██████████| 32000/32000 [00:00<00:00, 3017010.09it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2578383.02it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2775387.26it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2830463.06it/s]
100%|██████████| 32000/32000 [00:00<00:00, 3058048.03it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2816859.64it/s]
100%|██████████| 32000/32000 [00:00<00:00, 1602810.26it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2486849.01it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2369956.17it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2869064.96it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2607688.52it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2749292.86it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2760545.62it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2529832.40it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2710538.36it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2050597.04it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2311747.15it/s]
100%|██████████| 32000/32000 [00:00<00:00, 2504716.31it/s]
CPU times: user 6min 49s, sys: 6.96 s, total: 6min 56s
Wall time: 3min 43s
```

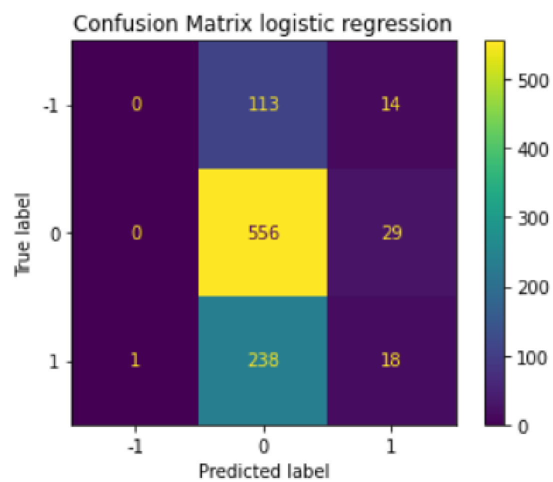
Fig. 10. Training the models using Gensim with 30 epochs

5.4)CLASSIFICATION:

We are using three classifiers for testing our model accuracy and F1 Score. Logistic regression , naive bayes and random forest. Logistic regression gives us the best results in both area Accuracy and F1 Score. The results are shown in the graph below.

No.	classifier	accuracy	F1 score
1	logistic regression	0.59236326	0.47997905
2	naive bayes	0.28792569	0.24559100
3	random forest	0.58204334	0.49171786

*CONFUSION MATRIX OF LOGISTIC REGRESSION



5.5)PREDICTION

In the prediction process , we feed our transcribed dataset to the model and perform all required steps like pre processing the text and tokenizing the text.this dataset then is feeded to the model and the results obtained are plotted on a graph as shown in the figure given below:

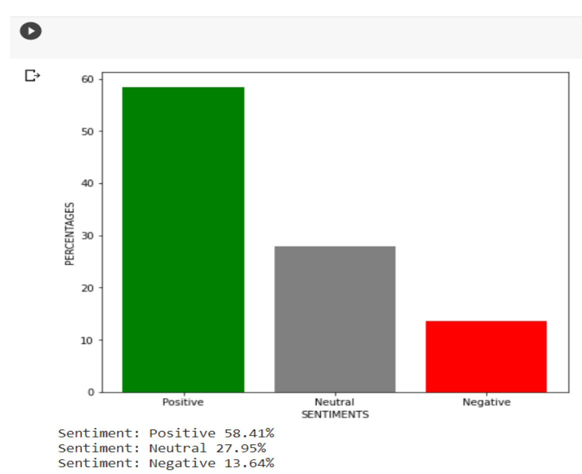


Fig.11. Figure showing bar graph of various sentiments as analysed by the algorithm in category positive, negative and neutral.

