MACHINE LEARNING ENGINNER

A SUMMER TRAINING REPORT

Submitted by

SRAJAN AGARWAL

19BCS4021

Submitted in partial fulfilment of summer training for the award of the degree of

BACHELOR OF ENGINEERING

IN

CSE WITH SPECIALIZATION IN CLOUD COMPUTING

CHANDIGARH UNIVERSITY
Discover. Learn. Empower.

APEX INSTITUTE OF TECHNOLOGY
CHANDIGARH UNIVERSITY, GHARUAN
MOHALI, PUNJAB

OCTOBER 2021

Page No.

CONTENTS

CERTIFICATE

CETPA INFOTECH PRIVATE LIMITED

(An ISO 9001:2015 Certified Company)

Certificate of Training

This is to certify that

SRAJAN AGARWAL

has successfully completed Six Weeks Summer Training on
"Machine Learning"
from 9 June to 20 July, 2021
at CETPA INFOTECH PVT. LTD., Noida.

Anil Kumar Singh
Director-Training

CETPA®
Because Knowledge Matters
www.cetpainfotech.com

Vikas Kalra
Director

Verify this Certificate by Registration Number ESMach9620216W1808367 at http://www.cetpainfotech.com

About the Company

CETPA INFOTECH PVT. LTD, certified by ISO 9001:2015 is Northern India's best training institute which offers training in many evergreen technologies. CETPA has designed its training program to fulfill the needs of engineering students by offering various types of training programs.

CETPA is a training, development and placement company which offer different technical training to technical students as well as working professionals. It has helped many job seekers, corporate professionals and learners to improve their technical as well as soft skills in order to become a proficient professional. To fulfill the increasing demand of training in foreign countries, CETPA, the best training institute provides excellent overseas training.

ACKNOWLEDGEMENT

I am highly indebted to CETPA INFOTECH for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the Training.

I would like to express my gratitude towards my parents & my trainer Mr. Atma Rai for their kind co-operation and encouragement which help me in completion of this Training and project.

I would like to express my special gratitude and thanks to industry persons for giving me such attention and time.
My thanks and appreciations also go to my colleague in developing the project and people who have willingly helped me out with their abilities.

ABSTRACT

In my training of Machine Learning I have learnt new algorithms and these algorithms can help the market in many ways, So I have chosen the project “Grocery Item Prediction”, As we are moving ahead with Digital India, for businesses to survive in the market should also understand the needs of their customer.

So by understanding the need of the market and expectations of the sellers from the ML ENGINEER I have chosen this Project to help the society.

List of Tables

List of Figures

List of Photographs

CHAPTER: 1 INTRODUCTION

1.1 Intro to Machine learning:

Machine learning is the study of getting PCs to act without being unequivocally modified. In the previous decade, Machine learning has given us self-driving vehicles, reasonable discourse acknowledgment, successful web search, and an unfathomably worked on comprehension of the human genome. Machine learning is so unavoidable today that you most likely use it many times each day without knowing it. Numerous scientists additionally think it is the most ideal way of gaining ground towards human-level AI.

Classical machine learning is often categorized by how an algorithm learns to become more accurate in its predictions. There are four basic approaches:

- supervised learning,
- unsupervised learning,
- semi-supervised learning
- reinforcement learning.

1.2 Different Types of Algorithms:

These type of algorithm data scientists choose to use depends on what type of data they want to predict:

- **Supervised learning:** In this type of machine learning, data scientists supply algorithms with labelled training data and define the variables they want the algorithm to assess for correlations. Both the input and the output of the algorithm is specified.
- **Unsupervised learning:** This type of machine learning involves algorithms that train on unlabelled data. The algorithm scans through data sets looking for any meaningful connection. The data that algorithms train on as well as the predictions or recommendations they output are predetermined.

- **Semi-supervised learning:** This approach to machine learning involves a mix of the two preceding types. Data scientists may feed an algorithm mostly labelled training data, but the model is free to explore the data on its own and develop its own understanding of the data set.
- **Reinforcement learning:** Data scientists typically use reinforcement learning to teach a machine to complete a multi-step process for which there are clearly defined rules. Data scientists program an algorithm to complete a task and give it positive or negative cues as it works out how to complete a task. But for the most part, the algorithm decides on its own what steps to take along the way.

CHAPTER: 2 THEORY

2.1 General Theory

Machine Learning Training which was given by the Cetpa Infotech has transformed me and upgraded my skills, they have complete infrastructure, they have used the complete practical approach to make me understand each and every concept of Machine Learning.

