

# Project Report

## Topic- Sentiment Analysis of Restaurant Reviews

### Using IBM Cloud



*Submitted by- Sneha Herlekar*  
[sneha.mohan2019@vitbhopal.ac.in](mailto:sneha.mohan2019@vitbhopal.ac.in)

## **1.Introduction**

### **1.1 Overview & Purpose of the Project**

Today Sentiment Analysis is a powerful marketing tool that enables a business company to understand the customer emotions towards the product or service. It allows to gain an overview of the wider public opinion behind a service or a product. Sentiment Analysis serves the purpose of helping a brand secure its edge over the other competitors by efficient decision making.

Everyday data is generated at a humongous rate and it is quite humanly impossible to analyze it. People share their opinions about a service or product over social media, curated websites, twitter etc. in order to make the process of analyzing these opinions or reviews, various machine learning techniques can be applied. This is where Sentiment analysis comes into picture.

Talking about the restaurants, although people still rely on the verbal review of any hotel, their online reviews and rating also matters. These reviews pivot customers to the restaurants. To improve their services and to get the overall feedback of the customers, sentiment analysis is done. Sentiment analysis is called finding the opinion from a large dataset which helps to analyze which restaurant is the best for customers who directly access the good reviews for the restaurant. System of sentiment analysis helps to convert unstructured information into structured information of public reviews, products, service, and brands. This helps in the field of commercial areas like marketing analysis, public dealings, reviews of product, promoters and scoring, feedback of product and service of products.

Sentiment analysis has different scopes which can be applied for three levels like document level, sentence level, sub sentence level. Using document level it can get hold of paragraphs or complete documents. By obtaining a single sentence it

helps to define a document . Sub expression can be achieved using sub sentences. Data analysis estimates that 80% of the world's data is unstructured. Most of the data starts from electronic mail, chats, community media, credentials and articles.

## **2. Literature Review**

Consumer satisfaction is one of the most important factors in marketing and research in terms of the behavior of consumers. For example, hotel consumers who receive excellent service will recommend it to others by word of mouth .

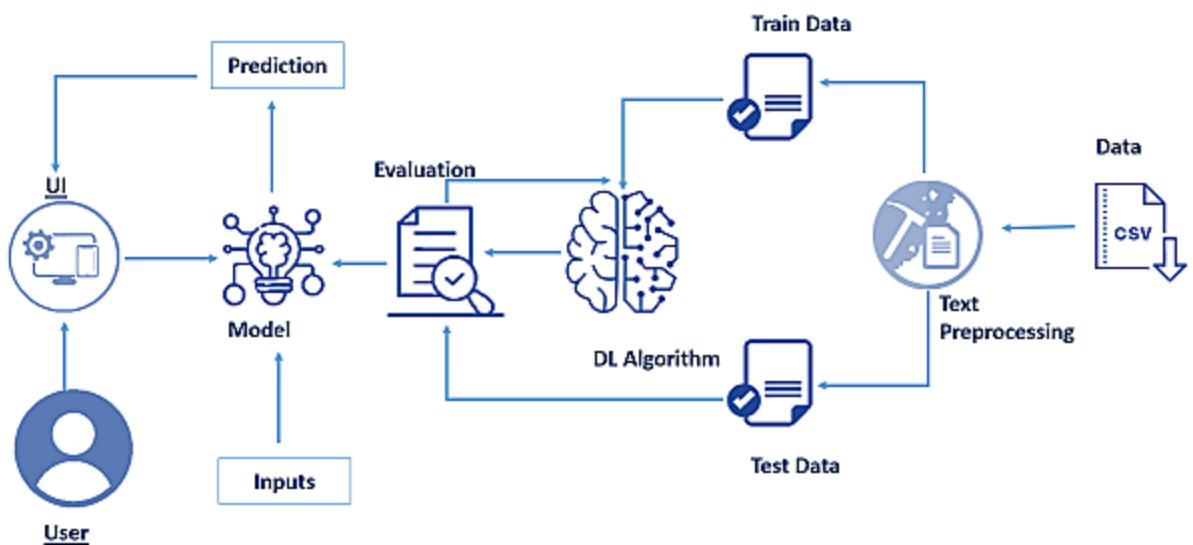
[3]Text mining, or the process of retrieving data from a collection of documents, is often done with software or analysis tools manuals. A text mining analysis can produce information that can be used to increase profits and improve services.

