A Systematic Study on Sentimental Analysis and its Applications

Karthik Konar[1], Riddhi Patel[2]

[1]MCA Student, Dept. of Computer Engineering, NMIMS Mukesh Patel School of Technology Management & Engineering, Vile Parle(West) Mumbai.
[2]MCA Student, Dept. of Computer Engineering, NMIMS Mukesh Patel School of Technology Management & Engineering, Vile Parle(West) Mumbai.

---***---

Abstract - *Sentiment analysis is the process of determining the opinion or feeling from a piece of text. In this paper, we have presented our experimental results on the performance evaluation of various sentiment analysis applications using various approaches. We have implemented a). Sentiment analysis using the Machine Learning approach. b). Sentiment analysis using the VADER Sentimental Analysis approach to find the accuracy of the found polarity through a text document. c). The Lexicon approach has been used in a Chabot to find the polarity and subjectivity of the product review. We have also discussed the implementation with the help of the flowchart and code snippet. The efficiency of approaches implemented is being discussed and comparatively studied.*

Key Words – Chabot, Lexicon, Machine learning, Polarity, Subjectivity, Sentimental Analysis, Vader.

1. INTRODUCTION

Sentimental analysis is that particular domain, where you try to understand human emotions with the help of a software. Human emotions are in written form, and we classify those sentiments as positive, negative, neutral. Sentiment analysis is also called opinion mining because we are trying to analyse the thoughts of a customer with respect to that particular product. Nowadays the customer plays a very big role in making a business successful. A customer can make or break the organization, therefore it is very important for the organization to understand the sentiments of its customers so that the organization can reach heights. Therefore the sentimental analysis is essential. The sentimental analysis determines useful information, those information can be used to understand the market strategy, generate needs.

There are various applications of sentimental analysis such as Review classification, Product review mining, etc. In review classification, we classify the user reviews into 3 categories i.e. positive, negative, and neutral. For e.g.: There are a lot of customers, who post a lot of reviews, how do we know the sentiment of each customer? This is where we can use sentimental analysis and classify the reviews to be positive, negative, neutral. The next application is product review mining. For e.g.: There is a company which manufactured a product say 'A' now the company wants to know what features which are there in the product are liked by its customers, to determine that company can take the help of sentimental analysis. Sentimental analysis can also be used during election times. for e.g.: say there are 2 candidates i.e. "A" and "B" with the help of sentimental analysis we can find which candidate is more popular.

2. RESULTS AND DISCUSSION

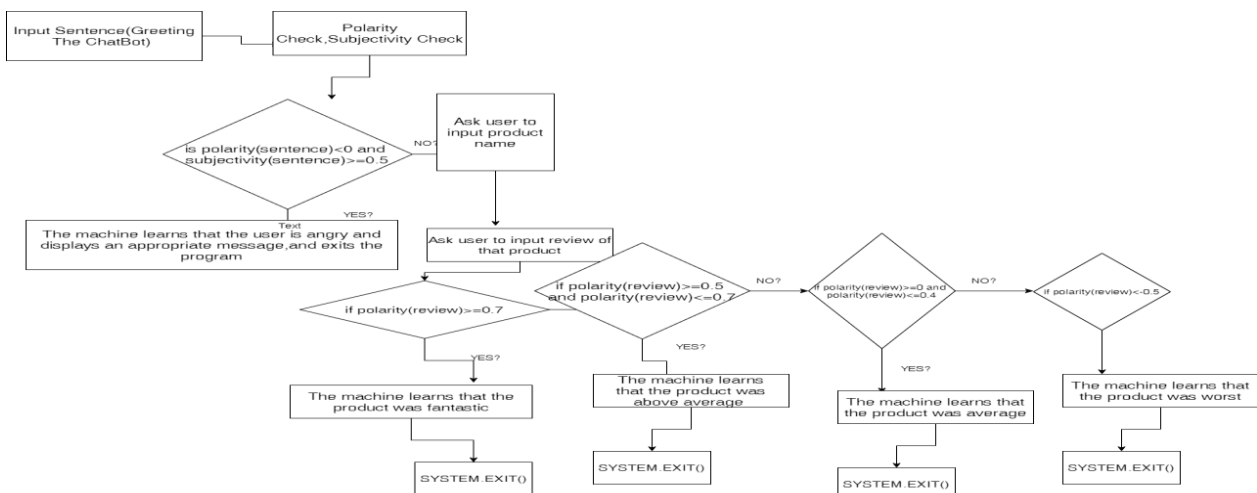### 2.1 System Architecture of ChatBot1 Developed for product review mining.

