

DemandEst - AI Powered Food Demand

Forecaster

# PROJECT REPORT

***Submitted by***

**ARUL.A 611619205006**

**ARUN KUMAR.A 611619205007**

**MUTHU BHARATHI.M 611619205026**

**VIDHYADHARAN.R 611619205057**

***in partial fulfilment for the award of the degree of***

**BACHELOR OF TECHNOLOGY IN INFORMATION TECHNOLOGY**

**MAHENDRA INSTITUTE OF TECHNOLOGY**

# (Autonomous)

**Mahendhirapuri,Mallasamudram**

**Namakkal- 637503**

**NOV 2022**

**MAHENDRA INSTITUTE OF TECHNOLOGY**

**(Autonomous)**

**Mahendrapuri,Mallasamudhram,**

**Namakkal-637503**

Department of Information technology

**BONAFIDE CERTIFICATE**

Certified of this project report **”DemandEst – AI PowerFood Demand Forecasater ” is bonafide work of “ ARUL.A (611619205026),ARUN KUMAR A(611619205007),MUTHU BHARATHI M(611619205026),VIDHYADHARAN R(611619205057) “** who carried out the project work under by supervisor.

**SIGNATURE SIGNATURE**

**Dr.N.Sathish.,M.E.,Ph.D., Mr.M.Premkumar,M.E.,**

**HEAD OF THE DEPARTMENT,** **SUPERVISIOR,**

Professor, Assistant Professor,

Department of Information Department of Information

Technology , Technology,

Mahendra Institute of Mahendra Institute of

Technology, Technology,

Namakkal-637503. Namakkal-637503.

**MAHENDRA INSTITUTE OF TECHNOLOGY**

**(Autonomous)**

**Mahendrapuri,Mallasamudhram,**

**Namakkal-637503**

Department of Information technology

**CERTIFICATE OF PROJECT APPROVAL**

This is to certify that the Project titled **“ DemandEst – AI Powered Demand Forecaster”ARUL.A(611619205006),ARUN KUMAR.A(611619205007),MUTHU BHARATHI.M (611619205026),VIDHYADHARAN.R(611619205057) “** in partial fulfillment for the award of the degree of Information Technology during the academic year 2022 – 2023.

**SUPERVISIOR HEAD OF THE DEPARTMENT**

(signature with seal) (signature with seal)

**Date:**

Submitted for the end semester viva vorce examination held on\_

**Internal Examiner External Examiner**

**ACKNOWLEDGEMENT**

We would like to take this opportunity to say our thanks to the people who have helped us make this project a reality.

We wish to express our sincere thanks to our honourable chairman , **Shri.M.G.Bharath Kumar B.Ed.,M.A., M.I.S.T.E.**, Of our Educational trust Kalipatty and the Managing Directors **Er.Ba.Mahendirann B.E.,** and **Er.Maha Ajay Prasad B.E.,** providing an extraordinary infrastructure.

We would like to express our sincere thanks to **Dr.Elango M.E.,Ph.D.,**the principal of our college,for their kind encouragement and blessings to do this project.

We also thank **Dr.N.Sathish M.E.,Ph.D.,** Head of the Department of Information Technology and Engineering for the encouragement,valuable suggestions and support in doing this project.

We would like to thank our internal guide **Mr.M.Prem Kumar M.E.,** Department of Information Technology and Engineering for the kind co -operative and support rendered in making our project as success.

We would like to say our science thanks to all other faculties,Department of Information Technology for their active and kind guidance and advices for our project

**ABSTRACT**

A food delivery service has to deal with a lot of perishable raw materials which makes it all,the most important factor for such a company is to accurately forecast daily and weekly demand.Too much inventory in the warehouse means more risk of wastage, and not enough could lead to out -of-stocks – and push customers to seek solutions from your competitors. The replenishment of majority of raw materials is done on weekly basis and since the raw material is perishable,the procurement planning is of utmost importamne,the task is to predict the demand for the next 10 weeks. The main aim of this project is to create an appropriate machine learning model to forecast the number of orders to gather raw materials for next ten weeks. To achieve this, we should know the information about of fulfilment center like area, city etc., and meal information like category of food sub category of food price of the food or discount in particular week. By using this data, we can use any classification algorithm to forecast the quantity for 10 weeks. A web application is built which is integrated with the model built.

