**Real-Time-Recommendation-System Using content-based filtering.**

**Abstract**

In this project, we were given to Design a real-time recommendation system for an e-commerce platform. The system should provide product recommendations based on the user’s recent browsing or purchase history using collaborative filtering or content-based filtering.

The followings steps will cover the solution of the Problem :-

**STEPS :-**

1. **Data Collection:- In this project I have used a Walmart Dataset from kaggle for building this project.**

**Dataset Link : -** [**https://www.kaggle.com/datasets/promptcloud/walmart-product-review-dataset**](https://www.kaggle.com/datasets/promptcloud/walmart-product-review-dataset)

**Also I have mentioned and uploaded in github**

[**https://github.com/AliptoChoudhury/Real-Time-Recommendation-System/blob/main/marketing\_sample\_for\_walmart\_com-walmart\_com\_product\_review\_\_20200701\_20201231\_\_5k\_data.tsv**](https://github.com/AliptoChoudhury/Real-Time-Recommendation-System/blob/main/marketing_sample_for_walmart_com-walmart_com_product_review__20200701_20201231__5k_data.tsv)

1. **Data Arranging:- So there are many columns which we will not use to training the model. This below mentioned columns will be used for this model from the dataset**

'Uniq Id','Product Id', 'Product Rating', 'Product Reviews Count', 'Product Category', 'Product Brand', 'Product Name', 'Product Image Url', 'Product Description', 'Product Tags'

# Data Cleaning and Preprocessing :- For Data Cleaning and we need to make tags we will import spacy library. It is a NLP library. From this library we will use lang.en.stop\_words sub library and will import STOP\_WORDS. It will remove (is,or,from) which we will have no use will training the model. So, it will remove it. We will apply a English model which is present in the library called “en\_core\_web\_sm”. This English model we will apply in three Columns “Category”, “Brand” and “Description”.

1. **Model Selection : - we will be using content-based filtering to make this Recommendation System.**

**Content Base Recommendation System suggest items to users based on the attributes or features of the items and the user’s previous preferences.**

**TF-IDF Vectorizer :- It helps in transforming text data into numerical data, which is useful for training the model.**

**Cosine Similarity – It is a metric used to measure how similar two vectors are. Two vectors similarity score is how much will decide how much both the vectors are similar.**

**The Full Code Implementation is uploaded in github:-**

[**https://github.com/AliptoChoudhury/Real-Time-Recommendation-System**](https://github.com/AliptoChoudhury/Real-Time-Recommendation-System)