[2]Sentiment analysis is used to find the author's opinions about a specific entity. A sentiment analysis of a review is an analysis of an opinion about a product. Based on this information Natural Language Processing (NLP), text analysis, and computational parts are used in sentiment analysis to eliminate or extract unnecessary parts of the sentence to determine if it is negative or positive.

The computing power provided by A. Sadilekhas [1] to monitor health and epidemiology continues to grow. Future work will include developing an end-to-end system that can identify restaurant risks automatically. On Twitter, online users can discover individuals who may be ill with a foodborne illness from a colony.

### 3. Theoretical Analysis

#### 3.1 Architecture Diagram

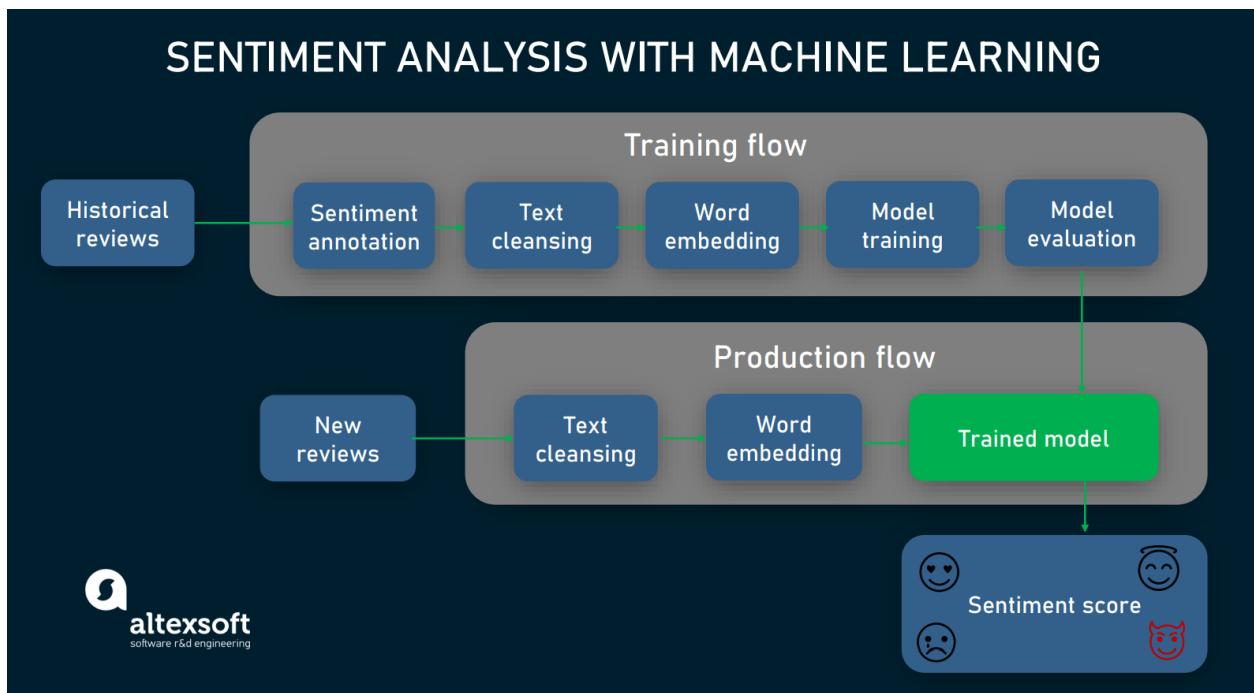


On the basis of reviews and ratings, most customers choose the best restaurants. Review systems and models are essential to any kind of model or system. This problem is approached by using natural language processing (NLP) and machine learning techniques based on review text content analysis. Using the above method, we can classify the review as positive or negative, and as well as see how many positive or negative reviews there are in total. By combining a predefined dictionary of words with an algorithm, we can classify the reviews on the basis of positive and negative reviews. Users can enter feedback in a web application, the feedback is analyzed by the model built, and predictions are displayed on the user interface.

