Project Report

Sentiment Analysis & Rating Prediction for Customer Reviews

Project Report submitted on the fulfilment of the requirements of Post graduate Diploma in Big data Analytics

Authors:

1. Kadam Vinod Maruti 220960925016
2. Kenge Durgesh Raosaheb 220960925017
3. Kothekar Harshada Vinayak 220960925019
4. Kulkarni Shreya Avinash 220960925020
5. Kunal Bendale 220960925021
6. Patel Bhaumik Narendrabhai 220960925024

Co-ordinators:

1. Ms. Roopa Panicker (Course Co-ordinator)
2. Ms. Divya Das (Project Guide)
3. Ms. Soorya M (Project Guide)

STDC
CDAC Thiruvananthapuram
Trivandrum, Kerala 695581

Index

Abstract

Review websites, such as TripAdvisor and Yelp, allow users to post online reviews for various businesses, products and services, and have been recently shown to have a significant influence on consumer shopping behaviour. An online review typically consists of free-form text and a star rating out of 5. The problem of predicting a user's star rating for a product, given the user's text review for that product, is called Review Rating Prediction and has lately become a popular, albeit hard, problem in machine learning. In this paper, we treat Review Rating Prediction as a multi-class classification problem, and build six different prediction models by combining two feature extraction methods, (i) Removal of stopwords and punctuation, (ii) Vectorization by applying Count Vectorizer, with four machine learning algorithms, (i) K Nearest Neighbour classifier, (ii) Decision Tree classification, (iii) Random Forest Classifier, and (iv) Support Vector Classification (v) Multinomial Naïve Bayes Classification (vi) Multilayer Perceptron Classifier. We analyse the performance of each of these six models to come up with the best model for predicting the ratings from reviews. We use the dataset provided by Yelp for training and testing the models.

Introduction

The Businesses get a lot of reviews in form of comments and stars sometimes the websites do not have a rating system but rather just a comment system. So going through all the comments is not efficient. Our project seeks to provide solution to these businesses and predict the ratings along with the sentiments of customers from the comment.

In this project we have used Yelp business dataset which consist of Yelp's businesses, reviews, and user data. In this dataset, you'll find information about hotel businesses across 8 metropolitan areas in the USA and Canada.

This project proposes the use of Natural Language Processing (NLP) and various Machine Learning (ML) and Deep Leaning algorithms, such as decision tree, random forest, support vector machine, multinomial naïve bayes, KNN and multilayer perceptron to classify sentiments as positive, average, and negative.

The performance of these algorithms will be evaluated using precision, recall, and accuracy metrics. The results of this project will demonstrate the effectiveness of using NLP and ML algorithms to classify the sentiments of the customers and predict the ratings thereby helping the business to take wise business decision.

Literature Survey

Most of the recent work related to review rating prediction relies on sentiment analysis to extract features from the review text. Several studies have been conducted to evaluate the effectiveness of this approach in analysing the sentiments and classify them to take wise business decision.

Sunmin Lee et al

[1] The study focused on the initial stage of the algorithm by answering the research question that can the Bidirectional Encoder Representations from Transformers model determine whether a customer's review on Yelp is positive or negative, and the degree of said positivity or negativity, based on the review's content.

Boya Yu et al

[2] The main approach used in this paper is to use a support vector machine (SVM) model to decipher the sentiment tendency of each review from word frequency. Word scores generated from the SVM models are further processed into a polarity index indicating the significance of each word for special types of restaurants. Customers overall tend to express more sentiment regarding service.

Parikh et al

[3] This study's purpose was to identify factors of usage, trust, influence, and contribution of restaurant reviews on Yelp.com. This study found that information search reduction and community membership were the greatest factors encouraging Yelp.com use.