Fig 1.1: System Architecture of Chatbot for product Review Classification.

The algorithm for the above chatbot is shown as follows:

Step 1: A greeting message is displayed to the user. The user is asked to greet back the chatbot. The greeting sentence of the user is then subjected to undergo polarity and subjectivity checks. If the polarity of the sentence is less than 0 and the subjectivity of the sentence is greater than or equal to 0.5 the chatbot assumes that the user is angry and displays an appropriate message to the user and the chatbot terminates.

Step 2: if the polarity of the sentence is not less than 0 and the subjectivity of the sentence is less than 0.5 the chatbot assumes that the user is fine and happy and the chatbot asks the user to input the product name recently purchased by him/her, and then the chatbot asks the user to write a review about that particular product.

The review is then subjected to measure polarity on the text.

If the polarity of the sentence:

(**a**) Is greater than or equal to 0.7, the chatbot assumes that the product was fantastic and then it terminates.

(**b**) It is greater than or equal to 0.5 and less than 0.7, the chatbot assumes that the product was above average and then it terminates.

(**c**) Is greater than or equal to 0 and less than 0.4, the chatbot assumes that the product was average and then it terminates.

(**d**) It is less than -0.5, the chatbot assumes that the product was worse.

2.2 System Architecture which is developed for classifying the text as positive, negative, neutral using machine learning approach
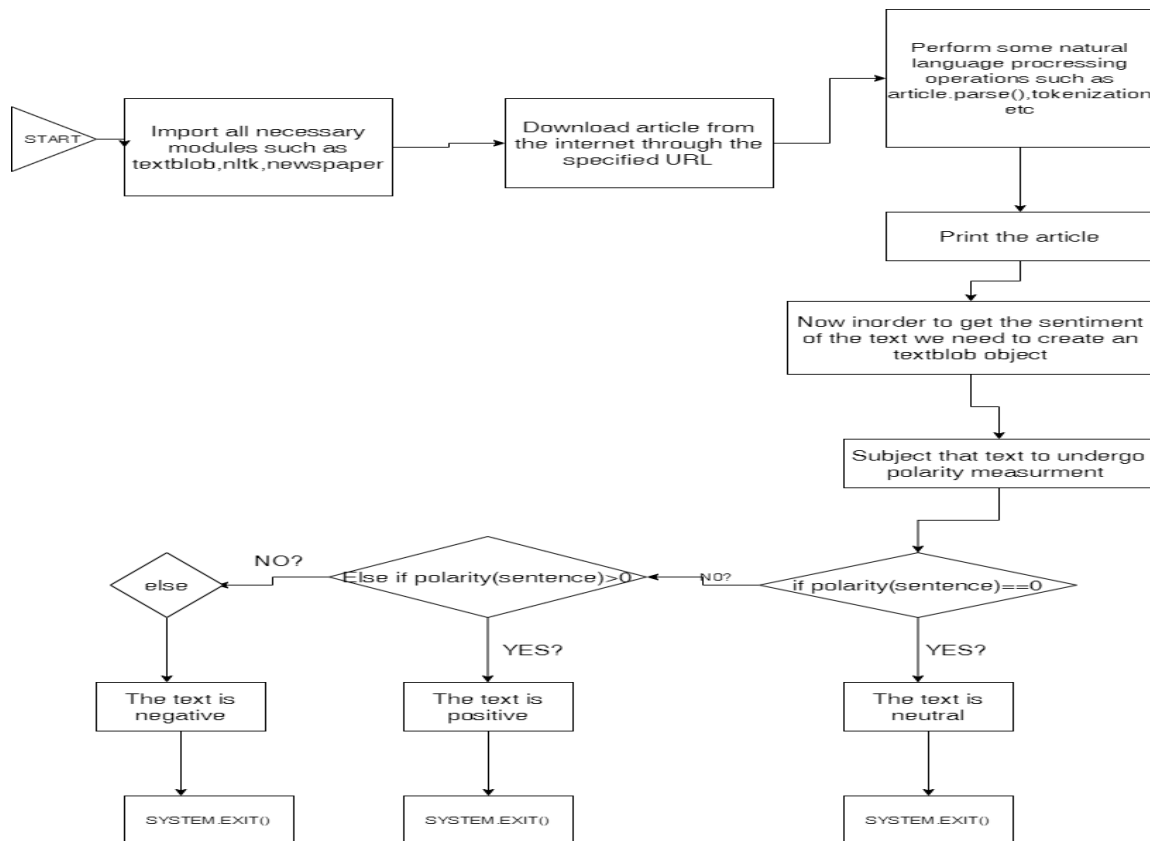

Fig 1.2: System Architecture developed for classifying the text as positive, negative, neutral using a machine learning approach.