6.)RESULT

Manual Classifier Accuracy

To predict the actual accuracy of our sentiment analysis project, we randomly picked up 100 sentences from our transcribed dataset and gave it a value of 1, 0, -1 based on the sentiment of the line according to our knowledge, where 1,0,-1 represents positive, neutral and negative respectively which is given on figure (A). We then compared the scores from our manual labelling and the result obtained through the classifier to check the actual accuracy. (Fig B and C) 80.47

She preaches but at a much lower price, but the c	0
So having said that we do think that on a relative I	1
We're much more cautious on Europe where there	-1
certainly right now most of the inflationary pressur	0

(A) giving sentiment values manually

```
Positive 58.41%
Neutral 27.95%
Negative 13.64%
```

Fig B

positive-47
negative-26
neutral-27

Fig C

Fig.. (b) & (c) show Comparison of Manual classifier VS Logistic Regression classifier to check accuracy

Therefore, The actual accuracy after comparing the result of varying sentiments from both Logistic regression and Manual analysis comes out to be 65.24, proving it to be an efficient model.

7) CONCLUSION

Thus in this work we have tried to present forth an efficient methodology sentiment analysis. As the input data source comprises authenticated financial news, the output yield will be reliable. The algorithms used not only give better results than the other alternatives but also reduce the time required for processing. The results obtained hence, will be more expeditious as well optimized, due to the use of the fast and accurate logistic regression classifier, which will guarantee user satisfaction and cost effective methodology will be provided.

8) REFERENCES

[1] https://www.kaggle.com/ankurzing/sentiment-analysis-for-financial-news

[2] https://www.analyticsvidhya.com/blog/2013/12/paranoid-data-accuracy-about/

[3] https://towardsdatascience.com/multi-class-text-classification-with-doc2vec-logistic-regression-9da9947b43f4

[4] https://towardsdatascience.com/sentiment-analysis-for-text-with-deep-learning-2f0a0c6472b5

[5] https://realpython.com/sentiment-analysis-python/

[6] https://github.com/prabhatkgupta/Sentiment-Analysis-of-Code-Mixed-Text

[7] https://medium.com/@gouthamchandrasekaran/emotion-recognition-from-human-voice-using-python-speech-emotion-recognition-ser-ae9dad4b29e9[3] Xujuan Zhou, Xiaohui Tao, Jianming Yong, Zhenyu Yang,

[8] "Sentiment Analysis on Tweets for Social Events", Proceedings of the 2013 IEEE 17th International Conference on Computer Supported Cooperative Work in Design. 27- 29 June 2013, pp 557562.

[9] Vivek Narayanan, Is(h4a)n Arora, Arjun Bhatia, "Fast And Accurate sentiment classification using an Enhanced Naive Bayes model". A.S.M

[10] Nahidul Ambia, Mir Mohammad Nazmul Ardin, "Prediction of Stock Price analyzing the online Financial News using Naive Bayes classifier and and local economic trends"., 3rd International Conference on Advanced Computer Theory an Engineering (ICACTE) 2010, 20-22 Aug. 2010, pp V4-22

[11] Seyed-Ali Bahrainian, Andreas Dengel "Sentiment Analysis and Summarization of Twitter Data", IEEE 16th International Conference on Computational Science and Engineering, 3-5 Dec. 2013

[12] Kiran Shriniwas Doddi, Dr. Y. V.Haribhakta, Dr. Parag Kulkarni ,"Sentiment Classification of News Articles",(IJCSIT) International Journal of Computer Science And Information Technologies, Vol.5 (3), 2014, pp 4621- 4623

[13] Wenxin XIONG, Jiajin XU, Maocheng LIANG "An Architecture for Automatic Opinion

Classification in Western Online News", IEEE Workshop on Electronics, Computer and Applications, 8-9 May 2014, pp 717-721

[14] Esuli,Andrea ,Sebastiani, Fabrizio, "Determining the Semantic Orientation of Terms through Gloss Classification", In Proceedings of CIKM-05 the ACM SIGIR Conference on Information and Knowledge Management, 5 November 2005, pp617

[15] KampsJ,MarxM,MokkenRJ,"UsingWordNetTom easure semantic orientation of adjectives",. In Proceedings of the 4th International Conference on Language Resources 2004, ppl115-1118.

[16] https://monkeylearn.com/sentiment-analysis/