

Merchandise Popularity Prediction

LaChandra Ash

Nov. 7, 2021

## Introduction

On the off chance that you get mistaken like that for the substance of your wardrobe, it's not difficult to perceive how you need assistance monitoring a huge number of SKUs in your stockroom. What's close accessible, what's submitted, what's apportioned to future shipments, what's on hand, what's been returned, resale, etc? You request the board and satisfaction framework to have the option to take a preview of this information for yesterday and today, however what might be said about tomorrow? How might the numerous multichannel advancements you have set up and have arranged influence interest? What number of SKUs will you need for which things fourteen days from now, or a month out?

Here's the place where product estimating can assist you with lessening orders. Indeed, there's an "application" for that, i.e., an application called "stock anticipating," which is commonly set up to serve at least one vertical business sectors, on the grounds that the requirements of a maker or distributor contrast fundamentally from a retailer, which thusly vary from those of online business or index shippers. Product estimating is a complex, numerically based cycle of "bend coordinating," where you attempt to adjust the business bend for every thing in the ebb and flow season with a bend from past seasons that best matches request by thing as every week in the season ticks by.

Since last season's complete, you can "foresee" where the ebb and flow season's bend is going. Direct traders need a framework that can:

- Figure interest by thing, in light of chronicled request and advancements (current and arranged); Propose buys (in view of distribution center restrictions, min/max

amounts, seller lead times, and authentic movement); Decrease rainchecks (by delivering precise estimates, obviously!); Limit overloads; Expand turns

## Frameworks Choices

There are four significant frameworks for direct business stock estimating:

- Galvin (galvinassociates.com) – This framework was first evolved in the last part of the 1980's and has been utilized by probably the biggest direct advertisers in the business.
- Its alternatives incorporate advertising, promoting, time-staged open-to-purchase, thing arranging, material prerequisites arranging, buy request the board, and thing question/report essayist.
- On the off chance that/SO from Harbinger Frameworks (forerunnersystems.com) keeps a harmony between the business or dissemination Plan and the product Plan; moves up all the individual thing gauges from various wellsprings of interest and conveys a day by day or week by week buy necessity dependent on "least danger/best return" rules; and takes into consideration thing explicit selling bends including changes for unique event occasions, including Web occasions. Clients incorporate Williams-Sonoma, Rebuilding Equipment, and Frontgate.

Forecast\*21 from Direct Tech (direct-tech.com) makes unit and dollar figures utilizing top-down and base up apparatuses; allots gauges by style, shading, and size to make SKU-level plans; applies week after week request bends to set up week after week request designs by SKU; and estimates interest for packs, sets and fabricated things. Clients incorporate Grounds-keeper's Inventory, PetEdge, and Appleseeds.

Barely (justenough.com) is fairly more retail-driven than the other three, and intended to work with Microsoft Elements/Hatchet, Netsuite, and SAP Retail, with modules for activities arranging, retail-level interest arranging, and advancements arranging and investigation. Clients incorporate Plan Reachable, B&H, and TJ Maxx.

### **Merchandise Popularity Predictions**

Gain proficiency with the nuts and bolts about predictive marketing and what independent ventures need to know. Perceiving predictive promoting via web-based media is a basic on and off chance that you look for an item and become quickly besieged with advertisements about the equivalent or comparative items, the organization is likely utilizing predictive marketing. A few advantages of predictive marketing are that it gives deals productivity and viability, it considers advancements focusing on, and makes for better stock administration. Some different advantages of predictive promotion are that it assists with hazarding the board and expands the exactness of monetary determination just as production networks cost the executives. Proof of an organization's utilization of predictive marketing is generally simple to spot via online media stages.

Quest for a thing on the web and unexpectedly you're besieged with promotions and proposals via online media offering the equivalent or comparable things. Yet, how do such countless organizations know which item you're probably purchasing soon?

### **What is predictive marketing?**

Predictive promoting (otherwise called predictive investigation and mechanized marketing) depends on innovation that catches and examines a wide scope of client information like buying propensities and history, perusing designs, individual inclinations, average spend, buy timing, and numerous other client exercises. This information is then used to make customized promoting and a custom shopping experience for every client that can construct faithfulness and increment store income. As of not long ago, just the biggest organizations had the assets to put resources into what were once organizations that claimed exclusive frameworks that assembled, broke down and utilized client information to dependably anticipate purchasing practices. Presently, because of cloud innovation, even independent companies can convey and profit by predictive marketing.