**CONTENTS**

# INTRODUCTION

* 1. Project Overview
  2. Purpose

# LITERATURE SURVEY

* 1. Existing problem
  2. References
  3. Problem Statement Deﬁnition

# IDEATION & PROPOSED SOLUTION

* 1. Empathy Map Canvas
  2. Ideation & Brainstorming
  3. Proposed Solution
  4. Problem Solution ﬁt

# REQUIREMENT ANALYSIS

* 1. Functional requirement
  2. Non-Functional requirements

# PROJECT DESIGN

* 1. Data Flow Diagrams
  2. Solution & Technical Architecture
  3. User Stories

# PROJECT PLANNING & SCHEDULING

* 1. Sprint Planning & Estimation
  2. Sprint Delivery Schedule
  3. Reports from JIRA

# CODING & SOLUTIONING

* 1. Data Dictionary
  2. Libraries Used
  3. Data Pre-Processing
  4. Feature Engineering
  5. Evaluation Metric
  6. Initial Approach
  7. Advanced Models

# TESTING

* 1. Test Cases
  2. User Acceptance Testing

# RESULTS

* 1. Performance Metrics

# ADVANTAGES & DISADVANTAGES

1. **APPLICATIONS**

# CONCLUSION

1. **FUTURE SCOPE**

# APPENDIX

Source Code Output Screenshots

GitHub & Project Demo Link

# INTRODUCTION

* 1. **OVERVIEW**

A food delivery service has to deal with a lot of perishable raw materials which makes it all, the most important factor for such a company is to accurately forecast daily and weekly demand. Too much inventory in the warehouse means more risk of wastage, and not enough could lead to out-of-stocks - and push customers to seek solutions from your competitors. The replenishment of the majority of raw materials is done on weekly basis and since the raw material is perishable, the procurement planning is of utmost importance, the task is to predict the demand for the next 10 weeks.

# PURPOSE

The main aim of this project is to create an appropriate machine learning model to forecast then number of orders to gather raw materials for next ten weeks. To achieve this, we should know the information about of fulﬁllment center like area, city etc., and meal information like category of food, sub category of food, price of the food or discount in particular week. By using this data, we can use any classiﬁcation algorithm to forecast the quantity for 10 weeks. For this a web application is built which is integrated with the model.

# LITERATURE SURVEY

* 1. **EXISTING PROBLEM**

The replenishment of the majority of raw materials is done on weekly basis and since the raw material is perishable, the procurement planning is of utmost importance. Also the recruiting of staff members at the fulﬁllment center is an prospect wherein the prediction of orders would be beneﬁcial. Although this is a process that can be done manually.

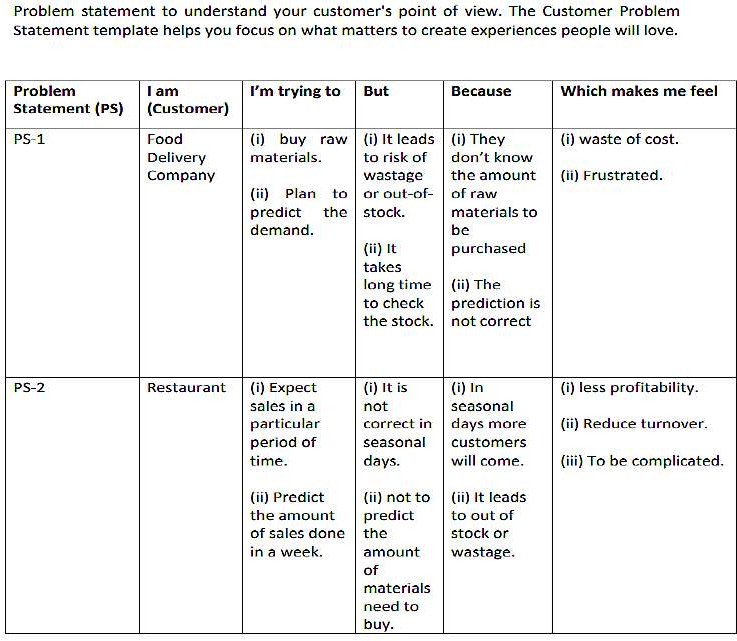
# REFERENCE

* + - Adi, G. N. (2018, March 9). Thousands of GO-CAR Drivers on Strike in Surakarta. The Jakarta Post. https:/[/www.thejakartapost.com/news/2018/03/08/thousands-of-go-car-](http://www.thejakartapost.com/news/2018/03/08/thousands-of-go-car-)

