**Symbiosis Skills and Professional University Kiwale, Pune**

------------------------------------------------------------------------------------------

**PROJECT REPORT**

**On**

**“Customer Decision Prediction”**

****

**Submitted by**

**Name of student-Priyanka Diliprao Jagtap(Registration Number)**

**ML-Batch-VI**

**Under The Guidance of**

**Trainers’ Name: Ankita Kanchan Ma’am**

**STUDENT DECLARATION AND ATTESTATION BYTRAINER**

This is to declare that this report has been written by me. No part of the report is plagiarized from other sources. All information included from other sources have been duly acknowledged. I aver that if any part of the report is found to be plagiarized, I shall take full responsibility for it.

Signature of student – Priyanka Diliprao Jagtap.

**Name of student: Priyanka Diliprao Jagtap.**

Registration Number: ML-6

Signature of trainer- Ankita Kanchan

**Name of trainer: Ankita Kanchan Ma’am**

**CERTIFICATE**

This is to certify that the report entitled, **“Customer Decision Prediction”**submitted by **“Priyanka Diliprao Jagtap**” to Symbiosis Skills and Professional University, Pune, Maharashtra, India, is a record of bonafide Project work carried out by him under my supervision and guidance and is worthy of consideration for the completion of certificate course in ‘Machine Learning Engineer”.

Signature of Trainer- Ankita Kanchan

Name of Trainer-Ankita Kanchan Ma’am

Date: 14 /7 / 2021

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Supervisor Supervisor

Date: 14/7/2021

**INDEX**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Customer Decision Prediction by use of Machine learning** | | | | | |
| **Sl. No** | **Index** |  |  |  | **Page No.** |
| 1 | **Acknowledgement** | |  |  | 5 |
| 2 | **Plan of Capstone Project** | |  |  |  |
|  | 2.1 | Purpose of Project | |  | 6 |
|  | 2.2 | Period of Project | |  | 7 |
|  | 2.3 | Problem Statement Detailing | |  | 8 |
| 3 | **Objective of the** Project | | |  |  |
|  | 3.1 | General Objective | |  | 9 |
| 4 | **Introduction (About the problem Statement)** | | |  |  |
|  | 4.1 | Introduction | |  | 10 |
|  | 4.2 | Predictive system | |  | 11 |
|  | 4.3 | Architecture diagram | |  | 12 |
|  | 4.4 | Methodology | |  | 13 |
|  | 4.5 | Results | | | 15 |
| 5 | **Future plans** | | |  |  |
|  | 5.1 | Future Plan | | | 16 |
| 6 | Learning from the Project | | |  | 17 |
| 7 | Conclusion | | | | 18 |

**1.1ACKNOWLEDGEMENTS**

I would like to express my special thanks of gratitude to my trainer Ankita Kanchan Ma’am who gave me the golden opportunity to do this wonderful project on the topic “Customer Decision Prediction” by using classification algorithm, which also helped me in doing a lot of Research and I came to know about so many new things.

I would also would like to special thanks to my group members Priyanka D. Jagtap,

Moinuddin Farooqui for effective support and important contribution in our project.

I am really thankful to them.

Secondly I would also like to thank my Parents and also my friends who helped me a lot in finishing this project within the limited time.

I am making this project not only for marks but to also increase my knowledge.

THANKS AGAIN TO ALL WHO HELPED ME.

**2.1Purpose of Project**

To understand the customer decision making process is key to identifying marketing challenges and opportunities .it’s important to align marketing efforts with the steps customer undertake to decide what to buy.

To study the customer behaviour and it is important because it helps marketers understand what influences consumer buying decisions. By understanding how consumer decide on a product, they can fill in the gap in the market and identify the products that are needed and the products that are absolete.

**2.2Period of Project**

1. First we collect data from internet .
2. Data preprocessing and Feature engineering- For data preprocessing and feature engineering we spent our 3 days.
3. Model building -For model building it took us 4 days.
4. Total period required for our project completion is 1 month.

