PRODUCT RECOMMENDATION

A MINI PROJECT REPORT SUBMITTED BY

Geshal Gerald Mendonca
4NM20CS074
VI Semester, B Section

Harshitha J
4NM20CS077
VI Semester, B section

UNDER THE GUIDANCE OF

MS. Vaishali Bangera
Assistant professor GD I
Department of Computer Science and Engineering

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

Bachelor of Engineering in Computer Science & Engineering

from

Visvesvaraya Technological University, Belagavi

N.M.A.M. INSTITUTE OF TECHNOLOGY
(An Autonomous Institution under VTU, Belgaum)
AICTE approved, (ISO 9001:2015 Certified), Accredited with 'A' Grade by NAAC
NITTE -574 110, Udupi District, KARNATAKA.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

B.E. CSE Program Accredited by NBA, New Delhi from 1-7-2018 to 30-6-2021

May 2023

NITTE EDUCATION TRUST | N.M.A.M. INSTITUTE OF TECHNOLOGY
(An Autonomous Institution affiliated to Visvesvaraya Technological University, Belagavi)
Nitte – 574 110, Karnataka, India
(ISO 9001:2015 Certified), Accredited with 'A' Grade by NAAC
☎: 08258 - 281039 – 281263, Fax: 08258 – 281265

Department of Computer Science and Engineering
B.E. CSE Program Accredited by NBA, New Delhi from 1-7-2018 to 30-6-2021

CERTIFICATE

"Product recommendation system" is a bonafide work carried out by Geshal Gerald Mendonca (4NM20CS074) and Harshitha J (4NM20CS077) in partial fulfilment of the requirements for the award of Bachelor of Engineering Degree in Computer Science and Engineering prescribedby Visvesvaraya Technological University, Belagavi during the year 2023-2024.

It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report. The Mini project report has been approved as it satisfies the academic requirements in respect of the project work prescribed for the Bachelor of Engineering Degree.

Signature of Guide

Signature of HOD

ABSTRACT

Product recommendation using machine learning is a technique used to suggest products to customers based on their past behavior, preferences, and characteristics. This method utilizes algorithms such as collaborative filtering, content-based filtering, and hybrid methods to analyze the customer's data and generate personalized recommendations. The goal is to provide relevant and accurate product suggestions to the customer, resulting in improved user experience, increased engagement, and higher sales conversion rates. This abstract provides an overview of the concept of product recommendation using machine learning and highlights its benefits in enhancing customer satisfaction and loyalty.

In addition to that, the ML algorithms use a variety of data sources, including customer browsing history, purchase history, demographics, and other factors, to generate personalized product recommendations. This approach not only helps businesses increase their revenue but also improves the overall customer experience by providing relevant and useful product suggestions. Overall, ML-based product recommendation is a powerful tool for businesses to enhance customer satisfaction and increase their bottom line.

Table of Contents

INTRODUCTION

Product recommendation is a technique used by businesses to suggest products to their customers based on their past behavior, interests, and other relevant data. This technique has been widely adopted in e-commerce, social media, and other online marketplaces to enhance customer experience, increase sales, and improve customer retention. However, with the growing amount of data generated by businesses, manual product recommendation is no longer feasible. This is where machine learning comes into play.

Machine learning is a subset of artificial intelligence that involves training algorithms to learn from data and make predictions or decisions without being explicitly programmed. In the context of product recommendation, machine learning algorithms can analyze large amounts of customer data and generate personalized recommendations based on the individual preferences and behavior of each customer.

There are various machine learning algorithms that can be used for product recommendation. The choice of algorithm depends on the type and size of data, the complexity of the recommendation task, and the desired performance metrics. Here are some of the most commonly used algorithms for product recommendation:

1. **Collaborative Filtering:** Collaborative Filtering is a widely used algorithm in recommendation systems. It uses past user behaviour and preferences to recommend products to users with similar behaviour and preferences. Collaborative Filtering can be divided into two categories: user-based and item-based. User-based collaborative filtering recommends products based on the preferences of similar users, while item-based collaborative filtering recommends products based on the similarity of products.
2. **K-means:** To use K-means for product recommendation, the algorithm can be applied to product attributes or user behaviour data to group similar items or users together. Once the clusters are generated, the algorithm can then recommend products to users based on the preferences of other users within the same cluster.