drivers-on-strike-in-surak arta.html

* + - Alkhatib, A., & Bernstein, M. (2019, May). Street-level algorithms: A theory at the gaps between policy and decisions. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-13).
    - Brown, Tim. (2008). Design Thinking. Harvard Business Review. 86. 84-92, 141.
    - Colley, A., & Häkkilä, J. (2018, November). Service Design Methods for Human Computer Interaction. In Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (pp. 563-566).
    - Clarke, S. (2006). Transformation Lessons from Coca-Cola Enterprises Inc.: Managing the Introduction of a Structured Forecast Process. Foresight: The International Journal of Applied Forecasting, (4), 21-25.

# PROBLEM STATEMENT DEFINITION

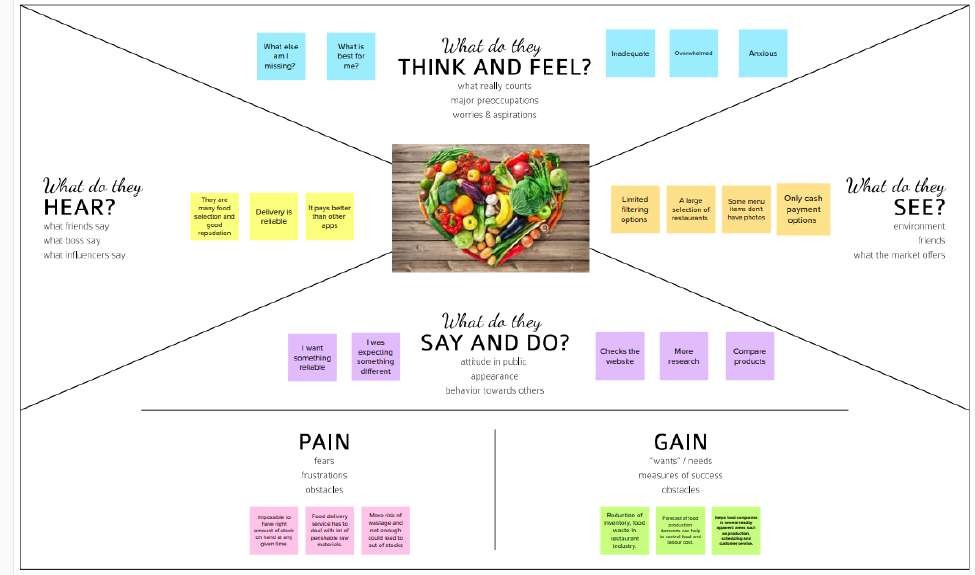


1. **IDEATION & PROPOSED SOLUTION**

# Empathy Map Canvas

An empathy map is a collaborative visualization used to articulate what we know about a particular type of user. It externalizes knowledge about users in order to 1) create a shared understanding of user needs, and 2) aid in decision making.

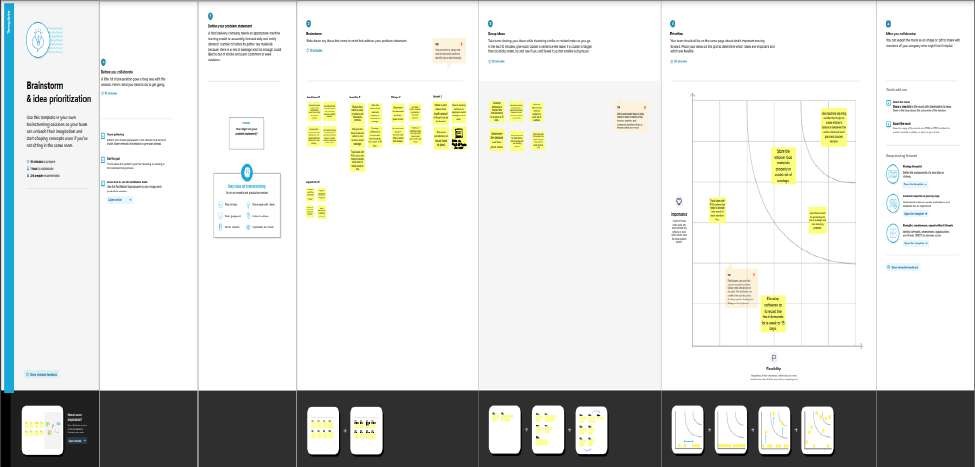
Traditional empathy maps are split into 4 quadrants (Says, Thinks, Does, and Feels), with the user or persona in the middle. Empathy maps provide a glance into who a user is as a whole and are not chronological or sequential.



