**Machine Learning and Data Mining Model for Automated Computerized System**

**Content**

1. Introduction

2. Data Acquisition

3. Data Visualization and Preparation

4. Selecting a model

5. Fine Tuning the Model

6. Presenting Data

7. Binding to the Main Software

8. Monitoring and Improving the System

**Introduction to ML and DM Model**

Victa Towers is a startup Manufacturing Company working in chemical and herbal products such as washing powder, Cleaning solutions, tea etc. Whereas it has large scale distributions with only a small number of employees. As per this project we developers are trying to automate their manufacturing and distribution process without increasing the number of employees.

As the Machine learning and Data Mining Model of the project, development of a system to predict the future sales of the company will be developed. Here the development team will see

1. how much manufacturing of a given product is happening at the moment
2. How Much sales of a given product is happening at the moment

Then by using these two pieces of information the idea is to develop a system that will predict how much product should be produced in the upcoming few days, according to the sales that happened in the past.

By developing this system we are able to

1. Reduce the wastage of production
2. Reduce Employer Overworking nature
3. Economic Advantages
4. Identify the Trends Opportunities and Threats in the Distribution System
5. Identifying the Weaknesses and Strengths of the system

**2.Data Visualization and Preparation**

To identify what to produce and when to produce the products we visualize the sales according to the month and according to the day of the week, since we do not have a column for the day of the sales we are writing a program to read the data from the data column.

For Reading and Functionality of the CSV file, the library “Pandas” is used and as the coding environment “Google Colab” is used

Then we compare the attributes “Place”, ”Date”, ”Month” and “Time” for each product that we have at the hand to learn about what to sell on excess on a given day and what to not to manufacture on a given day

The Data File that was taken had 17 columns, for the easiness in processing that was reduced to 5 columns

1. City - City where the sale was done

2. Product - The product that was sold

3. Date - Date that the product was sold

4. Time - Time at which the product was sold

5. Rating - Rating that was given to the product by the customer

All the data records that had null or empty values have been dropped because putting in missing values will create a bias in this instance.

Then we exported this set of data to a new csv file to preserve the original file