A Mini Project Report

*on*

D/Mart Number of Counters Problem

In Subject: Discrete Mathematics and Data Structures

*by*

|  |  |  |
| --- | --- | --- |
| Name | ROLL NO | PRN |
| Pushpak Surywanshi | 272056 | 22010400 |
| Soham Tolwala | 272058 | 22010024 |
| Saiprasad Toshatwad | 272060 | 22010203 |
| Prajwal Waykos | 272063 | 22010591 |



Department of Artificial Intelligence and Data Science

VIIT

**2020-2021**

**Contents**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr. No.** | **Topic** | | **Page No.** |
|  |  | |  |
| **Chapter-1** | **Introduction** | |  |
|  | 1.1 | Introduction | 1 |
|  | 1.2 | Requirements | 2 |
|  | 1.3 | Design & Problem Statement | 3 |
|  | 1.4 | Proposed work | 4 |
| **Chapter-2** | **Methodology** | |  |
|  | 2.1 | Datasets | 5 |
|  | 2.2 | Approach | 6 |
|  | 2.3 | Platform and Technology | 7. |
|  | 2.3 | Outcomes & Use cases | 8. |
|  | 2.4 | Challenges | 9 |
| **Chapter-3** | **Conclusion** | | 9 |
|  | **References** | | 10 |

Introduction

DMART is one of the most popular shopping malls in India.

It is a one -stop supermarket chain that offers customers a wide range of basic home and personal products under one roof in a cheap and reasonable price,and as a result the customers are always attracted towards it and so there is a great demand for the products in the mall,especially in the festive days than other days.

Apart from good quality products,the mall also offers good quality services,such as helping the customers to buy the right choice of products,increasing the number of counters in the mall during festivals that will help the customers to save their time,instead of standing in the queue for a long time for bill payment.And so the mall pays a large and hefty amount of money to give good quality services to customers.

Thus,our aim as data scientists is to help the owner of Dmart to save his revenue on the extra services which are not required with the help of various prediction models.

Requirements

1. The first and the foremost asset required to build any Machine Learning or Deep Learning Model is Data
2. All the data used in this project is entirely Hypothetical and we students have designed it on ourselves. Later, we can take any real time data and implement the similar model on it.
3. Also, there are few Software requirements such as
4. **Excel-** It is being Used to prepare data
5. **Python –** It is the programming language we have used throughout the project
6. **VsCode, -** It is the IDE that we have used to implement the actual code.
7. **Tableau –** Tableau is being used for preparing good visualisations
8. **Powerpoint –** Used to prepare presentations
9. **Animaker** – Used to create animated video

10. **Git-Hub** – Used to maintain flexibility for programmers.

Design And Problem statement



The above diagram represents that the problem is basically divided into 2 parts first is the data creation and second is the Model building

The Data creation part includes Defining the various data variables also called as features and second is the creation of records with respect to the different attributes withing the database like MySQL or Excel sheet.

Problem statement – **Design a Regression model that is able to predict the number counter boys that the mall owner should employ in order to meet the long queue problem**