# Ideation & Brainstorming

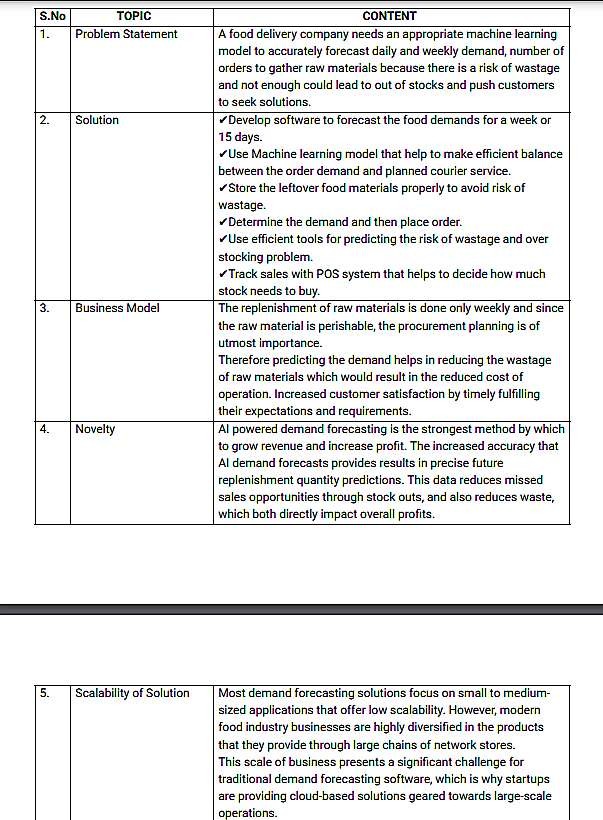
Brainstorming is a method design teams use to generate ideas to solve clearly deﬁned design problems. In controlled conditions and a free-thinking environment, teams approach a problem by such means as “How Might We” questions. They produce a vast array of ideas and draw links between them to ﬁnd potential solutions

Everyone in a design team should have a *clear* deﬁnition of the target problem. They typically gather for a brainstorming session in a room with a large board/wall for pictures/Post-Its. A good mix of participants will expand the experience pool and therefore broaden the idea space.