Proposed System

A. Dataset Pre-processing:

We first write some basic Python commands for exploratory data analysis on the data. Also created a column named ‘length’ to calculate the number of words in a review.

```
Few dataset entries:
              business_id        date               review_id  stars  \
0  9yKzy9PApeiPPOUJEtnvkg  2011-01-26  fWKvX83p0-ka4JS3dc6E5A      5
1  ZRJwVLyzEJq1VAihDhYiow  2011-07-27  IjZ33sJrzXqU-0X6U8NwyA      5
2  6oRAC4uyJCsJl1X0WZpVSA  2012-06-14  IESLBzqUCLdSzSqm0eCSxQ      4
3  _1QQZuf4zZOyFCvXc0o6Vg  2010-05-27  G-WvGaISbqqaMHlNnByodA      5
4  6ozycU1RpktNG2-1BroVtw  2012-01-05  1uJFq2r5QfJG_6ExMRCaGw      5

                                                text    type  \
0  My wife took me here on my birthday for breakf...  review
1  I have no idea why some people give bad review...  review
2  love the gyro plate. Rice is so good and I als...  review
3  Rosie, Dakota, and I LOVE Chaparral Dog Park!!...  review
4  General Manager Scott Petello is a good egg!!!...  review

                  user_id  cool  useful  funny
0  rLtl8ZkDX5vH5nAx9C3q5Q     2       5      0
1  0a2KyEL0d3Yb1V6aivbIuQ     0       0      0
2  0hT2KtfLiobPvh6cDC8JQg     0       1      0
3  uZetl9T0NcROGOyFfughhg     1       2      0
4  vYmM4KTsC8ZfQBg-j5MWkw     0       0      0
```

```
[ ] # SHAPE OF THE DATASET
    print("Shape of the dataset:")
    print(data.shape)
```

```
Shape of the dataset:
(10000, 10)
```

```
[ ] # COLUMN NAMES
    print("Column names:")
    print(data.columns)
```

```
Column names:
Index(['business_id', 'date', 'review_id', 'stars', 'text', 'type', 'user_id',
       'cool', 'useful', 'funny'],
      dtype='object')
```

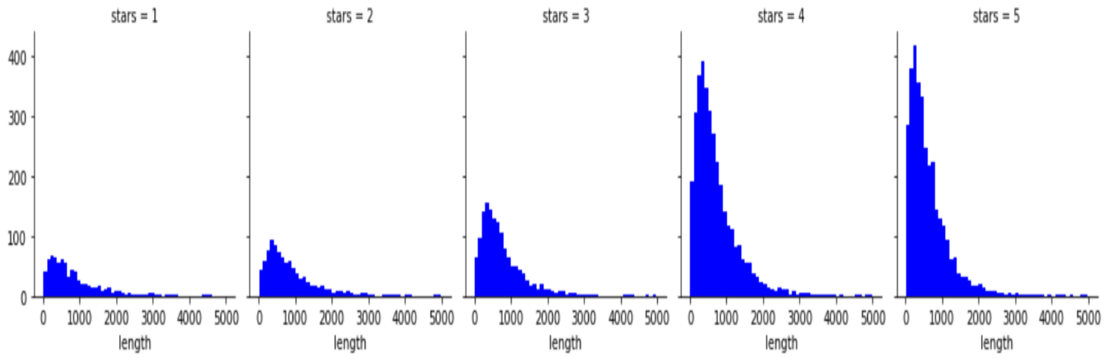
```
[ ] #CREATING A NEW COLUMN IN THE DATASET FOR THE NUMBER OF WORDS IN THE REVIEW
    data['length'] = data['text'].apply(len)
    data.head()
```

	business_id	date	review_id	stars	text	type	user_id	cool	useful	funny	length
0	9yKzy9PApeiPPOUJEtnvkg	2011-01-26	fWKvX83p0-ka4JS3dc6E5A	5	My wife took me here on my birthday for breakf...	review	rLtl8ZkDX5vH5nAx9C3q5Q	2	5	0	889
1	ZRJwVLyzEJq1VAihDhYiow	2011-07-27	IjZ33sJrzXqU-0X6U8NwyA	5	I have no idea why some people give bad review...	review	0a2KyEL0d3Yb1V6aivbIuQ	0	0	0	1345
2	6oRAC4uyJCsJl1X0WZpVSA	2012-06-14	IESLBzqUCLdSzSqm0eCSxQ	4	love the gyro plate. Rice is so good and I als...	review	0hT2KtfLiobPvh6cDC8JQg	0	1	0	76
3	_1QQZuf4zZOyFCvXc0o6Vg	2010-05-27	G-WvGaISbqqaMHlNnByodA	5	Rosie, Dakota, and I LOVE Chaparral Dog Park!!...	review	uZetl9T0NcROGOyFfughhg	1	2	0	419
4	6ozycU1RpktNG2-1BroVtw	2012-01-05	1uJFq2r5QfJG_6ExMRCaGw	5	General Manager Scott Petello is a good egg!!!...	review	vYmM4KTsC8ZfQBg-j5MWkw	0	0	0	469