The algorithm for the above architecture is as follows:

Step 1: Import all the necessary libraries in python such as Text Blob, newspaper, nltk.

Step 2: Specify the URL of the website which you wish to download to check the sentiment of the text.

Step 3: Perform some natural language processing operations.

Step 4: Print the article.

```
Computer Science (CS)

Computer Science (CS) is the study of the principles and use of computers. Simply put it's the science that deals with the theory and methods of processing information
in digital computers, the design of computer hardware and software, and the applications of computers. A computer is simply a tool for a computer scientist, like a tel
escope for an astronomer. Teach yourself !
Computer Science (CS)Computer Science (CS) is the study of the principles and use of computers.
Simply put it's the science that deals with the theory and methods of processing information in digital computers, the design of computer hardware and software, and th
e applications of computers.
A computer is simply a tool for a computer scientist, like a telescope for an astronomer.
```

Step 5: Now in order to know the sentiment of the text which is downloaded through the specified URL, you need to create an object of the TextBlob.

Step 6: After performing step (5) measure the polarity in each sentence.

Step 7: if polarity (sentence) ==0:

Then print "The text is neutral".

Else if polarity (sentence)>0:

Then print "The text is positive".

Else:

Print "The text is negative".

2.3 System Architecture developed for classifying the accuracy of positive and negative reviews using SentimentIntensityAnalyzer Module.

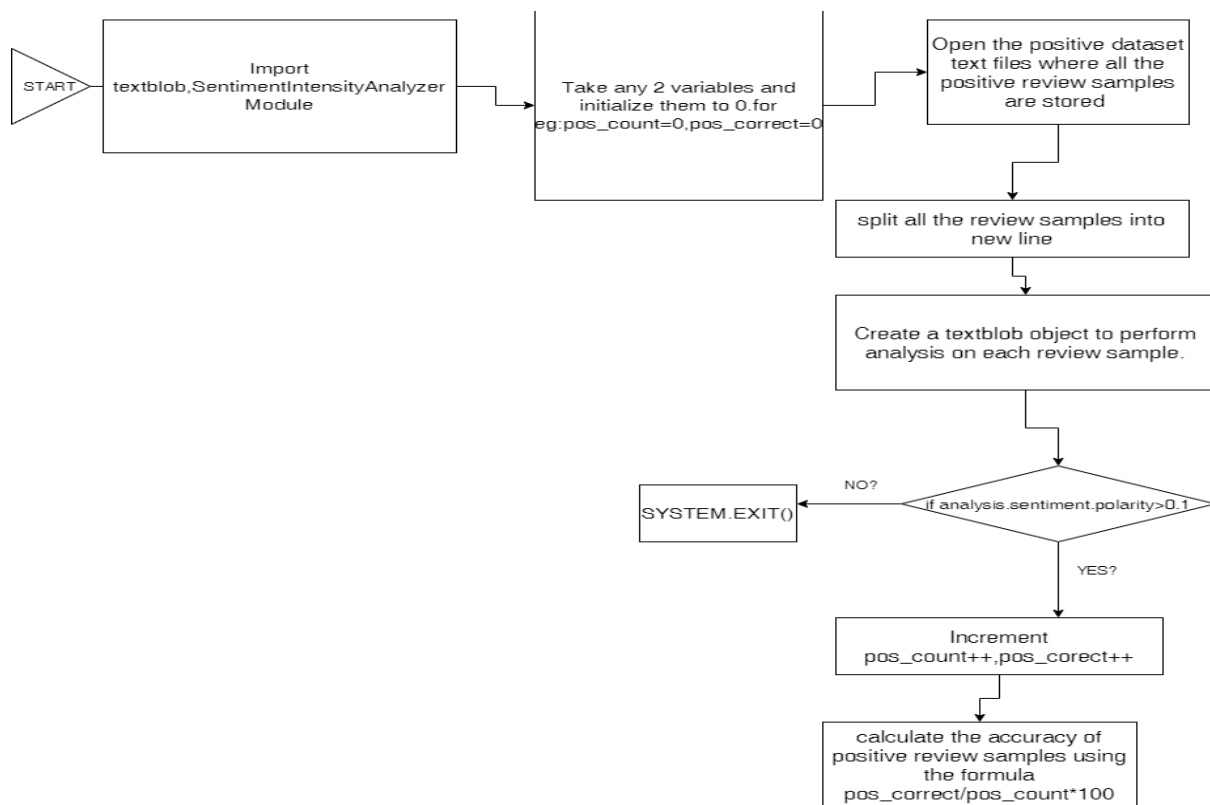**(a)To calculate the accuracy of positive dataset movie reviews.**