SOFTWARE REQUIREMENTS

- **Operating System :** Windows/Linux

- **Jupyter Notebook:**

- Jupyter Notebook is an open-source web application that allows users to create and share documents containing live code, equations, visualizations, and narrative text. The application was originally designed for the Python programming language, but it now supports many other programming languages, including R, Julia, and Scala.

- One of the key features of Jupyter Notebook is its ability to support interactive computing. Users can write code in cells, execute them, and view the results in real-time. This makes it an ideal platform for data analysis, scientific computing, and machine learning, where users need to explore and experiment with data.

- Jupyter Notebook also allows users to create rich, multimedia documents by combining code, text, and visualizations in a single document. These documents can be saved as notebooks and shared with others, making it easy to collaborate on projects or share findings with a wider audience.

- In addition, Jupyter Notebook is highly customizable, with a wide range of extensions and plugins available to enhance its functionality. Users can also run Jupyter Notebook on their local machine or in the cloud, making it a flexible and versatile platform for data analysis and scientific computing.

HARDWARE REQUIREMENTS

- **RAM :** Minimum 4GB (8GB recommended)
- **Processor :** Intel i3 or above
- **HDD :** 20 GB Disk space

DESIGN AND ANALYSIS

IMPLEMENTATION

- ***Installing Jupyter Notebook:***

  Install Python on your local machine from the official website. Open a command prompt or terminal and type "pip install jupyter".Launch Jupyter Notebook by typing "jupyter notebook" in the command prompt or terminal. Install additional packages using pip if needed, for example "pip install pandas"

- ***Importing libraries:***

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# %matplotlib inline
plt.style.use("ggplot")

import sklearn
from sklearn.decomposition import TruncatedSVD
```

- ***Loading the dataset:***

```
amazon_ratings = pd.read_csv('../input/amazon-ratings/ratings_Beauty.csv')
amazon_ratings = amazon_ratings.dropna()
amazon_ratings.head()
```

- ***Utility matrix:***

```
# Subset of Amazon Ratings

amazon_ratings1 = amazon_ratings.head(10000)
```

```
ratings_utility_matrix = amazon_ratings1.pivot_table(values='Rating', index='UserId', columns='P
roductId', fill_value=0)
ratings_utility_matrix.head()
```

- ***Decomposing the matrix:***

```
SVD = TruncatedSVD(n_components=10)
decomposed_matrix = SVD.fit_transform(X)
decomposed_matrix.shape
```

- ***Correlation matrix:***

```
correlation_matrix = np.corrcoef(decomposed_matrix)
correlation_matrix.shape
```

- ***Isolating product id from correlation matrix:***

```
X.index[99]
```

```
'6117036094'
```

```
i = "6117036094"

product_names = list(X.index)
product_ID = product_names.index(i)
product_ID
```

```
99
```

```
correlation_product_ID = correlation_matrix[product_ID]
correlation_product_ID.shape
```

```
(886,)
```

- ***Recommending top 10 highly correlated products in sequence:***

```
Recommend = list(X.index[correlation_product_ID > 0.90])

# Removes the item already bought by the customer
Recommend.remove(i)

Recommend[0:9]
```

```
['0733001998',
 '1304139212',
 '1304139220',
 '130414089X',
 '130414643X',
 '130414674X',
 '1304174778',
 '1304174867',
 '1304174905']
```

- ***Item to item based recommendation system based on product description:***

```
product_descriptions = pd.read_csv('../input/home-depot-product-search-relevance/product_descriptions.csv')
product_descriptions.shape

# Missing values

product_descriptions = product_descriptions.dropna()
product_descriptions.shape
product_descriptions.head()
```

```
product_descriptions1 = product_descriptions.head(500)
# product_descriptions1.iloc[:,1]