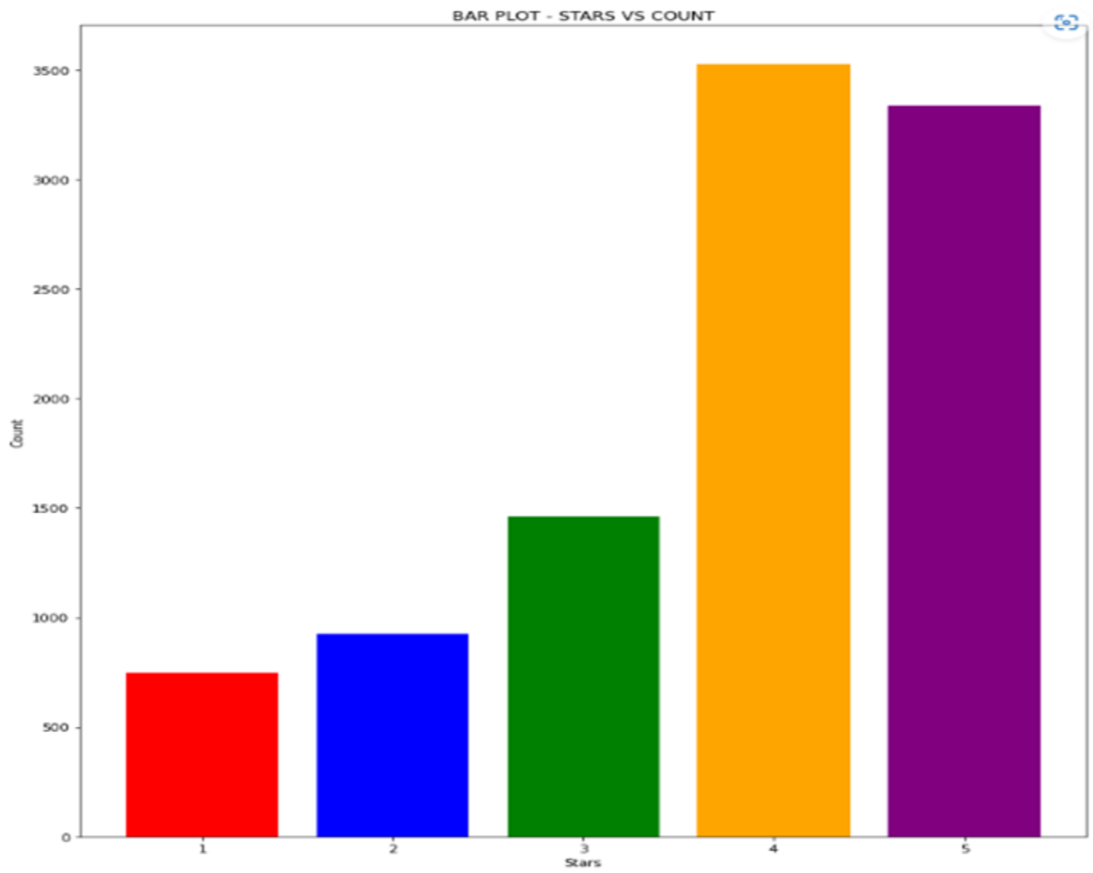
We carried out some visualization methods to better understanding of data.

```
In [28]:
# COMPARING TEXT LENGTH TO STARS
graph = sns.FacetGrid(data=data,col='stars')
graph.map(plt.hist,'length',bins=50,color='blue')
```

```
Out[28]: <seaborn.axisgrid.FacetGrid at 0x7f6c10c03f10>
```