Fig 1.3: System Architecture developed for classifying the text as positive, negative, neutral using a feature vector approach.

The algorithm for the above architecture is as follows:

Step 1: Import all the necessary libraries in python such as TextBlob, SentimentIntensityAnalyzer.

Step 2: Now declare two variables for positive review ie.pos_count, pos_correct, and initialize them to 0.

Step 3: Open the positive review file where all the positive text is stored.

Step 4: Now run a for loop to read that text file and split the line in a new line.

Step 5: Declare a new variable to store the split line.

Step 6: After performing step (5) measure the polarity in each sentence.

Step 7: if polarity(sentence)>0.1:

    Then increment pos_correct

Increment pos_count

Step 8: Once the polarity is been checked now calculate the accuracy of positive review and print the accuracy

    Print "Positive accuracy= {} % via {} samples". Format (pos_correct/pos_count*100.0, poscount)

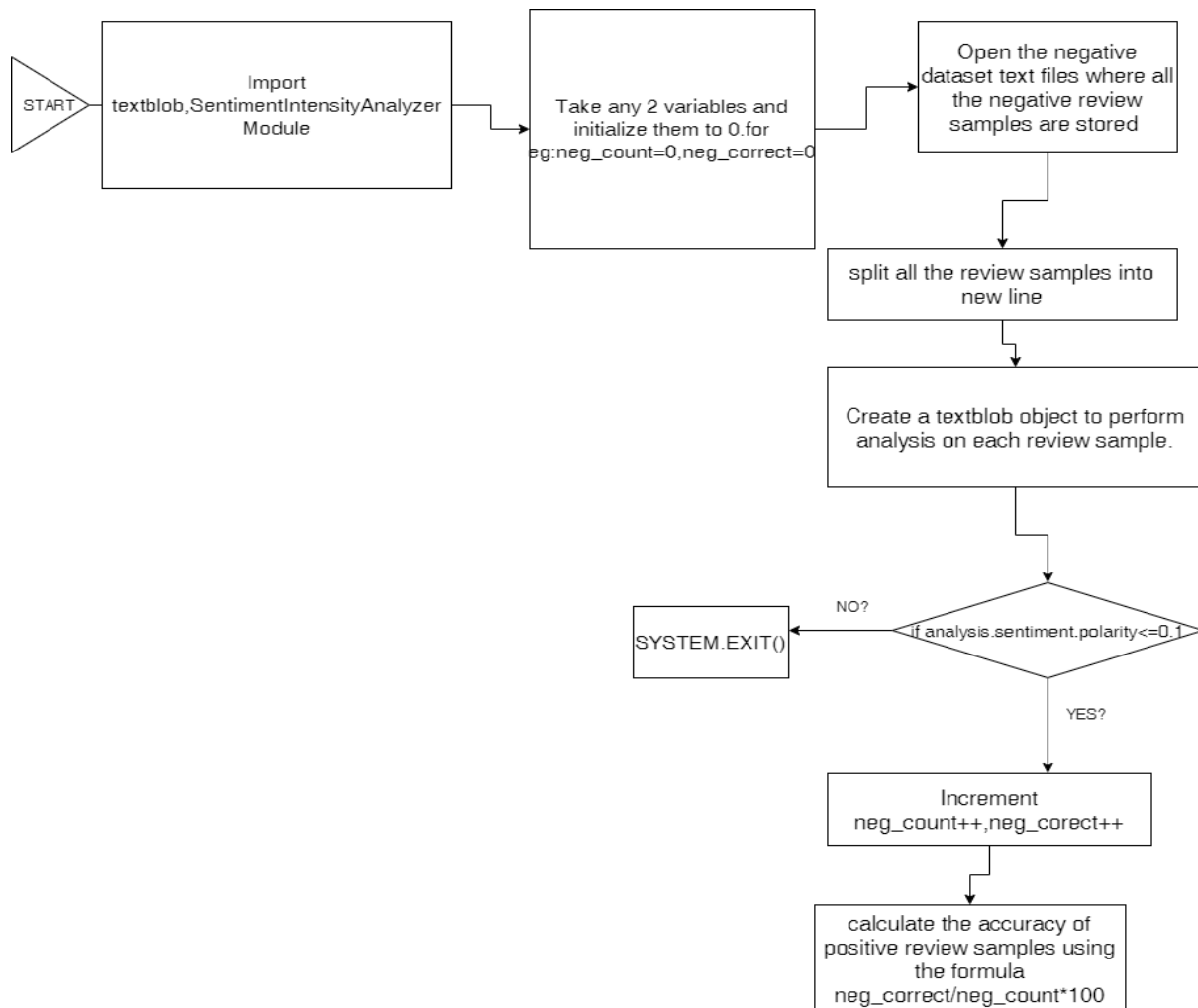**(b)To calculate the accuracy of negative dataset movie reviews.**