### 3.2 Project Requirements

- Python 3
- Anaconda Navigator
- Jupyter Notebook
- Spyder
- Flask
- Python Libraries
  - Pandas
  - Numpy
  - Scikit Learn
  - Keras
  - Pickle
  - NLTK

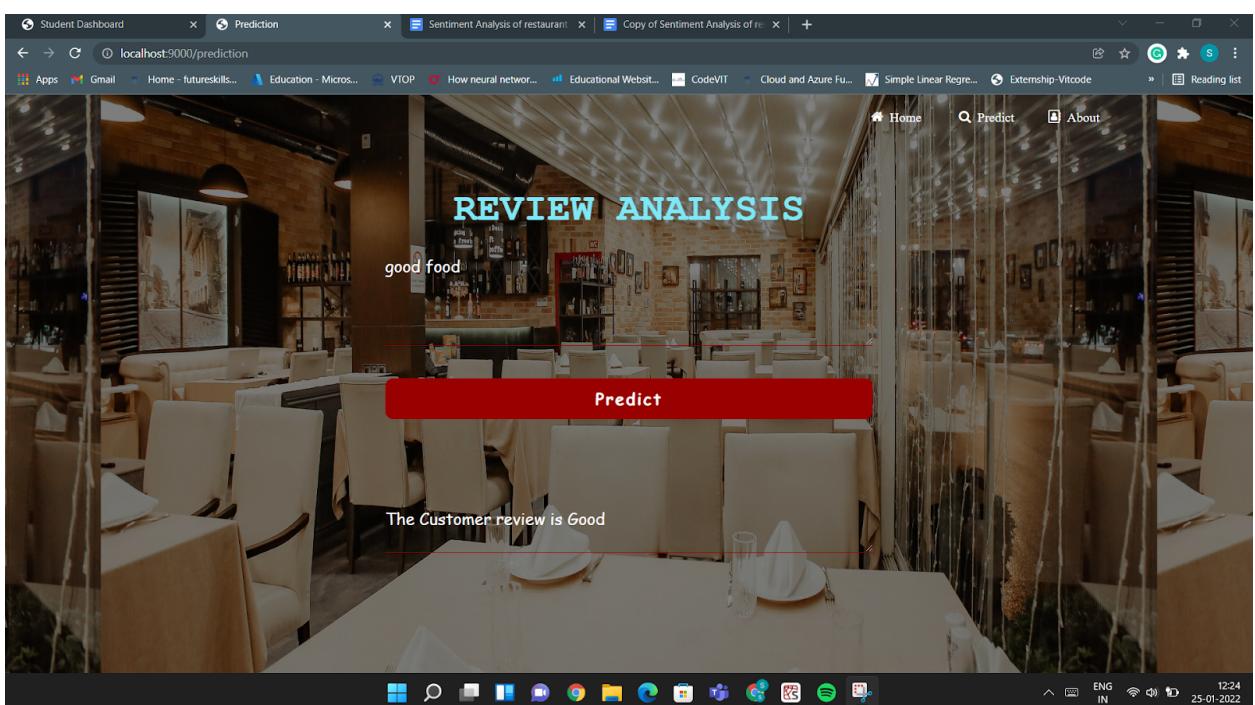
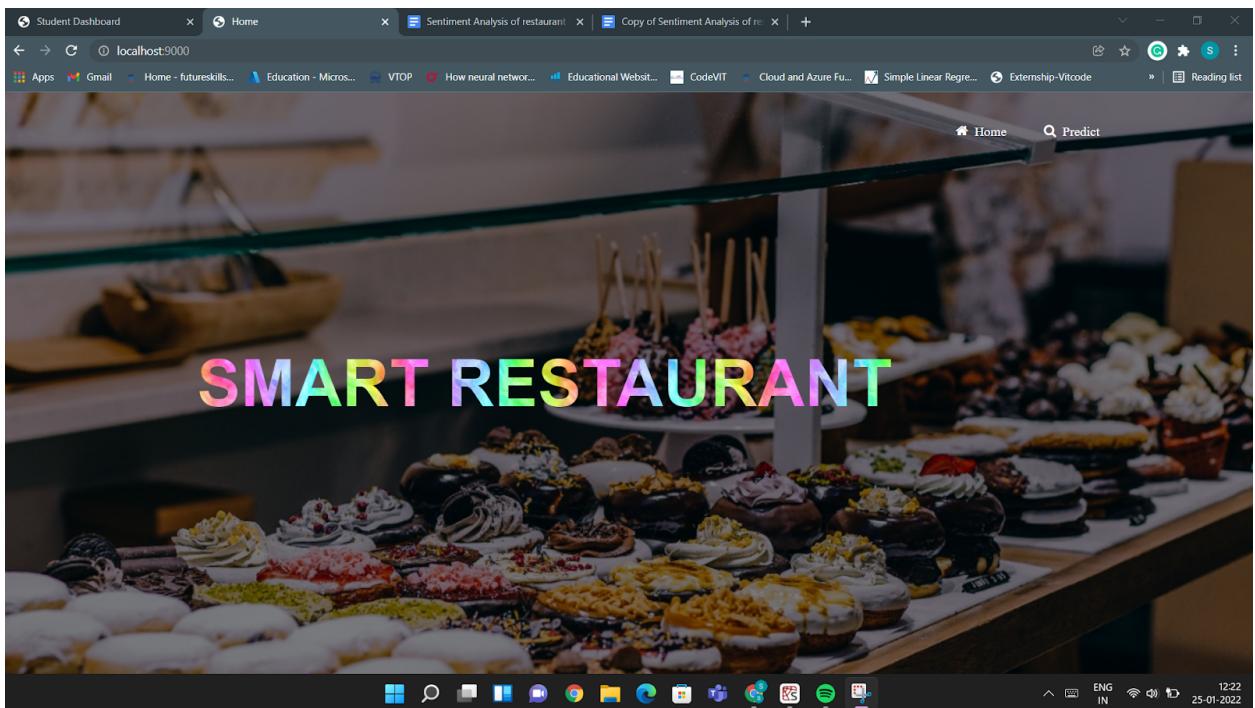
## **4. Working**