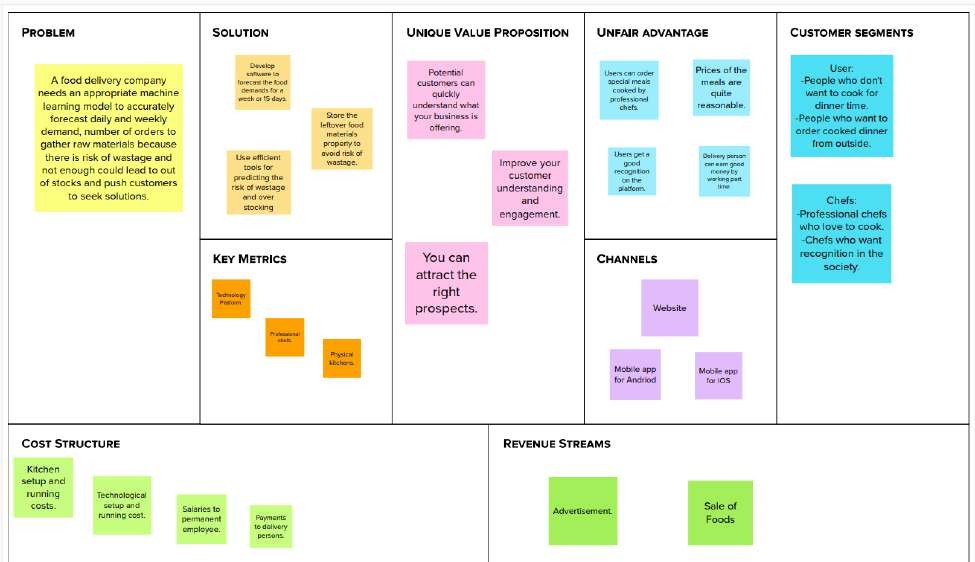
# 3.2 Proposed Solution

[Proposed Solution](https://www.lawinsider.com/dictionary/proposed-solution) means the technical solution to be provided by the Implementation agency in response to the requirements and the objectives of the Project.[Proposed Solution](https://www.lawinsider.com/dictionary/proposed-solution) means the Proposed System with modiﬁcations that meet the Agency’s requirements as set forth in this RFP.[Proposed Solution](https://www.lawinsider.com/dictionary/proposed-solution) means the combination of software, hardware, other products or equipment, and any and all services (including any installation, implementation, training, maintenance and support services) necessary to implement the solution described by Vendor in its Proposal.



# 3.4 Problem Solution Fit

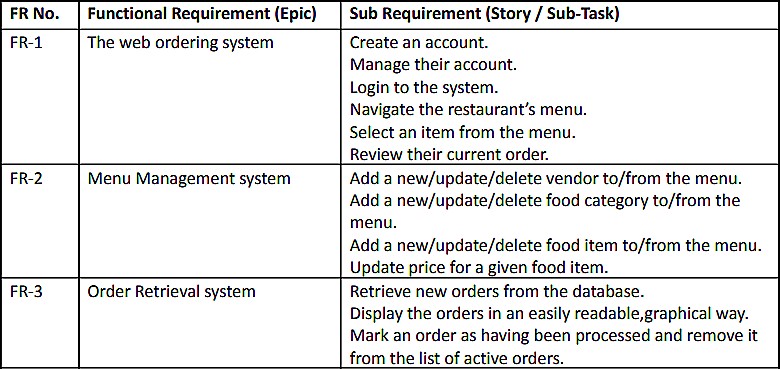
The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem.



# REQUIREMENT ANALYSIS

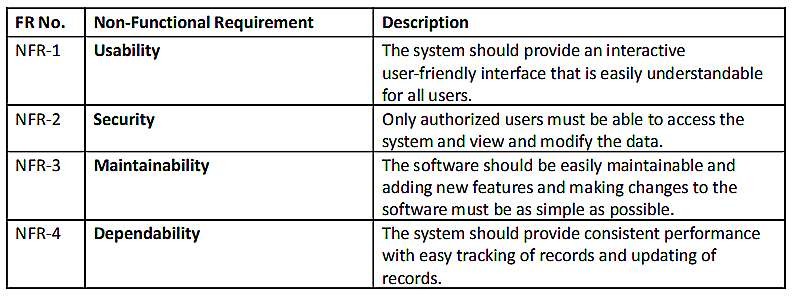
* 1. **Functional requirement**

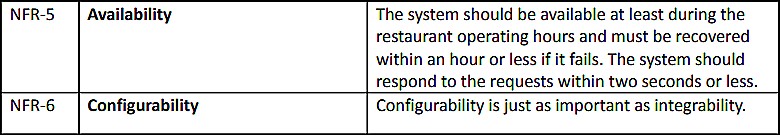
Functional requirements may involve calculations, technical details, data manipulation and processing, and other speciﬁc functionality that deﬁne what a system is supposed to accomplish. Behavioral requirements describe all the cases where the system uses the functional requirements, these are captured in use cases.



# Non-Functional requirements

In [systems engineering](https://en.wikipedia.org/wiki/Systems_engineering) and [requirements engineering](https://en.wikipedia.org/wiki/Requirements_engineering), a non-functional requirement (NFR) is a [requirement](https://en.wikipedia.org/wiki/Requirement) that speciﬁes criteria that can be used to judge the operation of a system, rather than speciﬁc behaviours.