### **What are a few advantages of predictive marketing?**

Deals, productivity, adequacy, and promoting achievement depends on introducing the correct message to the perfect client at the perfect time. Predictive marketing advances assemble and break down all the information that organizations require to figure out who are the best clients and how to keep them. Advancements focusing on the correct advancement to the correct client is one of the large advantages of predictive promoting. From the client's viewpoint, the offers they get are those that are most actually important. Organizations see higher change rates from the advancements they run and, at last, higher income.

Better stock administration and predictive promoting can empower independent ventures to more precisely estimate stock necessities, which decreases overloading and related expenses. More compelling publicizing procedure utilizes predictive models dependent on your business'

client information, promoting dollars can be spent all the more successfully to reach and change over possibilities that take after the practices of your best clients. Improved situating over contenders.

Predictive promoting helps organizations with all things considered, and types contend more viable in the present commercial center. This critical apparatus assists with night the battleground with regards to building and holding a bigger client base while additionally improving business procedure and smoothing out tasks.

### **MPP Models**

On the off chance that you are keen on seeing instances of predictive promoting, as indicated by Acuvate, beneath are a few models:

- Conduct examination: By gathering information on clients, organizations can more readily comprehend their clients to make more powerful deals crusades. The information gathered can give knowledge permitting organizations to more readily perceive their high-esteem clients, the intentions behind their buys, their purchasing behaviors, and which are the best channels to market to them and why.
- Customized in-store insight: By utilizing new human following innovation, stores are currently ready to break down in-store conduct to more readily survey the effect of marketing that has been created lately. This can permit organizations to more readily customize the in-store insight and support steadfast clients by giving impetus to visit customers.

- Examination on activity and inventory network: Given that items have much quicker life cycles and complex tasks, enormous information investigations are regularly used to more readily comprehend supply ties and to help diminish costs. To acquire a serious edge, organizations need to all the more likely comprehend examinations on tasks and store networks.
- Exchange advancement enhancement: Reports have shown that organizations can lose up to the cash they put resources into exchange advancements. This is on the grounds that it very well may be close to incomprehensible for leaders to quantify its viability and return for capital invested. In any case, when utilized related to prescriptive investigation, predictive examination can permit organizations to bridle ongoing organized and unstructured information from different business sectors just as customer touch points. This information would then be able to be utilized to make significant suggestions to help organizations discover exchange advancement techniques that really work.

As well as improving deals income and client maintenance, predictive promoting can help in other significant regions of business, for example:

- Exactness of monetary gauging
- Production network cost the executives
- Danger the board
- Benefit and income development projections exactness

- In general, essential business dynamic
- Business insight and revealing
- Information sharing, coordinated effort and perception

Today, most private ventures gather and store client information. However, a common independent company normally can't keep up deciphering and investigating the huge measure of data accessibility from the developing number of information sources messages, site visits, uphold calls, different deals channels, online media and others. In addition, most private ventures don't have the assets or aptitude to break down the information they as of now assemble consistently. Predictive marketing or predictive insightful applications are extraordinarily quick. Further, they regularly utilize a lot more extensive arrangements of information and can follow and break down the data that a couple of private ventures can act progressively.

### **Hypothesis/Questions**

- How can businesses benefit from MPP?
- Can MPP increase sales?
- How can businesses use MPP?
- What are the pros for using MPP?
- What are cons of using the predictions?
- Which categories had the highest amounts of sales?

- What is the accuracy percentage of merchandising predictions?
- Does the dataset provide enough information to predict the merchandise?
- Can the business have a competitor advantage if they used the merchandise predictor model?
- Are there any types of merchandise that cannot be used by the predictor?