Fig 1.4: System Architecture developed for classifying the text as positive, negative, neutral using a feature vector approach.

The algorithm for the above architecture is as follows:

Step 1: Import all the necessary libraries in python such as TextBlob, SentimentIntensityAnalyzer.

Step 2: Now declare two variables for negative review ie.neg_count, neg_correct and initialize them to 0.

Step 3: Open the positive review file where all the negative text is stored.

Step 4: Now run a for loop to read that text file and split the line in a new line.

Step 5: Declare a new variable to store the split line.

Step 6: After performing step (5) measure the polarity in each sentence.

Step 7: if polarity (sentence) <0.1: Then increment neg_correct

Increment neg_count

Step 8: Once the polarity is been checked now calculate the accuracy of a negative review and print the accuracy

Print "Negative accuracy={}% via {}samples". Format(neg_correct/neg_count*100.0,poscount)

2.4 RESULTS

(I)RESULTS OBTAINED FOR THE MODULE WHICH WAS DEVELOPED USING LEXICON APPROACH. (FOR PRODUCT REVIEW MINING)

```
RESTART: C:\Users\Kiiciklaj\AppData\Local\Programs
Hello!! My name Is ChatBot.Press Enter To Continue

Hey there.I oouldn't recognize you.you are...?
you are the worst chatbot ever
hey..you need to calm down..i just asked your name
come back when you are calm..byee
```

Fig 1.5: In the above figure the polarity measured from the Greeting sentence is less than 0 .So therefore the chatbot learns that the user is angry, therefore displays appropriate messages and then terminates.

```
Hello!! My name Is ChatBot.Press Enter To Continue

Hey there.I couldn't recognize you.you are...?
my name is karthik
are you sure?
yes
Please input your name again to continue
karthik
karthik..Gotcha..!!
All Right.Now what's your gender.Male or Female?
male
ohkk......so let's start talking
SO how was the Product that you purchased?
i purchased iphone
Oh wow.. I'll look forward to purchase this product too..!! How is it?
it is awesome
I knew it would be fantastic. Thankyou for buying
Have you eaten yet? Its past 19
no
So let's first eat then talk.Byee
```

Fig 1.6 : In the above figure the polarity measured from the sentence is greater than 0.7 .so therefore the chatbot learns that the product is wonderful.

```
What was the last product you buyed?
nokia lumia
Oh wow.. I want to buy that  too..!! How is it?
it was above average phone
Oh..so it was a mediocre product
```

Fig 1.7: In the above figure the polarity measured from the sentence is greater than 0 and less than or equal to 0.4 .so therefore the chatbot learns that the product is average.

```
What was the last product you buyed?
earphones
Oh wow.. I want to buy that  too..!! How is it?
it is worst
Oh No...Was it that bad? I hope your money get refunded
```

Fig 1.8: In the above figure the polarity measured from the sentence is less than -0.5 so the chatbot learns that the product is worse