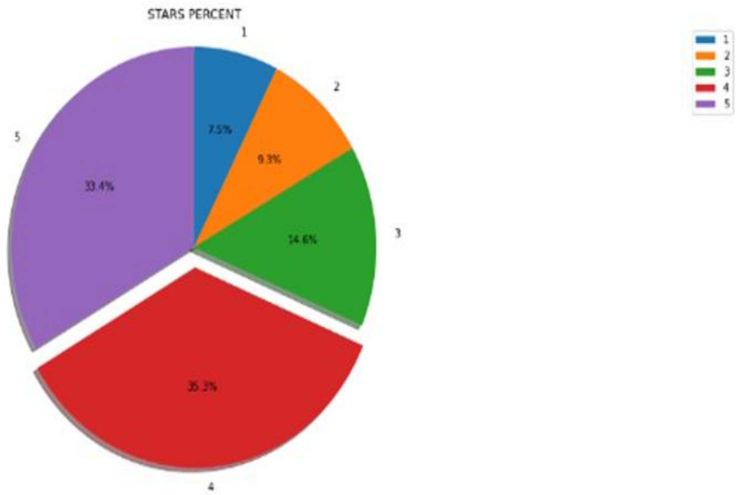



```
In [29]:
# BAR PLOT - STARS VS COUNT
x = [ 1, 2, 3, 4, 5]
y = [ 749, 927, 1461, 3526, 3337]
colors = ['red', 'blue', 'green', 'orange', 'purple']

fig, ax = plt.subplots(figsize=(12, 15))

plt.bar(x, y, color=colors)

# ADDING LABLES AND TITLES
plt.xlabel('Stars')
plt.ylabel('Count')
plt.title('BAR PLOT - STARS VS COUNT')

plt.show()

```



In [30]:

```python
# PIE CHART - EACH STAR PERCENT
Stars = [ 1, 2, 3, 4, 5]
Count = y = [ 749, 927, 1461, 3526, 3337]
myexplode = [0, 0, 0, 0.1, 0]

plt.figure(figsize=(20, 8))
plt.pie(Count, labels=Stars, autopct='%1.1f%%', shadow=True, startangle=90, explode=myexplode, counterclock=False)

plt.legend(labels=Stars, loc=1)
plt.axis('equal')
plt.title('STARS PERCENT')

# Show the plot
plt.show()
```



Creating a bar plot for average ratings vs month and average words vs stars. We have found out the mean value (stval) of the vote columns w.r.t the stars on the review and also the correlation (corr) between the vote columns.

(6). Mean Value of the Vote columns

```
[35] # GETTING THE MEAN VALUES OF THE VOTE COLUMNS WRT THE STARS ON THE REVIEW
     stval = data.groupby('stars').mean()
     stval
```

stars	cool	useful	funny	length
1	0.576769	1.604806	1.056075	826.515354
2	0.719525	1.563107	0.875944	842.256742
3	0.788501	1.306639	0.694730	758.498289
4	0.954623	1.395916	0.670448	712.923142
5	0.944261	1.381780	0.608631	624.999101

(7). Correlation between the voting columns:

```
# FINDING THE CORRELATION BETWEEN THE VOTE COLUMNS
corr = stval.corr()
corr
```

	cool	useful	funny	length
cool	1.000000	-0.743329	-0.944939	-0.857664
useful	-0.743329	1.000000	0.894506	0.699881
funny	-0.944939	0.894506	1.000000	0.843461
length	-0.857664	0.699881	0.843461	1.000000

B. Feature Extraction:

Classifying the dataset and splitting it into the reviews and stars.

In [38]:

```
# CLASSIFICATION
data_classes = data[(data['stars']==1) | (data['stars']==3) | (data['stars']==5)]
data_classes.head()
```

Out[38]:

	business_id	date	review_id	stars	text	type	user_id	cool	useful	funny	length
0	9yKzy9PApeiPPOUJEtnvkg	2011-01-26	fWKvX83p0-ka4JS3dc6E5A	5	My wife took me here on my birthday for breakf...	review	rLtl8ZkDX5vH5nAx9C3q5Q	2	5	0	889
1	ZRJwVLyzEJq1VAihDhYiow	2011-07-27	IjZ33sJrzXqU-0X6U8NwyA	5	I have no idea why some people give bad review...	review	0a2KyEL0d3Yb1V6aivbIuQ	0	0	0	1345
3	_1QQZuf4zZOyFCvXc0o6Vg	2010-05-27	G-WvGaISbqqaMHlNnByodA	5	Rosie, Dakota, and I LOVE Chaparral Dog Park!!...	review	uZetl9T0NcROGOyFfughhg	1	2	0	419
4	6ozycU1RpktNG2-1BroVtw	2012-01-05	1uJFq2r5QfJG_6ExMRCaGw	5	General Manager Scott Petello is a good egg!!!...	review	vYmM4KTsC8ZfQBg-j5MWkw	0	0	0	469
6	zp713qNhx8d9KCJJnrw1xA	2010-02-12	riFQ3vxNpP4rWLk_CSri2A	5	Drop what you're doing and drive here. After I...	review	wFweIWhv2fREZV_dYkz_1g	7	7	4	1565

Yelp allows users to write text reviews in free form. This means that a user may excessively use capital letters and punctuation marks (to express his/her intense dislike, for example) and slang words within a review. Moreover, stop words, like 'the', 'that', 'is' etc, occur frequently across reviews and are not very useful.