2.2 Types of System Available

- Capability to run Online/offline classes together
- Practical classes on google Colab and Jupyter
- 1 Year membership that allow students to study n no. of times

2.3 Research Works Carried

I have worked upon a large dataset of amazon which taken my immense time to understand the problem and sort out the solution, By spending hours on the Dataset, I have filtered it and found around 10000 useful entries which helped by project to work with 93% accuracy.

CHAPTER 3: MARKET ANALYSIS TO SCALE UP BUISNESS

3.1 Introduction To Problem:

I have taken the immense time to understand the problem of the market, and came up with the solution as my project “Grocery Item Prediction”, In this we have understand the need of sellers and in this project used the large dataset with 180 products and around 15000 entries. We have created the model that can give 93%. This will help the company to connect our application to their software which will show the suggestions to customer when they add or buy something from the seller’s platform.

3.2 Need and objective of Study:

ML helps in extorting meaningful data from innumerable raw data. If ML is implemented properly, it can smartly predict altering customer behaviours and act as a problem solver for a variety of complex business problems. Big technology giants like Microsoft, Google, Amazon, etc., have landed with their Cloud ML platforms. Below are listed the reasons why a start-up should quickly adopt ML.

3.3 Scope of ML

- **Instantaneous Decision Making**

Each start-up depends on exact data in order to settle on ideal best choices. Without the assistance of smart technology like ML, it is close to difficult to extricate the right information from such colossal information. ML grants new businesses to change enormous arrangements of information into significant and information knowledge. Consequently, to react to this constantly changing business conditions and market requests, this data can be utilized to do ordinary business functional exercises. Thusly, the new companies can settle on moment business choices and stay at the edge of their competition.

- **Eliminates Manual Tasks**

Inaccurate and duplicate data entry can become the biggest problem for start-ups. There are many new businesses that have gone over this issue. Predictive Machine Learning algorithms

can, to a greater extent, avoid any inaccuracies or errors caused because of data entries done manually. Machine Learning discovers duplicate or inaccurate data effectively. Thus, instead of digging the data to figure out the errors, the businessman and his staff can utilize that time to carry out fruitful tasks that would add value to his business.

- **Customer Behaviour Prediction**

Customer segmentation and customer behaviour prediction are the two biggest challenges among other challenges that the marketers of the current period face. Organizations reserve the option to get to a ton of information, which they can utilize successfully to determine critical business bits of knowledge. Data mining and Machine Learning can help start-ups predict complex client practices, their purchasing patterns. On the basis of their browsing as well as purchase histories, it helps the new companies in sending the best possible deals to the individual customers.

- **Product Recommendations**

Almost 80% of online business sites are utilizing ML to make item suggestions. The ML algorithms use the shopping history of the consumer and then match it with the vast product inventory for identifying hidden patterns as well as for grouping similar products together. Then the customers have suggested these products and motivated to purchase the product. Predicting more accurately “the related products” will help a start-up increase its revenue per consumer.

- **Measuring Effectively the Risk Factors**

Ascertaining risk isn't quite as simple as running a blade on the bar of softened margarine. Countless variables have to be considered and complex decisions have to be made by the managers. With the assistance of Machine Learning cheats, mistakes and other danger components can be perceived and estimated viably.