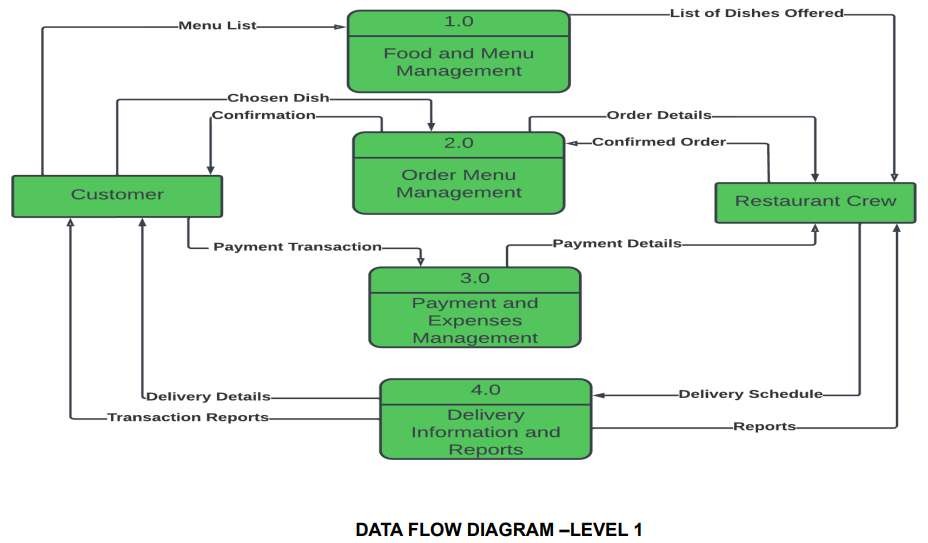


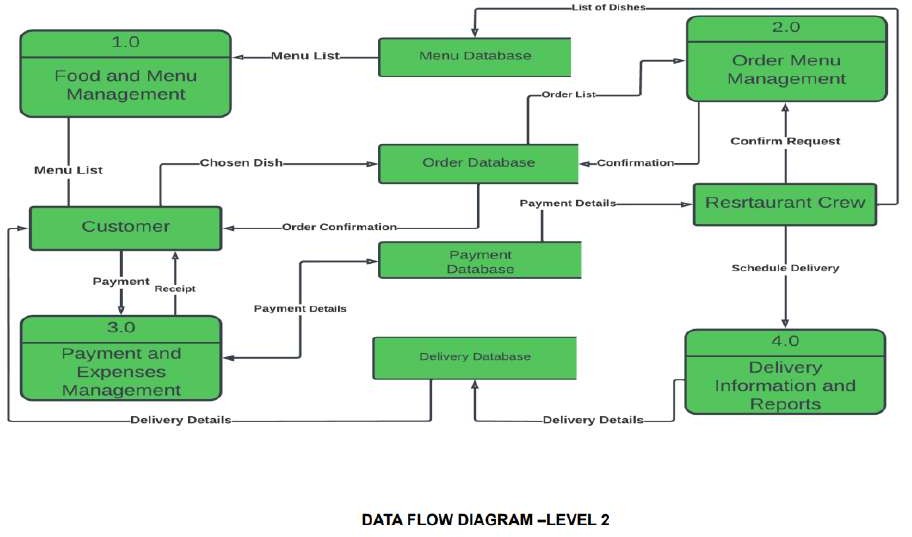


# PROJECT DESIGN

* 1. **Data Flow Diagrams**

A Data Flow Diagram (DFD) is a traditional visual representation of the information ﬂows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