Therefore, it is necessary to pre-process the reviews in order to extract meaningful content from each of them. To do this, we use standard Python libraries to remove capitalizations, stop words and punctuations.

[ ]

```
# CLEANING THE REVIEWS - REMOVAL OF STOPWORDS AND PUNCTUATION
def text_process(text):
    nopunc = [char for char in text if char not in string.punctuation]
    nopunc = ''.join(nopunc)
    return [word for word in nopunc.split() if word.lower() not in stopwords.words('english')]
```

Converting the text data into vectors by vectorization.

```
[ ] # CONVERTING THE WORDS INTO A VECTOR
    vocab = CountVectorizer(analyzer=text_process).fit(x)
    print(len(vocab.vocabulary_))

    31336
```

Applying fit transform.

```
# Testing review
r0 = x[0]
print(r0)

My wife took me here on my birthday for breakfast a

Do yourself a favor and get their Bloody Mary.  It

While EVERYTHING on the menu looks excellent, I had

Anyway, I can't wait to go back!
```

```
[ ] # Transforming
    vocab0 = vocab.transform([r0])
    print(vocab0)

      (0, 292)      1
      (0, 1213)     1
      (0, 1811)     1
      (0, 3537)     1
      (0, 5139)     1
      (0, 5256)     2
```

Getting featured words back.

```
# Getting feature words
"""
    Now the words in the review number 78 have been converted into a vector.
    The data that we can see is the transformed words.
    If we now get the feature's name - we can get the word back!
"""
print("Getting the words back:")
print(vocab.get_feature_names_out()[11128])
print(vocab.get_feature_names_out()[24544])
```

```
Getting the words back:
amazing
pretty
```

Vectorization of the whole review set and checking the sparse matrix.

```
x = vocab.transform(x)
# Shape of the matrix:
print("Shape of the sparse matrix: ", x.shape)

#Non-zero occurences:
print("Non-Zero occurences: ",x.nnz)

# DENSITY OF THE MATRIX
density = (x.nnz/(x.shape[0]*x.shape[1]))*100
print("Density of the matrix = ",density)
```

```
Shape of the sparse matrix:  (5547, 31336)
Non-Zero occurences:  312457
Density of the matrix =  0.17975812697942373
```

C. Models Used:

a) K nearest neighbour

b) Decision Tree

c) Random Forest

d) Support Vector Machine (SVM)

e) Multinomial Naïve Bayes

f) Multilayer perceptron classifier

1. K Neighbours Classifier -

The K-Nearest Neighbours algorithm was applied to the Yelp dataset for classification. The Yelp dataset consists of text reviews from customers and their corresponding ratings. The objective was to classify the reviews into positive or negative sentiment.

The KNeighborsClassifier object was created with 10 neighbours, and the model was trained using the training dataset. The predicted target values for the test dataset were obtained using the predict method and stored in predknn.

Overall, the K-Nearest Neighbours algorithm was effective in classifying the sentiment of Yelp reviews, as evidenced by the accuracy score and classification report. This approach can be useful for businesses to monitor their online reputation by analysing customer reviews.

2. Decision Tree Classifier -

The Decision Tree algorithm was applied to the Yelp dataset to classify reviews into positive or negative sentiment. The DecisionTreeClassifier object was created, and the model was trained using the training dataset. The predicted target values for the test dataset were obtained using the predict method and stored in preddt.

The Decision Tree algorithm is particularly useful for the Yelp dataset, as it allows businesses to identify key factors that influence customer sentiment. For example, a decision tree can identify which words or phrases are commonly used in positive or negative reviews, allowing businesses to tailor their products or services accordingly.

In conclusion, the Decision Tree algorithm was effective in classifying Yelp reviews into positive or negative sentiment, as evidenced by the accuracy score and classification report. This approach can provide valuable insights for businesses looking to improve their online reputation and customer satisfaction.

3. Random Forest Classifier -

Random Forest is a type of supervised learning algorithm that is based on decision trees and can be trained on a set of labelled data to identify patterns and relationships in the data. It can handle high-dimensional data and identify complex patterns and relationships in the data.