- The important part is training the model on the existing data available.
- In this project we use the Zomato review dataset. Getting the dataset is a part of data collection.
- Then comes Cleansing the data, here the numbers and the text data is separated, special characters , punctuations are removed, the text is made in lowercase.
- Then tokenization, vectorization and stemming are performed on the dataset using the nltk library of python.
- This data is then split into test and training data.
- The training is then fed to the neural networks.
- Then the prediction on the trained model is made using new text and then sentiment analysis is made based on the previous data.

## 5. Results and outcomes

## Restaurant Review App:



## Flask Deployment:

The screenshot shows the Spyder Python IDE interface. The top menu bar includes File, Edit, Search, Source, Run, Debug, Consoles, Projects, Tools, View, Help. The left sidebar has icons for file operations like Open, Save, Find, and Run. The main area shows a code editor with Python code for a Flask application. The code imports numpy, pandas, flask, tensorflow, and other modules, handles stopwords, and defines routes for home and prediction pages. A file browser on the right lists files in the current directory, including static assets and project files. The bottom right contains a terminal window showing command-line output and a status bar with various toolbars and system information.

```
1 import numpy as np
2 import pandas as pd
3 from flask import Flask, render_template, request
4 from tensorflow import WSGIServer
5 from gevent import pywsgi
6 import os
7 import tensorflow.compat.v1 as tf
8 tf.disable_v2_behavior()
9 #global graph
10 #graph = tf.get_default_graph()
11
12 from tensorflow.keras.models import load_model
13 import pickle
14
15 with open('cv.pkl', 'rb') as file:
16     cv = pickle.load(file)
17
18 import re
19 import nltk
20
21 nltk.download("stopwords")
22 from nltk.corpus import stopwords
23 from nltk.stem.porter import PorterStemmer
24
25 ps = PorterStemmer()
26
27 model = load_model("zomato_2_analysis.h5", compile=False)
28
29 app = Flask(__name__, template_folder="template")
30
31 @app.route('/')
32 def welcome():
33     return render_template('home.html')
34
35 @app.route('/prediction', methods=['GET', 'POST'])
36 def pred():
37     if request.method == "POST":
38         review = request.form['message']
39         print(review)
40         review = re.sub('[^a-zA-Z]', ' ', review)
41         review = review.lower()
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
259
260
261
262
263
264
265
266
267
268
269
269
270
271
272
273
274
275
276
277
278
279
279
280
281
282
283
284
285
286
287
288
289
289
290
291
292
293
294
295
296
297
298
299
299
300
301
302
303
304
305
306
307
308
309
309
310
311
312
313
314
315
316
317
318
319
319
320
321
322
323
324
325
326
327
328
329
329
330
331
332
333
334
335
336
337
338
339
339
340
341
342
343
344
345
346
347
348
349
349
350
351
352
353
354
355
356
357
358
359
359
360
361
362
363
364
365
366
367
368
369
369
370
371
372
373
374
375
376
377
378
379
379
380
381
382
383
384
385
386
387
388
389
389
390
391
392
393
394
395
396
397
398
399
399
400
401
402
403
404
405
406
407
408
409
409
410
411
412
413
414
415
416
417
418
419
419
420
421
422
423
424
425
426
427
428
429
429
430
431
432
433
434
435
436
437
438
439
439
440
441
442
443
444
445
446
447
448
449
449
450
451
452
453
454
455
456
457
458
459
459
460
461
462
463
464
465
466
467
468
469
469
470
471
472
473
474
475
476
477
478
479
479
480
481
482