## Methods

1. Obtained the train and test datasets from the referenced website.
2. Downloaded catboost via pip install

```
!pip install catboost
```

3. Imported libraries and modules

```
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from lightgbm import LGBMClassifier
from catboost import CatBoostClassifier
from xgboost import XGBClassifier
from sklearn.linear_model import LogisticRegression , Ridge
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier , GradientBoostingClassifier
from sklearn.model_selection import KFold , StratifiedKFold
from sklearn.metrics import log_loss
import warnings
warnings.filterwarnings('ignore')
```

4. Unzipped the dataset

```
!wget --header="Host: machinehack-be.s3.amazonaws.com" --
!unzip dataset.zip|
```

5. Read the train and test datasets into a dataframe

```
train = pd.read_csv("MPP_Dataset/Train.csv")
test = pd.read_csv("MPP_Dataset/Test.csv")
train
```

6. View the first 5 rows of the dataset

	Store_Ratio	Basket_Ratio	Category_1	Store_
0	0.407	0.00380	2	-
1	0.234	0.10500	0	-
2	0.668	0.72600	9	-
3	0.184	0.00561	4	-
4	0.231	0.13100	6	-
...	...	...	...	
18203	0.403	0.28100	6	-
18204	0.481	0.01140	4	-
18205	0.168	0.00908	5	-
18206	0.598	0.65800	1	-
18207	0.582	0.32400	2	-

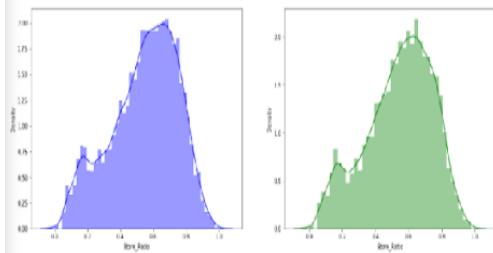
18208 rows × 12 columns

7. View, print, plot, and display the data in graphs

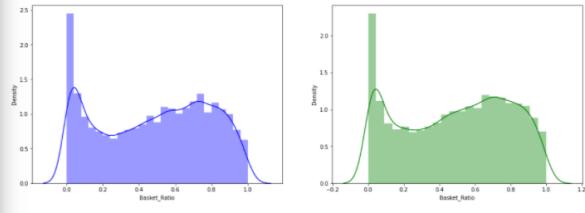
```
for column in test.columns:
    print(column)
    fig,ax = plt.subplots(nrows=1,ncols=2,figsize=(18,6))
    sns.distplot(train[column],ax=ax[0],color="blue")
    sns.distplot(test[column],ax=ax[1],color="green")
    plt.show()
```

