Group: 14
Group members:

Abhilash Rao Pallineni 0735412
Ketana Pereddigari 0734184
Amandeep kaur 0734293

Assessment 2

RECOMMENDATION OF WEBSITES BASED ON REVIEWS

Description

Reviews are the prior resources to provide information about the products or services based on users experience to consumers, sellers, and retailers. Most of the customer rely on reviews to make any decision related to purchases or any service utilization provided. Some of the reviews are given by non-users to influence the customer's decisions. In process of performing the project we would like to transform the business problem statement to a data analytics problem this will help us to work more efficiently to achieve the desired outcome. For better understanding of this translation we would like to express the functionality using reverse engineering step by step.

As a result, recommendation of websites based on user reviews and ratings could be achieved by performing sentimental analysis (Natural Language Processing) over text data to classify the information into three categories positive, negative and neutral feedback. For sentimental analysis we will apply several algorithms such as **SVM, Naive Bayes, Vector T-SNE, KNN, Latent Dirichlet allocation (LDA), maximum entropy** and other models which comes under supervised learning.

Before classifying we need to make sure that the data used should be taken which are given by customer who has purchased/used the product/services this can be achieved while collecting the data from e-commerce sites. Here analytics will play a major role for computational analysis to classify the users as purchaser or non-purchaser using the internal data from websites.

In order to suggest the best websites, we need to Collect reviews, ratings, product details and other related information from different websites. This can be done in different ways we would like to use **web scrappy tool (Parse hub application and other methods)** so that it can be beneficial to our learning curve and help in gathering only the required information for the analysis. **Data processing** will be the main task for our data analytics project using excel and python coding. Our main consideration is to remove reviews of non-purchasers and concentrate on purchased user reviews as the most focusing aspect of this project.

Group: 14
Group members:

Abhilash Rao Pallineni 0735412
Ketana Pereddigari 0734184
Amandeep kaur 0734293

Outcomes Assessed

In this analysis we will filter the reviews into buyers and non-buyers of any product/service. The categorization of the reviews by using any classifier we can get most preferable by evaluating the performance of models with cross validation, confusion matrix and RMSE. With SVM model we can analyze the reviews given as genuine or not. By exploration analysis on the data we can compare the reviews into positive, negative and neutral. With the traits of reviews and filtering out the non- buyers reviews we will be providing genuine information to the customers regarding the quality of recommendation list.