Secondly, it can handle missing data and noisy data, making it suitable for real-world scenarios where data is often incomplete or inaccurate.

The Random Forest algorithm is particularly useful for the Yelp dataset, as it combines multiple decision trees to provide more accurate and stable predictions. This approach can provide valuable insights for businesses looking to monitor their online reputation and customer satisfaction.

4. Support Vector Machine (SVM) -

The SVM algorithm is particularly useful for the Yelp dataset, as it can handle both linear and non-linear classification problems by finding the best separating hyperplane in a high-dimensional space.

This approach can provide valuable insights for businesses looking to monitor their online reputation and customer satisfaction.

In conclusion, the SVM algorithm was effective in classifying Yelp reviews into positive or negative sentiment, as evidenced by the accuracy score and classification report. This approach can provide valuable insights for businesses looking to improve their online reputation and customer satisfaction.

5. Multinomial Naive Bayes -

The Multinomial Naive Bayes algorithm is particularly useful for text classification problems such as sentiment analysis, as it can handle discrete data such as word counts. This approach can provide valuable insights for businesses looking to monitor their online reputation and customer satisfaction.

In conclusion, the Multinomial Naive Bayes algorithm was effective in classifying Yelp reviews into positive or negative sentiment, as evidenced by the accuracy score and confusion matrix. This approach can provide valuable insights for businesses looking to improve their online reputation and customer satisfaction.

6. Multilayer Perceptron Classifier -

The Multilayer Perceptron (MLP) Classifier is a type of neural network that was applied to the Yelp dataset to classify reviews into positive or negative sentiment. The MLPClassifier object was created, and the model was trained using the training dataset. The predicted target values for the test dataset were obtained using the predict method and stored in predmlp.

The MLP Classifier is a powerful algorithm for solving complex classification problems. It can learn non-linear relationships between input and output variables and is capable of handling high-dimensional datasets with a large number of features. In the context of the Yelp dataset, the MLP Classifier can help businesses gain insights into their customer sentiment and improve their online reputation.

In conclusion, the MLP Classifier was effective in classifying Yelp reviews into positive or negative sentiment, as evidenced by the accuracy score and confusion matrix. This approach can provide valuable insights for businesses looking to monitor their online reputation and customer satisfaction.

Discussion and Results

A. Output obtained:

1.Sentiment Classification:

We successfully classified the sentiments of the customers as per star ratings given by them. Results as shown in figure.

```
[ ] star_1=data_classes[(data_classes.stars==1)]
    star_1
```

	business_id	date	review_id	stars	text	type	user_id	cool	useful	funny	length
23	lJ0o6b8bJFAbG6MjGfBebQ	2010-09-05	Dx9sfFU6Zn0GYOckijom-g	1	U can go there n check the car out. If u wanna...	review	zRIQEDYd_HKp0VS3hnAffA	0	1	1	594
31	vvA3fbps4F9nGlAEYKk_sA	2012-05-04	S9OVpXat8k5YwWCn6FAgXg	1	Disgusting! Had a Groupon so my daughter and ...	review	8AMn6644NmBf96xGO3w6OA	0	1	0	361
35	o1GIYYZJjM6nM03fQs_uEQ	2011-11-30	ApKbwpYJdnhhgP4NbjQw2Q	1	I've eaten here many times, but none as bad as...	review	iwUN95LIaEr75TZE_JC6bg	0	4	3	1198
61	l4vBbCL9QbGiwLuLKwD_bA	2011-11-22	DJVxOfj2Rw9zklC9tU3i1w	1	I have always been a fan of Burlington's deals...	review	EPROVap0M19Y6_4uf3eCmQ	0	0	0	569
64	CEswyP-9SsXRNLR9fFGKKw	2012-05-19	GXj4PNAi095-q9ynPYH3kg	1	Another night meeting friends here. I have to...	review	MjLAe48XNfYITeFYca5gMw	0	1	2	498

```
[ ] star_3=data_classes[(data_classes.stars==3)]
    star_3
```