8. Display the visualizations

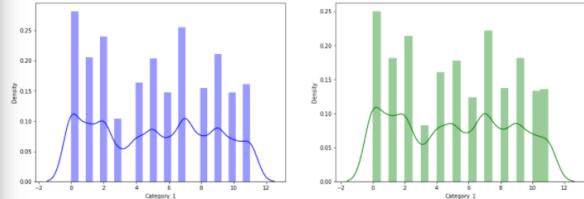
### Store\_Ratio



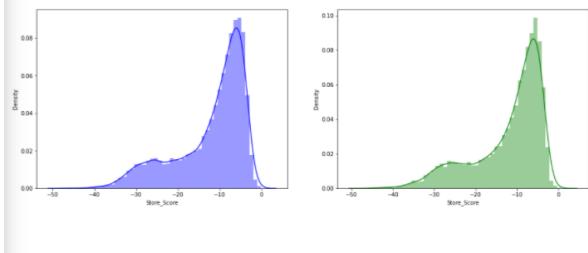
### Basket\_Ratio



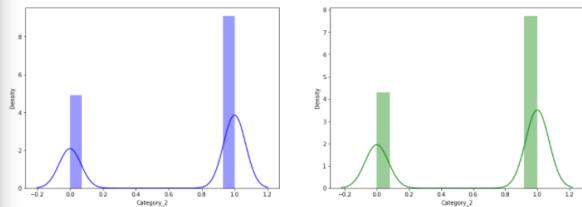
### Category\_1



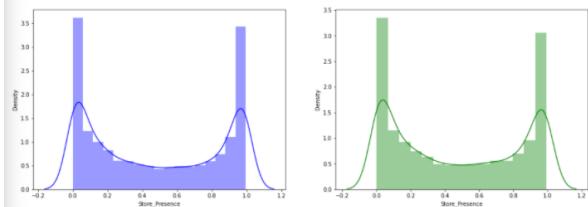
### Store\_Score

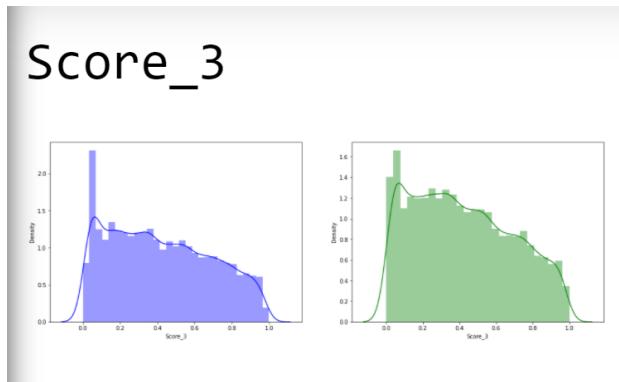
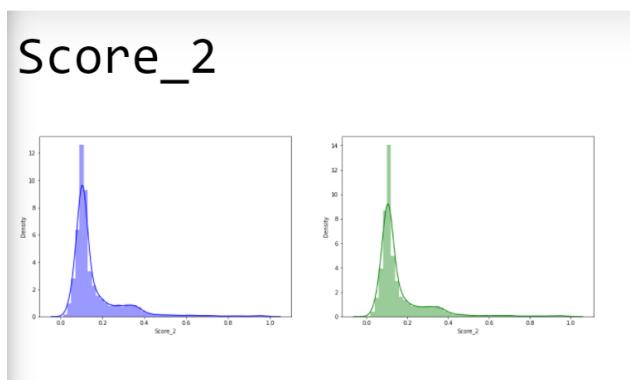
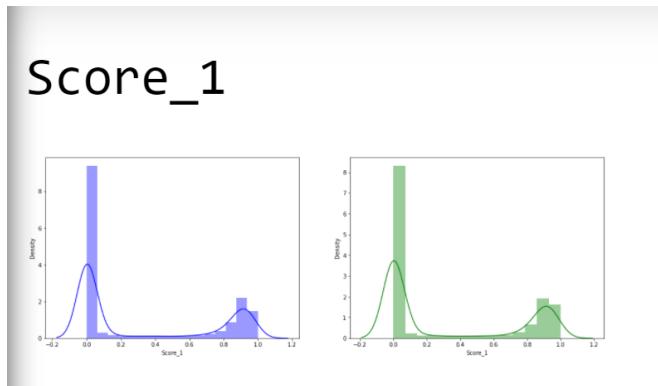