483
484
485
486
487
488
489
489
490
491
492
493
494
495
496
497
498
499
499
500
501
502
503
504
505
506
507
508
509
509
510
511
512
513
514
515
516
517
518
519
519
520
521
522
523
524
525
526
527
528
529
529
530
531
532
533
534
535
536
537
538
539
539
540
541
542
543
544
545
546
547
548
549
549
550
551
552
553
554
555
556
557
558
559
559
560
561
562
563
564
565
566
567
568
569
569
570
571
572
573
574
575
576
577
578
579
579
580
581
582
583
584
585
586
587
588
589
589
590
591
592
593
594
595
596
597
598
599
599
600
601
602
603
604
605
606
607
608
609
609
610
611
612
613
614
615
616
617
618
619
619
620
621
622
623
624
625
626
627
628
629
629
630
631
632
633
634
635
636
637
638
639
639
640
641
642
643
644
645
646
647
648
648
649
650
651
652
653
654
655
656
657
658
659
659
660
661
662
663
664
665
666
667
668
669
669
670
671
672
673
674
675
676
677
678
679
679
680
681
682
683
684
685
686
687
687
688
689
689
690
691
692
693
694
695
696
697
697
698
699
699
700
701
702
703
704
705
706
707
708
709
709
710
711
712
713
714
715
716
717
718
719
719
720
721
722
723
724
725
726
727
728
729
729
730
731
732
733
734
735
736
737
738
739
739
740
741
742
743
744
745
746
747
748
748
749
750
751
752
753
754
755
756
757
758
759
759
760
761
762
763
764
765
766
767
768
769
769
770
771
772
773
774
775
776
777
778
779
779
780
781
782
783
784
785
786
787
788
789
789
790
791
792
793
794
795
796
797
797
798
799
799
800
801
802
803
804
805
806
807
808
809
809
810
811
812
813
814
815
816
817
818
819
819
820
821
822
823
824
825
826
827
828
829
829
830
831
832
833
834
835
836
837
838
839
839
840
841
842
843
844
845
846
847
848
848
849
850
851
852
853
854
855
856
857
858
859
859
860
861
862
863
864
865
866
867
868
869
869
870
871
872
873
874
875
876
877
878
879
879
880
```

The screenshot shows the Spyder Python IDE interface. The top menu bar includes File, Edit, Search, Source, Run, Debug, Consoles, Projects, Tools, View, Help. The left pane contains a code editor with Python code for a web application. The right pane features a file browser, a variable explorer, and a terminal window.

Code Editor Content:

```
File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\kavik\VIT\flask_practice\web.py
web.py  home.html  prediction.html  project_data.html

30
31 @app.route('/')
32 def welcome():
33     return render_template('home.html')
34
35
36 @app.route('/prediction', methods=['GET', 'POST'])
37 def pred():
38     if request.method == 'POST':
39         review = request.form['message']
40         print(review)
41         review = re.sub("[^a-zA-Z]", ' ', review)
42         review = review.lower()
43         review = review.split()
44         review = [ps.stem(word) for word in review if not word
45                     in set(stopwords.words('english'))]
46         review = ' '.join(review)
47         review = cv.transform([review]).toarray()
48
49         with graph.as_default():
50             y_p = model.predict(review)
51             print(y_p)
52             if y_p.argmax() == 0:
53                 output = "Average"
54             elif y_p.argmax() == 1:
55                 output = "Good"
56             else:
57                 output = "Poor"
58             return render_template('prediction.html', prediction=
59             ("The Customer review is " + output))
60     else:
61         return render_template('prediction.html')

64 @app.route('/project')
65 def project():
66     return render_template("project_data.html")
67
68
69 if __name__ == '__main__':
70     app.run(host='localhost', port=9000, debug=False, threaded=False)
```