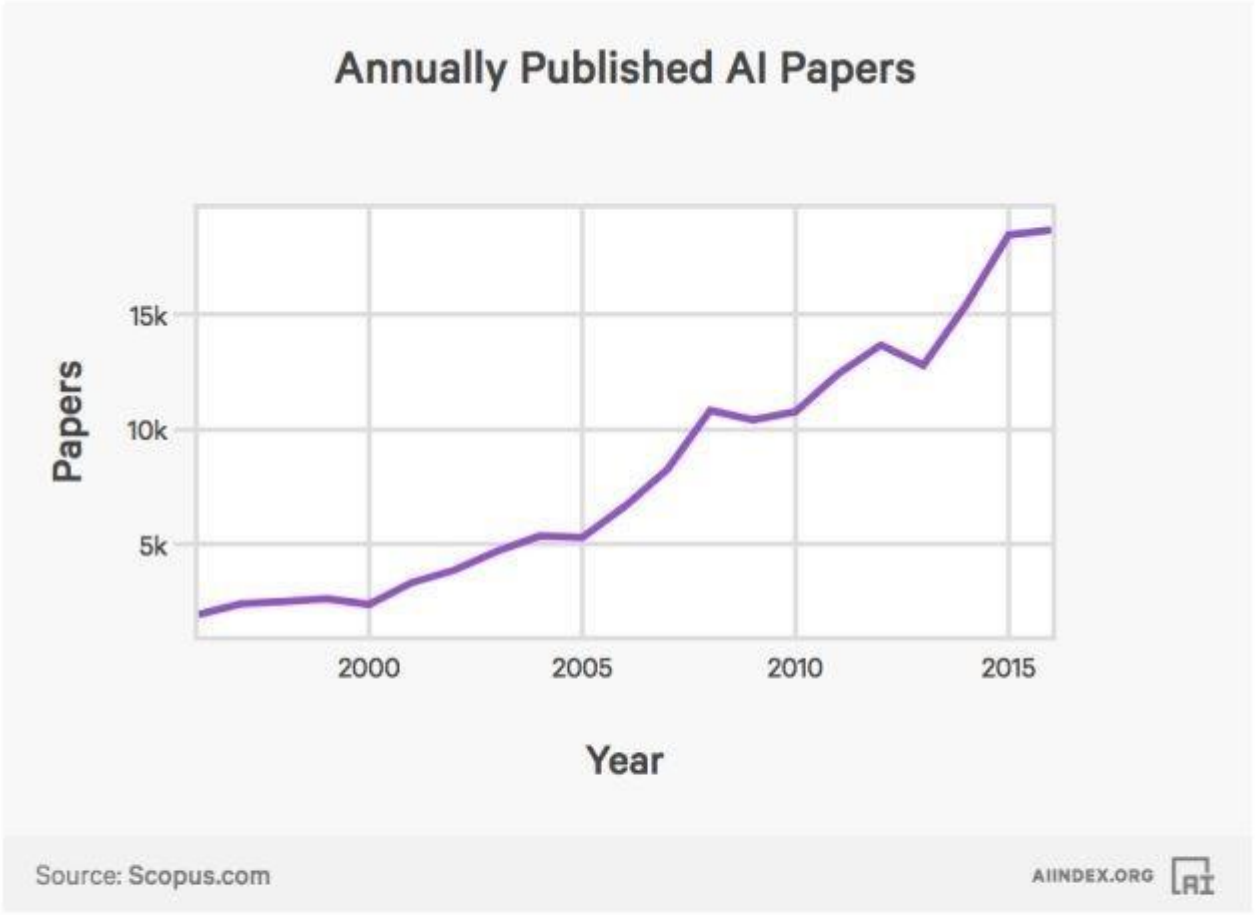


Figure 3.1: Scope of ML predicted by Past

CHAPTER 4: METHODOLOGY ADOPTED

After trying Immense technologies, I have came to the decision to use the Apriori and Association rule algorithm as I have got the best accuracy with the combination of these algorithms.

Used python and created the automatic code that can study the 15000 entries and shortlisted the approx. 10000 useful entries which helps us to get the best accuracy and do the predictions.

