Mohd-Talha-ZeoTap / Mohd_Talha_Lookalike.ipynb

Mohdtalha01 Add files via upload   b8c9c18 · 10 minutes ago

135 lines (135 loc) · 3.34 KB

Preview   Code   Blame   Raw

```python
In [1]:  from sklearn.metrics.pairwise import cosine_similarity
         from sklearn.preprocessing import MinMaxScaler
```

```python
In [5]:  import pandas as pd
```

```python
In [6]:  customers = pd.read_csv('Customers.csv')
         products = pd.read_csv('Products.csv')
         transactions = pd.read_csv('Transactions.csv')
```

```python
In [8]:  customer_features = transactions.groupby('CustomerID').agg({
             'TotalValue': 'sum',
             'Quantity': 'sum',
             'TransactionID': 'count'
         }).rename(columns={'TransactionID': 'TransactionCount'}).reset_index()
```

```python
In [9]:  customer_features = customer_features.merge(customers[['CustomerID', 'Regi
```

```python
In [10]:  customer_features = pd.get_dummies(customer_features, columns=['Region'],
```

```python
In [11]:  scaler = MinMaxScaler()
          scaled_features = scaler.fit_transform(customer_features.iloc[:, 1:])

          # Compute cosine similarity
          similarity_matrix = cosine_similarity(scaled_features)
```

```python
In [12]:  lookalike_results = {}
          for i, cust_id in enumerate(customer_features['CustomerID'][:20]):
              similarity_scores = list(enumerate(similarity_matrix[i]))
              similarity_scores = sorted(similarity_scores, key=lambda x: x[1], reve
              lookalike_results[cust_id] = [(customer_features['CustomerID'][j], sco
```

```python
In [14]:  import csv
          with open('Mohd_Talha_Lookalike.csv', 'w') as f:
              writer = csv.writer(f)
              writer.writerow(['cust_id', 'lookalike_cust_id_1', 'score_1', 'lookali
              for cust_id, lookalikes in lookalike_results.items():
                  writer.writerow([cust_id] + [item for sublist in lookalikes for it
```