REINFORCEMENT LEARNING – PROJECT

RECOMMENDER SYSTEM - ONLINE RETAIL DATA SET*

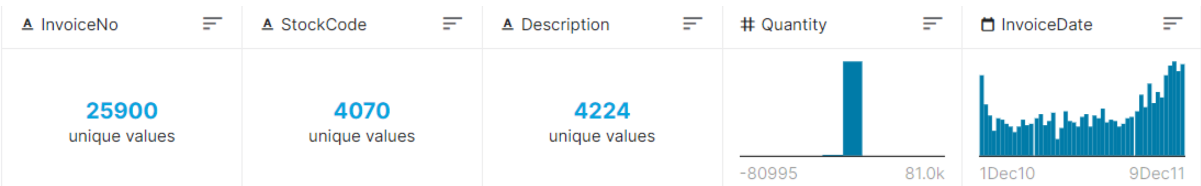
Business Problem

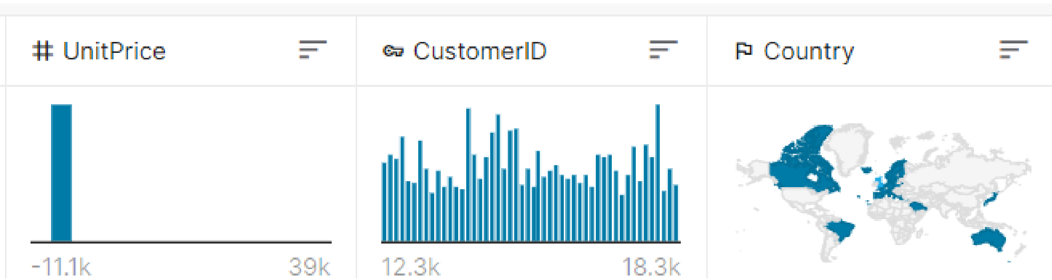
Based on customer purchase behavior data, the company needs to recommend products to customers.
The customer data snapshot is given below along with Data Source.

Data Source/ Code Reference:

Data downloaded from Kaggle: Link given below
[Online Retail Data Set | Kaggle]

No Direct Code available for download: Code stitched using different video tutorials and other web sources.





Change/Innovation in Code:

No change or innovation made to the code.
The code was stitched from multiple videos and online links.
Entire code is hand written from scratch.

Data Preprocessing and Modelling steps:

1. Preprocessing of data was done using:
    a. Pandas data frame techniques
    b. KMeans clustering analysis for feature generation

*Dataset downloaded from Kaggle

    c. Final data set was developed using merging, groupby and sorting commands
2. Modelling steps
    a. Each product were grouped and assigned a group - Segment
    b. Unique ID with name stateId was generated. This is based on the segmentation, month and day of purchase
    c. Dictionaries were created to record qty. of purchase
    d. Policy dictionary was created using gaussian distribution with random value near to quantity of product bought
    e. Rewards dictionary was updated using gaussian distribution with random value near to quantity of product bought
    f. A customer selection was manually generated (stateId)
    g. Using the stateId , recommendation was given to the customer. Here the initialized policy and rewards table is used.
    h. The recommendation is given based on exploitation and exploration method.
    i. Once the recommendation list is generated, then customer action to click on a particular product or purchase is manually triggered
    j. Then, using the trigger ad associated probability, the policy and rewards table are updated.
    k. This is how the RL works.

Comment on results:

<u>Result: Output of recommendation</u>

```
the context exists
product list ['21068', '21326', '20668', '22086', '84879', '21822', '35
957', '22158', '22834', '84947']
product 21068
count after update 217
Recommendation count after update 1
product 21326
count after update 193
Recommendation count after update 1
product 20668
count after update 169
Recommendation count after update 1
product 22086
count after update 157
Recommendation count after update 1
product 84879
count after update 153
Recommendation count after update 1
product 21822
count after update 1
Recommendation count after update 1
Policy value after update 0
Reward value after update -0.89
product 35957
count after update 145
Recommendation count after update 1
product 22158
count after update 137
Recommendation count after update 1
product 22834
count after update 122
Recommendation count after update 1
```

*Dataset downloaded from Kaggle

```
product 84947
count after update 113
Recommendation count after update 1
```

- The result of product recommendation is based on exploration and exploitation – greedy approach.
- The present rewards and values are based on normal average approach.
- The results are based on dummy testing of customer action.
- Product basket segmentation is done using a heuristic approach. Individual customer preferences can also be used.
- Results will get fine tunes as the RL algorithm learns more from future consumer actions.

Policy and Reward Update after customer action:

```
reward 4.0
reward value after customer action 141.42
product inside policy update 22158
current value of product in policy 135.4
No. of occurences of product 1
updated policy value -131.4
increment value of product before update -131.4
policy value after customer action 4.0

reward 3.35
reward value after customer action 196.76999999999998
product inside policy update 21326
current value of product in policy 191.12
No. of occurences of product 1
updated policy value -187.77
increment value of product before update -187.77
policy value after customer action 3.3499999999999943

reward -1.51
reward value after customer action 214.15
product inside policy update 21068
current value of product in policy 216.66
No. of occurences of product 1
updated policy value -218.17
increment value of product before update -218.17
policy value after customer action -1.5099999999999991

reward -0.93
reward value after customer action 166.73999999999998
product inside policy update 20668
current value of product in policy 167.32
No. of occurences of product 1
updated policy value -168.25
increment value of product before update -168.25
policy value after customer action -0.9300000000000068

reward -1.05
reward value after customer action 153.97
product inside policy update 22086
current value of product in policy 156.4
No. of occurences of product 1
```

*Dataset downloaded from Kaggle

```
updated policy value -157.45000000000002
increment value of product before update -157.45000000000002
policy value after customer action -1.0500000000000114

reward -2.79
reward value after customer action 149.84
product inside policy update 84879
current value of product in policy 151.94
No. of occurences of product 1
updated policy value -154.73
increment value of product before update -154.73
policy value after customer action -2.789999999999992

reward -2.53
reward value after customer action -3.42
product inside policy update 21822
current value of product in policy 0
No. of occurences of product 1
updated policy value -2.53
increment value of product before update -2.53
policy value after customer action -2.53

reward -1.53
reward value after customer action 142.33
product inside policy update 35957
current value of product in policy 144.52
No. of occurences of product 1
updated policy value -146.05
increment value of product before update -146.05
policy value after customer action -1.5300000000000011

reward -1.56
reward value after customer action 119.97
product inside policy update 22834
current value of product in policy 120.48
No. of occurences of product 1
updated policy value -122.04
increment value of product before update -122.04
policy value after customer action -1.5600000000000023

reward -0.6
reward value after customer action 110.84
product inside policy update 84947
current value of product in policy 113.91
No. of occurences of product 1
updated policy value -114.50999999999999
increment value of product before update -114.50999999999999
policy value after customer action -0.5999999999999943
```

*Dataset downloaded from Kaggle

Industry Application:

- Product recommendation for e-commerce players

*Dataset downloaded from Kaggle