**2.3Problem Statement Detailing**

Consumer behaviour does not stay identical or constant in each scenario it changes time to time. There are various factors which affect consumer behaviour. As the change comes in these factors , consumer behaviour also changes. The demographic factors that have an effect on consumer behaviour are Age, Gender, marital status, financial status ,family background, education ,occupation, family size. Then, To analyse the impact of demographic factors on consumer behaviour .

**3.1General Objective**

The dataset we’ll use for this python project- we’ll call it Data of Customer.csv. This dataset has a shape of 25000×9. The columns are id, state, age, occupation, income, customer sice, loyalty program, past purchase ,purchase .To build a model to accurately classify a customer will purchase product or not. Here we used machine learning classification algorithm, to check the customer will purchase product or not and to check the which model is best model for accurately classify.

**4.1Introduction**

Consumer Behaviour is the study of individuals , groups , or organization and all the activities associated with the purchase , use and disposal of goods and services, including the customer emotional , mental and behavioural responses that precede or follow these activities . The consumer is that the focus of all the selling activities .Knowledge of his behaviour is one of the foremost and vital aspects of the selling activities. Consumer behaviour does not stay identical or constant in each scenario it changes time to time .there are various factors which affect consumer behaviour .As the change comes in these factors , consumers behaviour also changes.

**4.2Prediction System**

In this project the focus on machine learning technique to predict the customer decision . The placement prediction is done by machine learning Classification algorithm and to check the which model is best model for accurately classify .

Our Data set have following columns:

1. ID - Unique identifier of a customer

2. STATE - Customer's state of residence

3. AGE - Customer's age

4. OCCUPATION - Customer's occupation

5. INCOME\_GROUP - Customer's income band

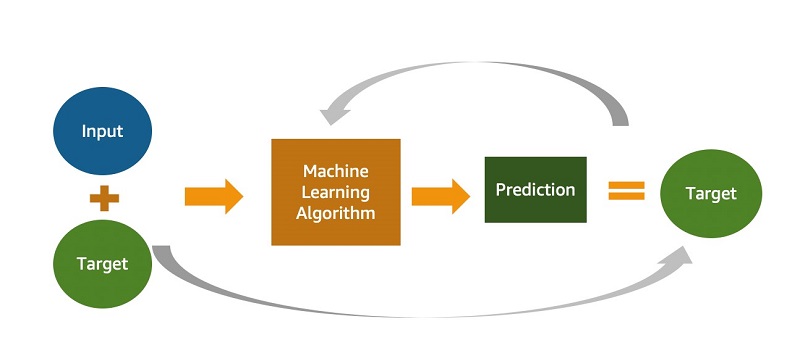
6. CUSTOMER\_SINCE - Year in which the customer made the first purchase at the store