Proposed work

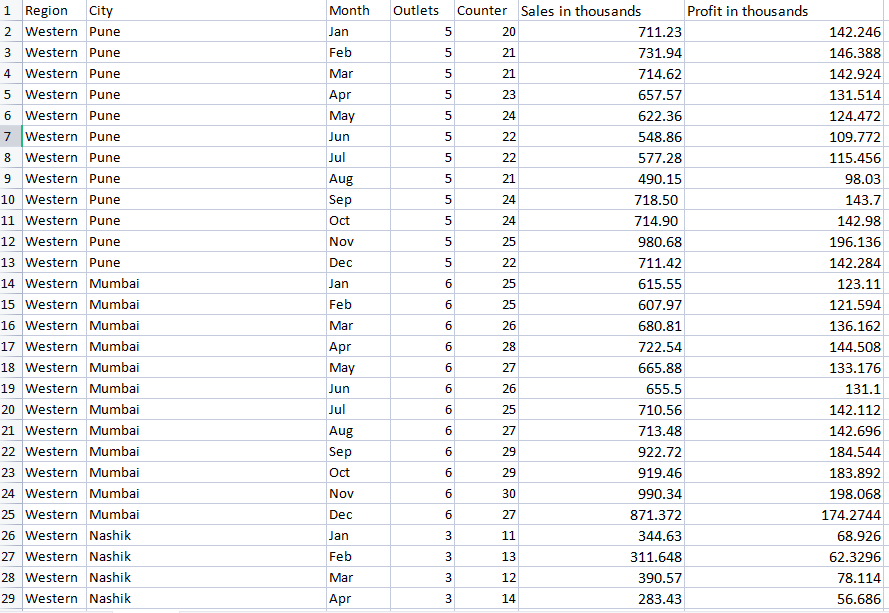
We have divided the problem statement into two parts,i.e data creation and model building.

The data creation part includes defining the various data variables also called as features and second is the creation of records with respect to the different attributes using the dataset.

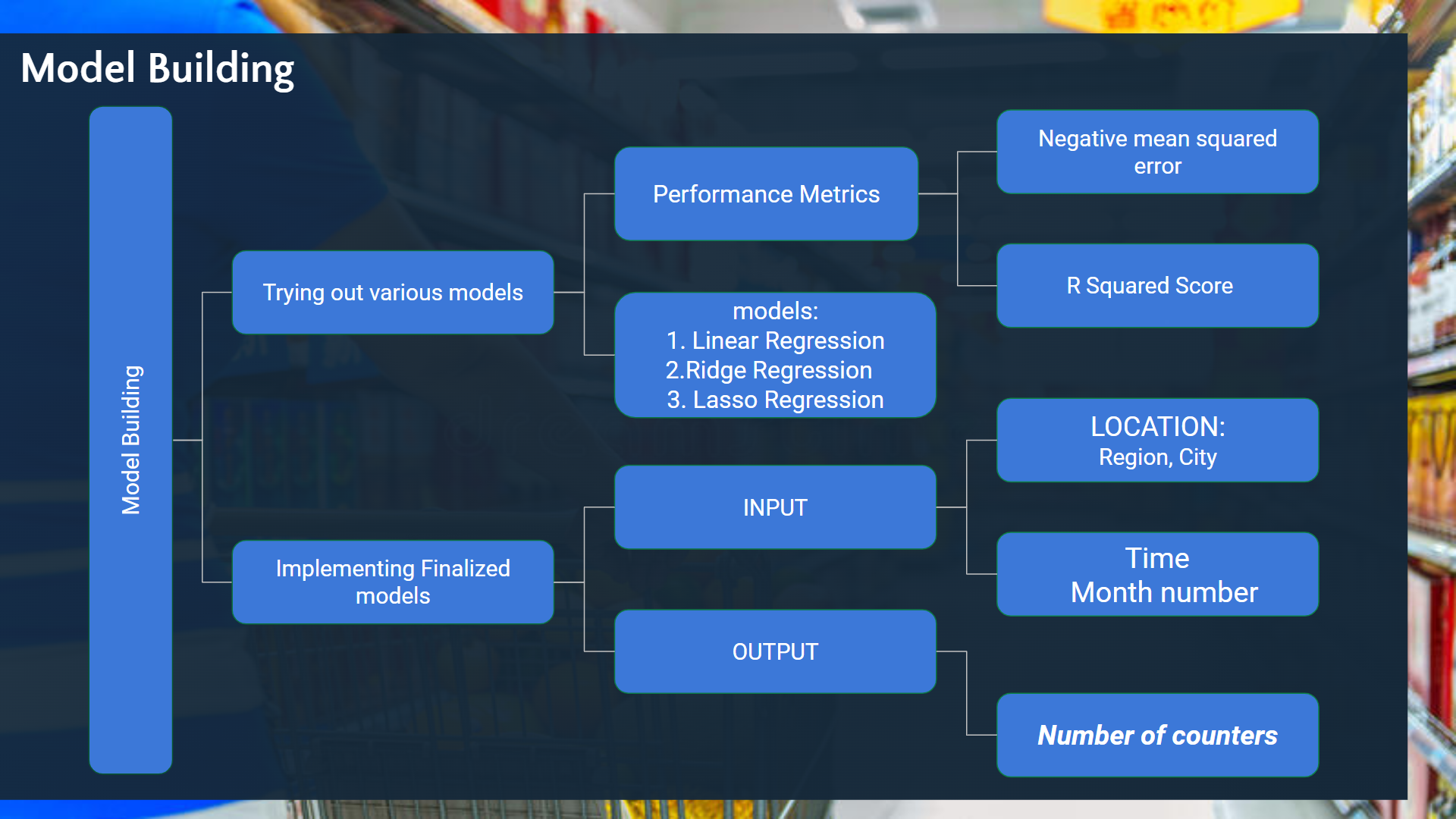
The dataset includes various varaibles such as the region and city in which the mall is located,the number of outlets in each city,the sales revenue of the mall,and the profit gained by it.