product_descriptions1["product_description"].head(10)
```

```
0    Not only do angles make joints stronger, they ...
1    BEHR Premium Textured DECKOVER is an innovativ...
2    Classic architecture meets contemporary design...
3    The Grape Solar 265-Watt Polycrystalline PV So...
4    Update your bathroom with the Delta Vero Singl...
5    Achieving delicious results is almost effortle...
6    The Quantum Adjustable 2-Light LED Black Emerg...
7    The Teks #10 x 1-1/2 in. Zinc-Plated Steel Was...
8    Get the House of Fara 3/4 in. x 3 in. x 8 ft. ...
9    Valley View Industries Metal Stakes (4-Pack) a...
Name: product_description, dtype: object
```

- ***Feature extraction from product descriptions:***

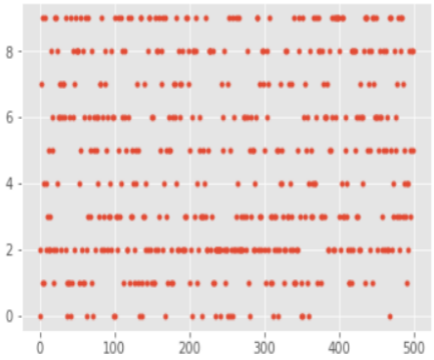
```
vectorizer = TfidfVectorizer(stop_words='english')
X1 = vectorizer.fit_transform(product_descriptions1["product_description"])
X1
```

- ***Visualizing product clusters in sequence of data:***

```
# Fitting K-Means to the dataset

X=X1

kmeans = KMeans(n_clusters = 10, init = 'k-means++')
y_kmeans = kmeans.fit_predict(X)
plt.plot(y_kmeans, ".")
plt.show()
```



RESULT

- ***Top words in each cluster based on product description:***

```
# # Optimal clusters is

true_k = 10

model = KMeans(n_clusters=true_k, init='k-means++', max_iter=100, n_init=1)
model.fit(X1)

print("Top terms per cluster:")
order_centroids = model.cluster_centers_.argsort()[:, ::-1]
terms = vectorizer.get_feature_names()
for i in range(true_k):
    print_cluster(i)
```

```
Top terms per cluster:
Cluster 0:
 concrete
 stake
 ft
 coating
 apply
 epoxy
 drying
 sq
 garage
 formula
Cluster 1:
 wood
 patio
 bamboo
 natural
 frame
 outdoor
 rug
 size
 steel
 dining
Cluster 2:
 used
 trim
 painted
 65
 proposition
 nbsp
 residents
 california
 project
 32
Cluster 3:
 door
 lbs
 easy
 dog
 nickel
 solid
 roof
 plastic
 house
 adjustable
Cluster 4:
 cutting
 saw
 tool
 blade
 design
 cut
 pliers
 grip
 metal
 non
Cluster 5:
 wall
```

```
 piece
 finish
 tile
 design
 use
 color
 easy
 installation
 water
Cluster 6:
 light
 watt
 bulb
 led
 fixture
 volt
 bulbs
 lighting
 use
 power
Cluster 7:
 helps
 water
 easy
 snow
 handle
 nozzle
 year
 features
 tool
 control
Cluster 8:
 air
 ft
 water
 unit
 room
 installation
 fan
 cooling
 use
 easy
Cluster 9:
 post
 fence
 gate
 ft
 screen
 vinyl
 posts
 aluminum
 brackets
 spline
```

- ***Predicting clusters based on key search words:***

```
def show_recommendations(product):
    #print("Cluster ID:")
    Y = vectorizer.transform([product])
    prediction = model.predict(Y)
    #print(prediction)
    print_cluster(prediction[0])
```

```
show_recommendations("cutting tool")
```

```
Cluster 4:
 cutting
 saw
 tool
 blade
 design
 cut
 pliers
 grip
 metal
 non
```

CONCLUSION

In conclusion, Product recommendation using machine learning is an effective technique for businesses to enhance customer experience, increase sales, and improve customer retention. By analyzing customer data, machine learning algorithms can generate personalized product recommendations that meet the individual needs of each customer. The benefits of using machine learning in product recommendation include enhancing customer experience, increasing sales, and saving time and resources.

The use of machine learning in product recommendation is becoming increasingly prevalent in e-commerce, social media, and other online marketplaces due to its proven effectiveness. Moreover, as machine learning algorithms continue to evolve, they are becoming more accurate and effective in identifying customer preferences and providing relevant product recommendations.

However, businesses must also be aware of the challenges associated with this technology, such as ensuring data privacy and security and the need for large amounts of data. Despite these challenges, machine learning-based product recommendation is a powerful tool for businesses that want to gain a competitive advantage by providing a better shopping experience for their customers, increasing sales, and improving their bottom line.

REFERENCES

- **Kaggle**
  https://www.kaggle.com/

- **Geeksforgeeks**
  https://www.geeksforgeeks.org/