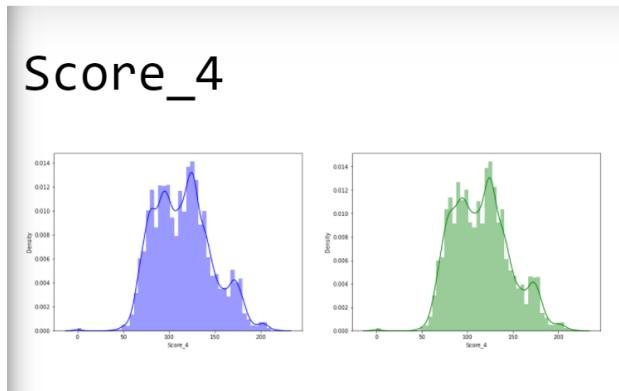
### Category\_2



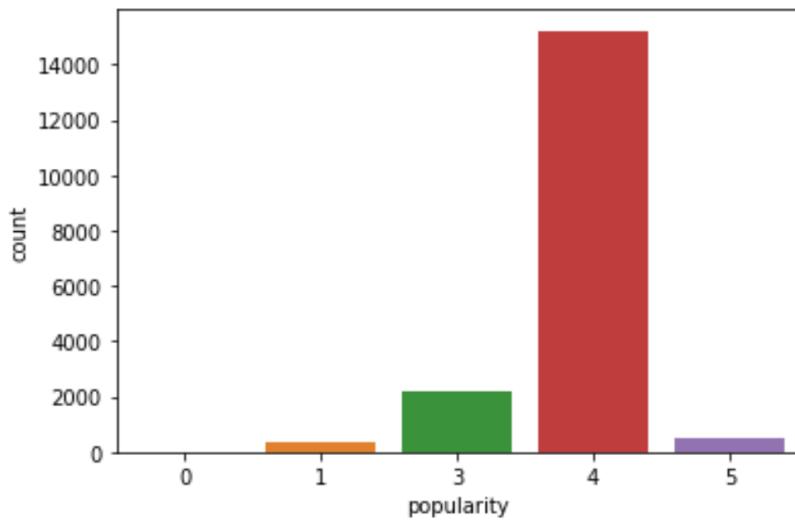
### Store\_Presence







```
sns.countplot(train.popularity)
```



```
X = train.drop('popularity',axis=1)
Y = train['popularity']

from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
X_ = scaler.fit_transform(X)
X = pd.DataFrame(X_,columns=X.columns)

folds = StratifiedKFold(n_splits=5,shuffle=True ,random_
oof_preds = []
oof_scores = []

for i , (train_idx,test_idx) in enumerate(folds.split(X,
    train_set = (X.iloc[train_idx],Y.iloc[train_idx]))
    test_set = (X.iloc[test_idx],Y.iloc[test_idx])

model = LogisticRegression(max_iter=5000)
model.fit(*train_set)

preds = model.predict_proba(test_set[0])
score = log_loss(test_set[1],preds)

oof_pred_fold = model.predict_proba(scaler.transform(t
oof_preds.append(oof_pred_fold)

print(score)
oof_scores.append(score)

print("Mean OOF Score : ",np.mean(oof_scores))
final_preds = np.mean(np.array(oof_preds),axis=0)
```

```
submission = pd.DataFrame(final_preds)
submission.to_csv("submission.csv", index=None)
submission|
```

## Conclusion

The results displayed category 4 had the highest number of popularity predictions. I advised the business to maintain a constant supply for category 4. The business should consider keeping category 3 in stock because it had the second highest number of popularity. The other categories displayed a lower number of popularity. I advised the business to create a plan on how they will use specific tactics to increase the sales and its popularity ranking score.

## References

- Chandra, D. K., Ravi, V., & Bose, I. (2009). Failure prediction of dotcom companies using hybrid intelligent techniques. *Expert Systems with Applications*, 36(3), 4830-4837.
- Golbandi, N. G., Katzir, L. K., Koren, Y. K., & Lempel, R. L. (2013, February). Expediting search trend detection via prediction of query counts. In *Proceedings of the sixth ACM international conference on Web search and data mining* (pp. 295-304).
- Jian, Z., Qingyuan, Z., & Liying, T. (2020). Market revenue prediction and error analysis of products based on fuzzy logic and artificial intelligence algorithms. *Journal of Ambient Intelligence and Humanized Computing*, 1-8.
- Kaggle (n.d.). Merchandise Popularity Predictions Dataset. Retrieved from <https://www.kaggle.com/datasets>.
- Park, J. H., & Park, S. C. (2003). Agent-based merchandise management in business-to-business electronic commerce. *Decision Support Systems*, 35(3), 311-333.
- Park, J. H., & Park, S. C. (2003). Agent-based merchandise management in business-to-business electronic commerce. *Decision Support Systems*, 35(3), 311-333.
- Jing, P., Ye, S., Nie, L., Liu, J., & Su, Y. (2019). Low-rank regularized multi-representation learning for fashion compatibility prediction. *IEEE Transactions on Multimedia*, 22(6), 1555-1566.

- Li, D., Zhao, G., Wang, Z., Ma, W., & Liu, Y. (2015, November). A method of purchase prediction based on user behavior log. In *2015 IEEE International Conference on Data Mining Workshop (ICDMW)* (pp. 1031-1039). IEEE.
- Lo, A. W., & Hasanhodzic, J. (2010). *The evolution of technical analysis: financial prediction from Babylonian tablets to Bloomberg terminals* (Vol. 96). John Wiley & Sons.
- Nguyen, L. T. K., Chung, H. H., Tuliao, K. V., & Lin, T. M. (2020). Using XGBoost and Skip-Gram Model to Predict Online Review Popularity. *SAGE Open*, 10(4), 2158244020983316.
- Yuan, H., Xu, W., Li, Q., & Lau, R. (2018). Topic sentiment mining for sales performance prediction in e-commerce. *Annals of Operations Research*, 270(1), 553-576.