File Browser:

Name	Date Modified
static	22-01-2022 05:29 PM
happy1.jpg	19-01-2022 10:00 PM
home.css	22-01-2022 05:27 PM
home.jpeg	19-01-2022 09:59 PM
neon.jpg	19-01-2022 09:59 PM
ok.gif	19-01-2022 10:00 PM
prediction.css	22-01-2022 05:27 PM
prediction.jpg	19-01-2022 10:00 PM
project_data.css	22-01-2022 05:27 PM
project_data.jpg	19-01-2022 10:00 PM
sad.jpg	19-01-2022 10:00 PM
template	24-01-2022 06:59 PM
cv.pkl	22-01-2022 03:45 PM
Restaurant Review Analysis.ipynb	19-01-2022 07:50 PM
web.py	24-01-2022 08:13 PM
zomato_2_analysis.h5	19-01-2022 04:27 PM

Console 2/A:

```
get_or_select_template
    return self.get_template(template_name_or_list, parent, globals)
File "C:\Users\kavik\anaconda\lib\site-packages\jinja2\environment.py", line 883, in get_template
    self._load_template(name, self._make_globals(globals))
File "C:\Users\kavik\anaconda\lib\site-packages\jinja2\environment.py", line 857, in _load_template
    template = self.loader.load(name, globals)
File "C:\Users\kavik\anaconda\lib\site-packages\jinja2\loaders.py", line 115, in load
    source, filename, uptodate = self._get_source(environment, name)
File "C:\Users\kavik\anaconda\lib\site-packages\flask\templating.py", line 60, in get_source
    return self._get_source_fast(environment, template)
File "C:\Users\kavik\anaconda\lib\site-packages\flask\templating.py", line 89, in _get_source_fast
    raise TemplateNotFound(template)
jinja2.exceptions.TemplateNotFound: project_data.html
127.0.0.1 - [24/Jan/2022 23:25:19] "GET /project HTTP/1.1" 500 -
127.0.0.1 - [24/Jan/2022 23:25:24] "GET / HTTP/1.1" 200 -
```

## **5. Advantages and Disadvantages**

An organization or group whose public perception is important for their success - no matter how it is defined - can benefit from sentiment analysis. On social media, blogs, and online forums, millions of people are busy discussing and reviewing businesses, companies, and organizations. The opinions of these people are being 'listened to' and analysed. Those being discussed are taking advantage of this enormous amount of data by using computer programs that not only locate all mentions of their products, but also determine the emotions and attitudes behind them. Sentiment analysis helps businesses understand what is being said about their products and services, so they can take action accordingly. They can quickly recognize any negative sentiments being expressed, and turn poor customer experiences into the kind of experiences that

customers appreciate. This can result in better products and services, and it will allow them to formulate marketing messages that will appeal to their target audience. This will ultimately lead to increased revenue and sales.

One of the biggest limitations of sentiment analysis is that it is not completely accurate. Sentiment analysis can identify and analyse many pieces of text automatically and quickly. The problem is that computer programs have difficulty recognizing things like sarcasm and irony, negations, jokes, and exaggerations, which a person would be able to easily detect. A failure to recognize them can cause results to be skewed.

## **6. Conclusion**

Sentiment analysis can be useful to many aspects of business from a monitoring of products analytics and from a customer production. By incorporating into existing systems and brands it is ready to work more rapidly with more accuracy for useful works .To examine critically and to bring out the essential elements or give the essence to analyze a data. To examine carefully and in detail so as to key factors,

possible results.

## **7. Future Scope**

Today Sentiment analysis is looked at only as a classification of negative or positive. But in future it will delve deeper in understanding and signify the conversations and what they reveal about the consumers. This will help brands to more effectively customize and deliver their services.

## **8. References**

- [1] A. Sadilek, S. Brennan, H. Kautz, and V. Silenzio. Nemesis: Which restaurants should you avoid today? In HCOMP, 2013.
- [2] R. Feldman, "Techniques and applications for sentiment analysis ,"  
Communications of the ACM, vol. 56, no . 4, pp.82-89,2013
- [3] E. Haddi, X. Liu, and Y. Shi, "The role of text pre-processing in  
sentiment analysis," Procedia Computer Science, vol. 17, pp.26-32,2013