	business_id	date	review_id	stars	text	type	user_id	cool	useful	funny	length
16	supigcPNO9IKo6olaTNV-g	2008-10-12	HXP_0Ul-FCmA4f-k9CqvaQ	3	We went here on a Saturday afternoon and this ...	review	SBbftLzfYYKItOMFwOTIJg	3	4	2	1469
18	b5cEoKR8iQliq-yT2_O0LQ	2009-03-06	v0cTd3PNpYCkTyGKSpOfGA	3	I met a friend for lunch yesterday. \n\nLoved ...	review	UsULgP4bKA8RMzs8dQzcsA	5	6	4	1161
20	8FNO4D3eozplij0k3q5Zbg	2008-10-08	MuqugTuR5DdIPcZ2IVP3aQ	3	DVAP....\n\nYou have to go at least once in yo...	review	C6lOtaaYdLIT5fWd7ZYIuA	2	4	1	565
34	3oLy0rlzRI_xiqfQHqC4_g	2011-03-27	Bk7F8lyBuOHVp6w3BAKVow	3	There's two ways to look at this place. One is...	review	1guJDgUfTiDdbgKqBhsZFQ	1	3	1	610
45	qB-qsaSnhbHCt18_AN4Quw	2011-12-21	1FvrC35rTJ6BWFvRog7tuA	3	Everything was nice. The ice cream was delicio...	review	66PQJEHC0tCWGMI4V9KT-Q	0	0	0	243
...	...	...	...	...	...	...	...	...	...	...	...

```
star_5=data_classes[(data_classes.stars==5)]
star_5
```

	business_id	date	review_id	stars	text	type	user_id	cool	useful	funny	length
0	9yKzy9PApeiPPOUJEtnvkg	2011-01-26	fWKvX83p0-ka4JS3dc6E5A	5	My wife took me here on my birthday for breakf...	review	rLtl8ZkDX5vH5nAx9C3q5Q	2	5	0	889
1	ZRJwVLyzEJq1VAihDhYiow	2011-07-27	IjZ33sJrzXqU-0X6U8NwyA	5	I have no idea why some people give bad review...	review	0a2KyEL0d3Yb1V6aivbIuQ	0	0	0	1345
3	_1QQZuf4zZOyFCvXc0o6Vg	2010-05-27	G-WvGaISbqqaMHINnByodA	5	Rosie, Dakota, and I LOVE Chaparral Dog Park!!...	review	uZetl9T0NcROGOyFfughhg	1	2	0	419
4	6ozycU1RpktNG2-1BroVtw	2012-01-05	1uJFq2r5QfJG_6ExMRCaGw	5	General Manager Scott Petello is a good egg!!!...	review	vYmM4KTsC8ZfQBg-j5MWkw	0	0	0	469
6	zp713qNhx8d9KCJJnrw1xA	2010-02-12	riFQ3vxNpP4rWLk_CSri2A	5	Drop what you're doing and drive here. After I...	review	wFweIWhv2fREZV_dYkz_1g	7	7	4	1565
...	...	...	...	...	...	...	...	...	...	...	...

2. Rating Prediction:

The model successfully predicted the ratings of the customers as per provided text comment input. Results as shown in figure.

(15). Rating Prediction on basis of review text.

```
[ ]  # POSITIVE REVIEW
     pr = data['text'][9999]
     print(pr)
     print("Actual Rating: ",data['stars'][9999])
     pr_t = vocab.transform([pr])
     print("Predicted Rating:")
     mlp.predict(pr_t)[0]
```

```
4-5 locations.. all 4.5 star average.. I think Arizona really has some
Actual Rating:  5
Predicted Rating:
5
```

```
# AVERAGE REVIEW
ar = data['text'][9995]
print(ar)
print("Actual Rating: ",data['stars'][9995])
ar_t = vocab.transform([ar])
print("Predicted Rating:")
mlp.predict(ar_t)[0]
```

```
First visit...Had lunch here today - used my Groupon.

We ordered the Bruschetta, Pretzels and Steak & Cheese

-We both thought there was WAY too much Balsamic used.

-We tried the butter and salt pretzel & cinnamon sugar

-The calzone was good. We liked the dough and it was f

Overall, we thought it was average as far as the food 

We have another Groupon to use so maybe we'll try a pi
Actual Rating:  3
Predicted Rating:
3
```

```
# NEGATIVE REVIEW
nr = data['text'][9987]
print(nr)
print("Actual Rating: ",data['stars'][9987])
nr_t = vocab.transform([nr])
print("Predicted Rating:")
mlp.predict(nr_t)[0]
```

```
The food is delicious.  The service:  discriminatory.
Actual Rating:  1
Predicted Rating:
1
```

B. Evaluation measures:

1. Precision
2. Recall
3. F1 Score
4. Accuracy

1. **Precision -**

Precision is a term commonly used in statistics and machine learning to measure the exactness or accuracy of a measurement or prediction. It is defined as the ratio of true positives (correctly identified positives) to the total number of positive predictions, which includes both true positives and false positives (incorrectly identified positives). In other words, precision measures how often a model's positive predictions are correct.