7. LOYALTY\_PROGRAM - Is the customer enrolled for the loyalty program?

8. PAST\_PURCHASE - Total amount of past purchase from the same store

9. PURCHASE - Target variable(purchase or not)

**4.3ARCHITECHTURE DIAGRAM**



**4.4Methodology**

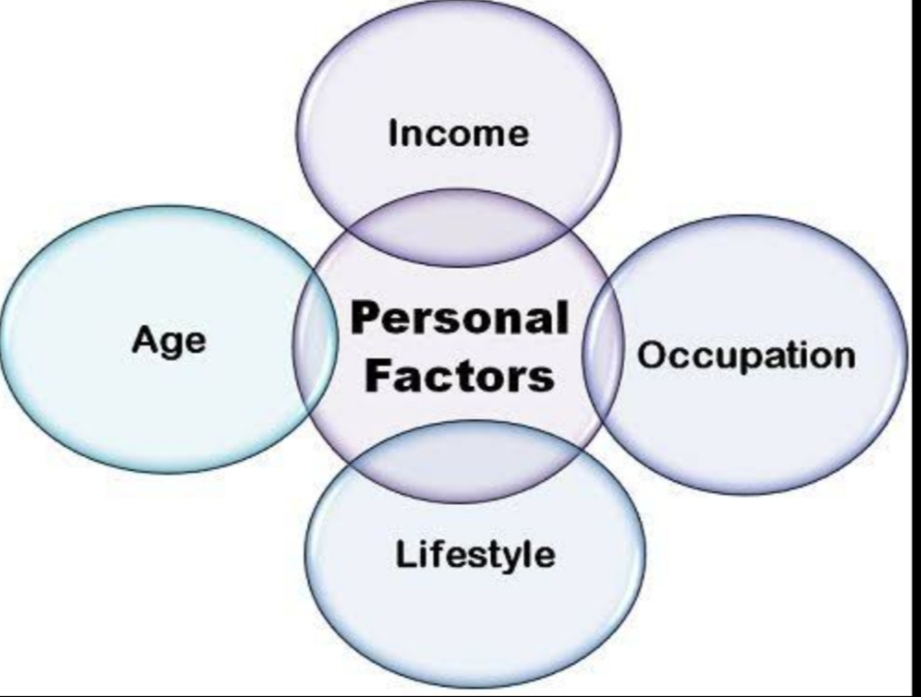
Choosing Machine Learning algorithm :

The output of report expected binary data that is binary format so that we can run following algorithms.

1. Logistic Regression
2. KNN
3. SVM
4. Random Forest Classifier
5. Decision Tree Classifier
6. Adaboost Classifier
7. Gradient Boosting Classifier
8. Bagging Classifier
9. XGBoost Classifier

**How a consumer Makes purchase Decisions**

**Factors affects Customer** **Behaviour**

****

**4.5Results**

The Accuracy for LogisticRegression is : 0.8384

The Accuracy for DecisionTreeClassifier is : 0.7842

The Accuracy for SVM is : 0.7606

The Accuracy for KNN is : 0.7106

The Accuracy for GNB is : 0.8406

The Accuracy for RandomForestClassifier is : 0.855

The Accuracy for AdaBoostClassifier is : 0.8564

The Accuracy for GradientBoostingClassifier is : 0.8598

The Accuracy for BaggingClassifier is : 0.8466

The Accuracy for XGBClassifier is : 0.8614

**Best model is XGBClassifier with Accuracy 0.8614**

Classification report:

precision recall f1-score support

0 0.86 0.98 0.91 3803

1 0.88 0.49 0.63 1197

accuracy 0.86 5000

macro avg 0.87 0.73 0.77 5000

weighted avg 0.86 0.86 0.85 5000

Accuracy:

Applying **XGBClassifier** - 86.14%

**5.0Future plans**

If any of organization shows interest in our project then we would like to work more to this project to making effective solutions. we implement and added some feature like natural language processing and advance deep learning Techniques.

**6.0 Learning from Project**

We have learnt many things from project would like to put in list.

1. Collection of data.
2. How to preprocess data.
3. In order to create a machine learning model there is always requirement of cleaned data so we have leant to apply various data preprocessing techniques and feature engineering techniques.
4. We have applied feature engineering which suits best to our dataset. Then we have applied various machine learning algorithms on it and learnt algorithms by coding and as well as mathematics behind each algorithm.
5. XGBClassifier algorithm was the best algorithm which gave us accuracy of 86.14%
6. In this way we have learnt very different things.

**7.0Conclusion**

* For this data, the XGBClassifier is the best fitted model with Accuracy 86.14%.
* More customer does not purchase product from the store.
* Most of the customer does not enrolled for the loyalty program.
* More customer purchase total amount is maximum 20000.
* Self employed people purchase more than other.
* People of age group between 26 to 32 are heavy buyers.
* Person having low income prefer to buy item from the store.
* In year 2017 & 2018 more customer made the first purchase at the store.