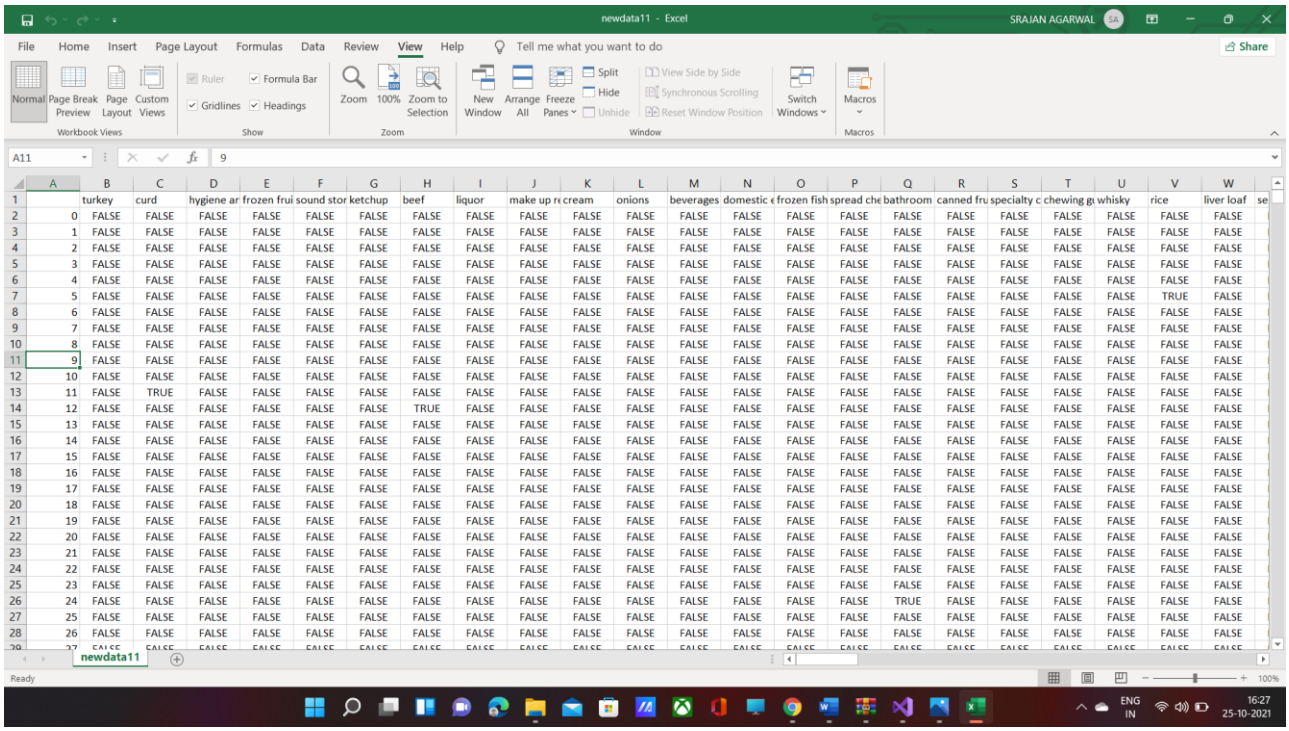


Table 4.1: Dataset

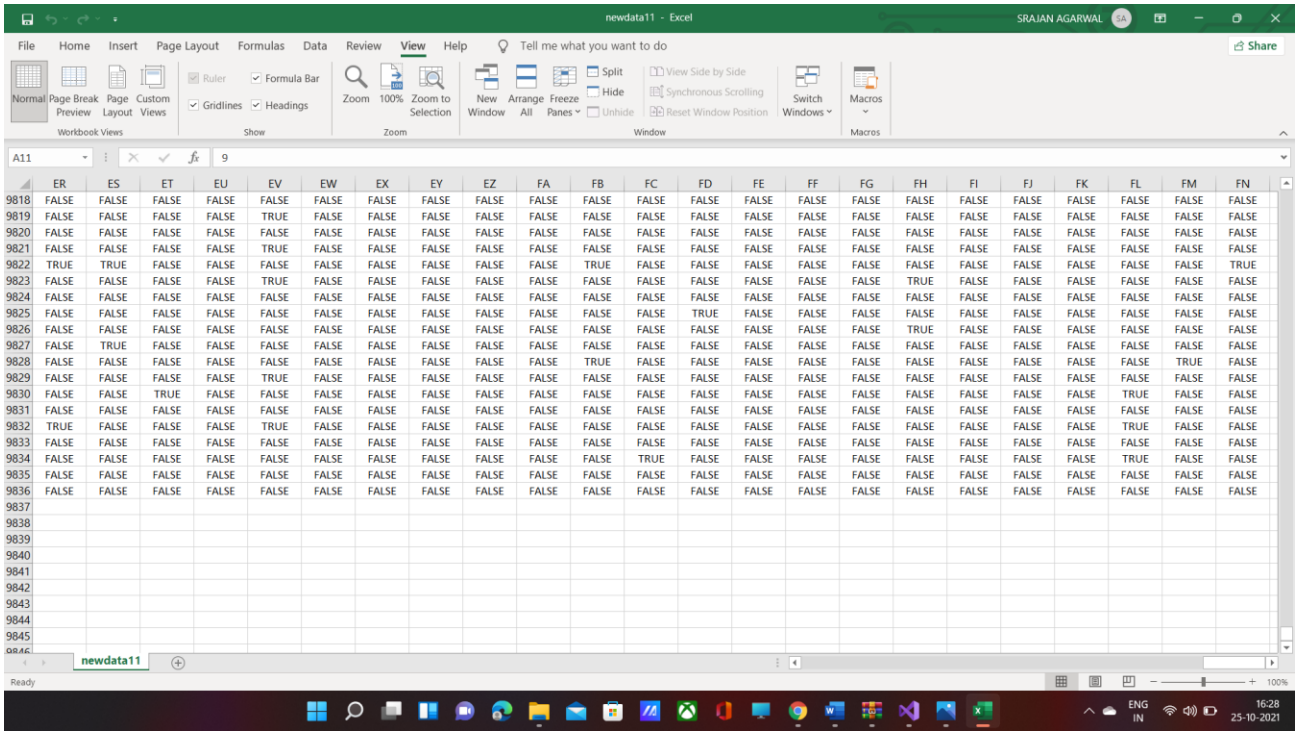
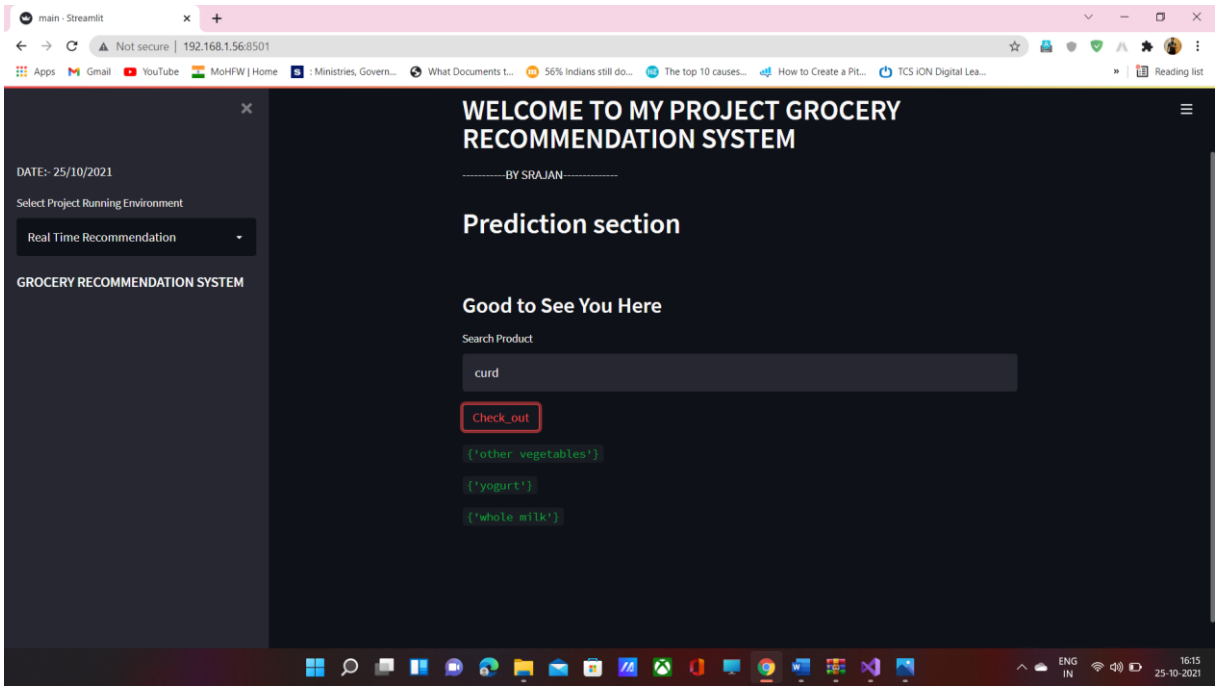


Table 4.2: Dataset

CHAPTER 5: RESULTS AND DISCUSSIONS

After the research on n no. of algorithms I have got the 93% accuracy, which claims that there are 93% chances that any customer who purchase any item from the seller's platform will also purchase at least 1 item from the suggestions provided by my model.

If customer bought at least 1 item from the suggestion then it will affect the monthly income of seller by increasing it by more then 10%.



Photograph 5.1: Output given by the model on entering the product.

CHAPTER 6 CONCLUSINS AND SCOPE FOR FUTURE STUDY

As per the expectations this system and my studies will help the sellers to survive in the market by increasing their sales as with this model, they will be able to sell their products to a right person who seriously wants to buy it.

In future we can work upon this model to increase its accuracy as well as it also became an independent grocery selling platform at discounted price as we know the target person so sales will automatically increase.

REFERENCES –

- **https://scikit-learn.org/stable/**
- **https://ml-cheatsheet.readthedocs.io/en/latest/**
- **https://www.netapp.com/artificial-intelligence/what-is-machine-learning/**

APPENDICES –

Reference for the ML algorithm has been taken from the documentation, similarly Streamlit is being used to develop the Front-End on the basis of the reference of the documentation.