2. **Recall -**

Recall is a term commonly used in statistics and machine learning to measure the completeness or sensitivity of a measurement or prediction. It is defined as the ratio of true positives (correctly identified positives) to the total number of actual positive cases, which includes both true positives and false negatives (incorrectly identified negatives). In other words, recall measures how many of the actual positives a model correctly identifies.

3. F1 score -

F1 score is a commonly used metric in statistics and machine learning that combines both precision and recall into a single score. It is the harmonic mean of precision and recall, and is calculated as follows:

F1 score = 2 * (precision * recall) / (precision + recall)

4. Accuracy -

Accuracy is a metric used in model evaluation to measure the proportion of correct predictions made by the model on a given dataset. It is defined as the ratio of the number of correct predictions made by the model to the total number of predictions made.

In other words, accuracy tells us how well the model is able to correctly classify instances in the dataset. A higher accuracy score indicates that the model is able to make more correct predictions, while a lower accuracy score indicates that the model is making more incorrect predictions.

While accuracy is an important metric, it should be used in conjunction with other evaluation measures such as precision, recall, and F1 score to get a more complete picture of the model's performance. Additionally, accuracy may not always be the best metric to use in certain cases, such as when the dataset is imbalanced or when the costs of false positives and false negatives are significantly different.

```
Confusion Matrix for Multilayer Perceptron Classifier:
[[ 95  33  34]
 [ 20 185  87]
 [ 14  62 580]]
Score: 77.48
Classification Report:
              precision    recall  f1-score   support

           1       0.74      0.59      0.65       162
           3       0.66      0.63      0.65       292
           5       0.83      0.88      0.85       656

    accuracy                           0.77      1110
   macro avg       0.74      0.70      0.72      1110
weighted avg       0.77      0.77      0.77      1110
```

Conclusion

We have implemented and experimented with several classification algorithms to predict star rating from review text, which gives a good result. We used count vectorizer as feature extractors. We have also predicted the customer's star ratings for restaurants using all the past reviews given by other customers and this customer predicting model.

In conclusion, the sentiment analysis of the Yelp dataset using machine learning algorithms was successfully conducted. The project involved analysing customer reviews from Yelp and predicting the sentiment of the reviews as either positive or negative.

We used various machine learning algorithms, including Multinomial Naive Bayes, Random Forest, Decision Tree, Support Vector Machine (SVM), K Nearest Neighbour and Multilayer Perceptron (MLP) to train and evaluate our models. We also performed data cleaning, pre-processing, and feature extraction to prepare the data for the machine learning models.

Based on our evaluation metrics, MLP performed the best with an accuracy of 77.48%. This indicates that our model is effective in predicting the sentiment of Yelp reviews. However, there is always room for improvement, and further research could be conducted to improve the accuracy of the model.

The results of this project can be applied in various industries, including marketing and customer service. Companies can use this type of analysis to better understand customer feedback and improve their products and services accordingly.

Overall, the sentiment analysis of the Yelp dataset using machine learning algorithms proved to be a valuable exercise in understanding the power of machine learning in analysing large amounts of data and extracting meaningful insights.

References

1. Sunmin Lee “Sentiment Analysis Using BERT on Yelp Restaurant Reviews” Department of Computer and Information Technology West Lafayette, Indiana August 2022.

2. Boya Yu, Jiaxu Zhou, Yi Zhang, Yunong Cao “Identifying Restaurant Features via Sentiment Analysis on Yelp Reviews” Center for Urban Science & Progress New York University, New York, The United States.

3. Anish A. Parikh,1 Carl Behnke, 2 Doug Nelson, 2 Mihaela Vorvoreanu,3 and Barbara Almanza2 “A Qualitative Assessment of Yelp.Com Users’ Motivations to Submit and Read Restaurant Reviews” 1Department of Management, Montclair State University, Montclair, New Jersey, USA 2School of Hospitality and Tourism Management, Purdue University, West Lafayette,Indiana,USA 3Department of Communication, Purdue University, West Lafayette, Indiana, USA

Dataset Reference:

https://www.kaggle.com/code/omkarsabnis/sentiment-analysis-on-the-yelp-reviews-dataset/input