Using the above dataset and records we create various prediction models to find out the number of counter boys that should be employed in order to meet the long queue problem.

Dataset



Approach



The Next step is the model building in this we had to do 2 things , first is the trying out of various regression models such as Linear Regression, Logistic Regression, Decision tree regressor, Random Forest Regressor, Lasso Regression and Ridge Regression.

We implemented all these models and calculated the negative mean squared error for all of them out of which the Ridge, Lasso and Linear Regressors had the least mean squared error which was around

-1.333515151313616, -1.34395205013, -1.345135326366 repectively

For further filtration we calculated the R square score of all the models which came out to be

Ridge = 98.115332

Lasso = 97.8353532

Linear = 97.923335

Hence from the above data we finalized the RIDGE REGRESSION as our final Machine Learning model for this particular model.

Platforms and Technology

1. The first and the foremost asset required to build any Machine Learning or Deep Learning Model is Data
2. All the data used in this project is entirely Hypothetical and we students have designed it on ourselves. Later, we can take any real time data and implement the similar model on it.
3. Also, there are few Software requirements such as
4. **Excel-** It is being Used to prepare data
5. **Python –** It is the programming language we have used throughout the project
6. **VsCode, -** It is the IDE that we have used to implement the actual code.
7. **Tableau** – Tableau is being used for preparing good visualisations
8. **Powerpoint –** Used to prepare presentations
9. **Animaker** – Used to create animated video
10. **Git-Hub –** Used to maintain flexibility for programmers.

Outcomes and Use Cases

The major outcomes are:

1.The algorithm is able to predict the number of counters based upon the given input of city, Region and Month.

2. the algorithm can be made better more than 100 times if we could get any Real Time data.

3. we can surely also add the GUI To enhance its user comfort

The major Use-cases are:

1. It can be used by mall owners.
2. It can be used by various franchise owners
3. It can also be owned by bug shop owners who have multiple shops
4. Also it can be expanded according to the availability of data.

Challenges

1. The first and the most difficult challenge was to create data.
2. The second challenge was to figure out the best model that fits the data.
3. Next was to create some good UI to such that the user could handle the program with ease.
4. Preparing visualizations and managing files were also one of the greatest concerns which was skillfully tackled by Git-hub and my responsible team mates.

Conclusions:

1.The algorithm is able to predict the number of counters based upon the given input of city, Region and Month.

2. Also we concluded that given the dataset and the other conditions the Ridge Regression works the best followed by Lasso and Linear regression.

References

1. Link of the dataset:-https://in.docworkspace.com/d/sIM3csc1juZ-7jQY
2. Python and Vs code software to compute the code and create the model.
3. Google-To get images of the mall.
4. Animaker-To create animated video.