(II)RESULTS OBTAINED FOR THE MODULE WHICH WAS DEVELOPED FOR TEXT CLASSIFICATION USING MACHINE LEARNING APPROACH. (FOR TEXT CLASSIFICATION).

```
Computer Science (CS)

Computer Science (CS) is the study of the principles and use of computers. Simply put it's the science that deals with the theory and methods of processing information
in digital computers, the design of computer hardware and software, and the applications of computers. A computer is simply a tool for a computer scientist, like a tel
escope for an astronomer. Teach yourself !
Computer Science (CS)Computer Science (CS) is the study of the principles and use of computers.
Simply put it's the science that deals with the theory and methods of processing information in digital computers, the design of computer hardware and software, and th
e applications of computers.
A computer is simply a tool for a computer scientist, like a telescope for an astronomer.
Teach yourself !
0.0
0.0
The text is neutral
```

Fig 1.8: In the above figure the polarity measured from the article is equal to 0 so the machine learns that the text is neutral.

(III)RESULTS OBTAINED FOR THE MODULE WHICH WAS DEVELOPED FOR CLASSIFYING THE ACCURACY OF POSITIVE AND NEGATIVE REVIEWS USING SENTIMENT INTENSITY ANALYZER(FOR MOVIE REVIEW CLASSIFICATION)

```
Positive accuracy=60.5026256564141% via 5332 samples
Negative accuracy=68.37959489872468% via 5332 samples
```

5332 samples were tested and classified into positive and negative reviews and their accuracy was calculated.

3. OBSERVATIONS.

SR.NO	TECHNIQUE	ADVANTAGES	DISADVANTAGES	REMARKS
1	Lexicon Approach	Easy to implement, Easy to understand, Less complex	Does not provide 100% accurate results.	It is implemented on a chatbot which classifies product reviews as positive, negative, neutral. Mainly it is implemented for product review mining.

| 2 | Machine Learning Approach | Provides more accurate results when compared with lexicon based approach. | Little complex to implement, understand, time consuming. | It is implemented for text classification. A program was developed which downloads articles from the internet and then classifies those text obtained from the article as positive, negative, neutral. It works by measuring polarity of each and every sentence. |
| 3 | VADER Sentiment analysis | It provides the most accurate results when compared with the above 2 approaches. | NIL. | 5332 datasets were tested using this approach to classify the accuracy of positive and negative reviews. |

Table 1.1: Summary of Comparison.

V. CONCLUSION

This paper provides a detailed comparison of various applications of sentiment analysis which was implemented using different approaches such as lexicon, machine learning, and VADER sentiment analysis. Based on the experiment performed we found that VADER sentiment analysis provides the highest accuracy when compared with the lexicon and machine learning approach.5332 datasets were tested using the VADER Sentiment analysis approach to classify the accuracy of positive and negative reviews and the time taken for execution is very less when compared with the lexicon and machine learning approach.

REFERENCES

[1]. Chetashri Bhadane, Hardi Dalal, Helena Doshi, " Sentiment analysis: Measuring opinions".International Conference on Advanced Computing Technologies and Applications (ICACTA-2015)

[2] V.K. Singh, R. Piryani, A. Uddin, P. Waila, Marisha, "Sentimental analysis of Textual Reviews Evaluating Machine Learning, Unsupervised and SentiWordNet Approaches". 2013 5th International Conference on Knowledge and Smart Technology (KST)

[3] Harsha Sinha, Arashdeep Kaur, "A Detailed Survey and Comparative Study of Sentiment Analysis Algorithms". IEEE 2nd International Conference on Communication, Control and Intelligent Systems (CCIS)

[4] Akshay Amolik, Niketan Jivane, Mahavir Bhandari, Dr.M.Venkatesan, " Twitter Sentiment Analysis of Movie Reviews using Machine Learning Techniques". International Journal of Engineering and Technology (IJET)

[5] Praveen Kumar, Umesh Chandra Jaiswal, "A Comparative Study on Sentiment Analysis and Opinion Mining". International Journal of Engineering and Technology (IJET)

AUTHORS

Karthik Konar is a first year student pursuing MCA from SVKM NMIMS Mukesh Patel School of Technology Management and Engineering(MPSTME), Mumbai, India. His passion for research led him to gain interest in exploring new domains such as artificial intelligence, wireless sensor networks, Cyber forensics. He has published 3 research papers in the fields of wireless sensor networks, artificial intelligence, cyber forensics

Riddhi Patel is a first year student pursuing MCA from SVKM NMIMS Mukesh Patel School of Technology Management And Engineering(MPSTME),Mumbai, India. She has completed her B.SC IT from jai hind college, Mumbai, India. Her Research domain is artificial intelligence, computer networks.