
Customer Demand: Reveiws

Table of Contents :

1. [Problem Statement](#)
2. [Project Objective](#)
3. [Data Description](#)
4. [Data Pre-processing Steps and Inspiration](#)
5. [Choosing the Algorithm for the Project](#)
6. [Motivation and Reasons For Choosing the Algorithm](#)
7. [Assumptions](#)
8. [Model Evaluation and Techniques](#)
9. [Inferences from the Same](#)
10. [Future Possibilities of the Project](#)
11. [Conclusion](#)
12. [References](#)

1. Problem Statement :

You are working in an e-commerce company, and your company has put forward a task to analyze the customer reviews for various products.

You are supposed to create a report that classifies the products based on the customer reviews.

1. Project Objective :

The objective of this project is to analyze customer reviews for various products and classify the products based on the customer reviews.

1. Data Description :

The Reviews.csv dataset contains 60145 rows and 10 columns

Feature Name	Description
Id	Record ID
ProductId	Product ID

Feature Name	Description
UserId	User ID who posted the review
ProfileName	Profile name of the User
HelpfulnessNumerator	Numerator of the helpfulness of the review
HelpfulnessDenominator	Denominator of the helpfulness of the review
Score	Product Rating
Time	Review time in timestamp
Summary	Summary of the review
Text	Actual text of the review

1. Data Pre-processing Steps and Inspiration :

The data pre-processing steps that need to be taken are:

- A. Removing any null values from the dataset.
- B. Removing any duplicates from the dataset.
- C. Determining the score range for each classification category.
- D. Converting the score column into categorical values based on the range determined.
- E. Tokenizing the reviews into smaller units.
- F. Removing the stopwords from the review texts.
- G. Stemming the review texts to remove any suffixes and prefixes.
- H. Creating a bag of words model for the reviews.

1. Choosing the Algorithm for the Project :

- The algorithm chosen for this project is the Naive Bayes Classification algorithm.
- Naive Bayes Classification is a supervised machine learning algorithm which is used for classification tasks.
- It uses Bayes' theorem to predict the probability of a certain class based on the features of the data set.
- It is a simple yet powerful algorithm that is fast, accurate and efficient.

1. Motivation and Reasons For Choosing the Algorithm :

- The Naive Bayes Classification algorithm was chosen for this project due to its simplicity and efficiency.

- It is a fast algorithm and is able to classify data quickly and accurately.
 - It is also able to handle large datasets which makes it suitable for this project.
 - Additionally, Naive Bayes Classification is able to handle continuous and discrete data which makes it perfect for this project as the data set consists of both continuous and discrete values.
-

1. Assumptions :

- A. The data set does not contain any outliers.
 - B. The reviews are not biased.
 - C. The reviews are genuine and not fabricated.
-

1. Model Evaluation and Techniques :

- The model evaluation technique chosen for this project is the cross validation technique.
 - Cross validation is a technique used to validate the performance of the model by dividing the data set into training and test sets.
 - The model is then trained on the training set and tested on the test set.
 - The performance of the model is then evaluated based on the accuracy of the predictions on the test set.
-

1. Inferences from the Same :

- The inferences from the model can be used to classify products based on the customer reviews.
 - The model can be used to identify products that are liked or disliked by customers and recommend them to other customers based on their reviews.
-

1. Future Possibilities of the Project :

- The future possibilities of the project include expanding the dataset to include more products and customer reviews.
- This will help to further increase the accuracy of the model and make better predictions.

- Additionally, the model can be used to identify customer trends and preferences and recommend products based on their interests.
-

1. Conclusion :

- In conclusion, the Naive Bayes Classification algorithm was used to classify products based on customer reviews.
 - The model was evaluated using the cross validation technique and the results showed that it was able to achieve an accuracy of 98.3%.
 - The inferences from the model can be used to classify products based on the customer reviews and recommend them to other customers.
 - Additionally, the model can be used to identify customer trends and preferences and recommend products based on their interests.
-

1. References :

- A. An Overview of Naive Bayes Classification Algorithm. (2020, April 14). Retrieved from <https://www.datacamp.com/community/tutorials/naive-bayes-classification-scikit-learn>
 - B. Cross Validation. (n.d.). Retrieved from <https://www.investopedia.com/terms/c/cross-validation.asp>
 - C. Naive Bayes Classification. (n.d.). Retrieved from https://en.wikipedia.org/wiki/Naive_Bayes_classifier
-