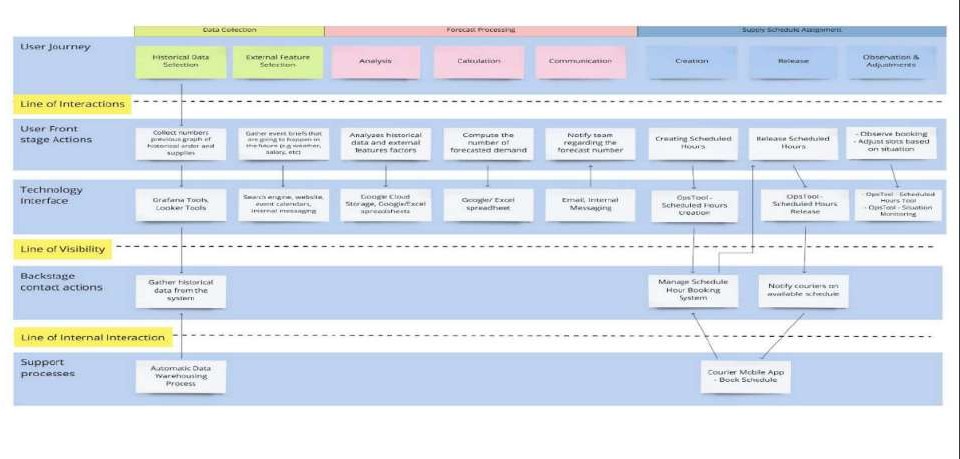


# Solution & Technical Architecture

Solution Architecture:

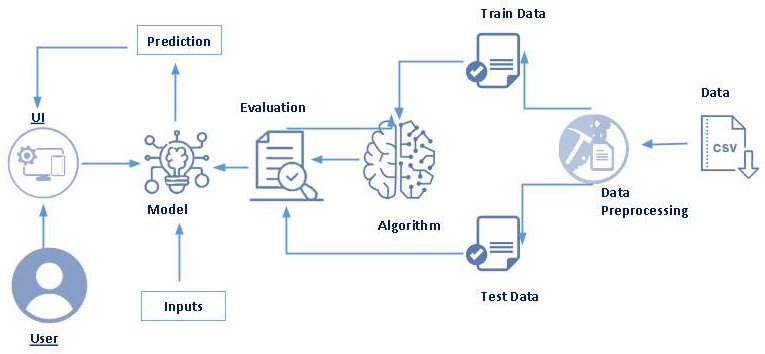
A solution architecture (SA) is an architectural description of a speciﬁc solution.

SAs combine guidance from different enterprise architecture viewpoints (business, information and technical), as well as from the enterprise solution architecture (ESA).



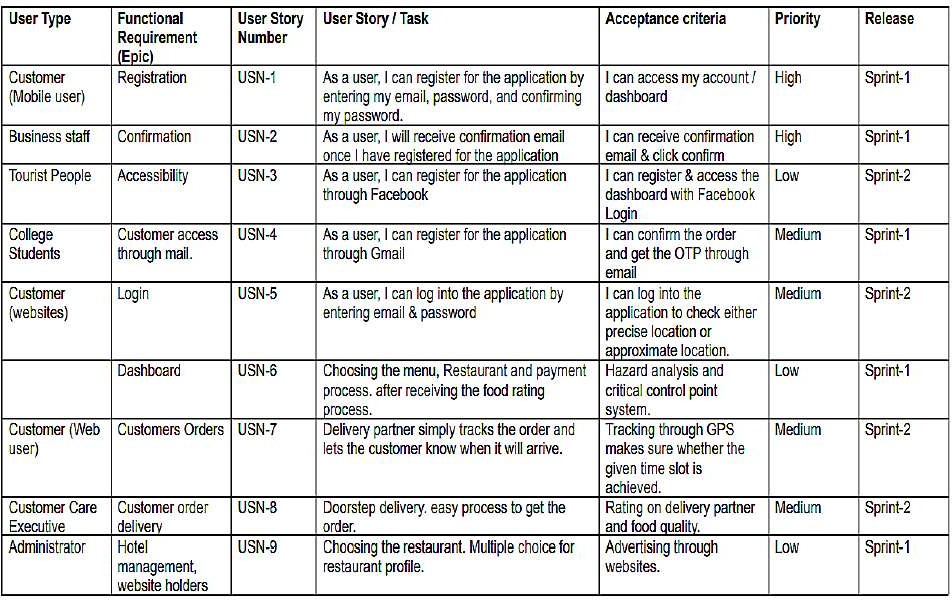
Technical Architecture:

Technical Architecture (TA) is a form of IT architecture that is used to design computer systems. It involves the development of a technical blueprint with regard to the arrangement, interaction, and interdependence of all elements so that system- relevant requirements are met.



# User Stories

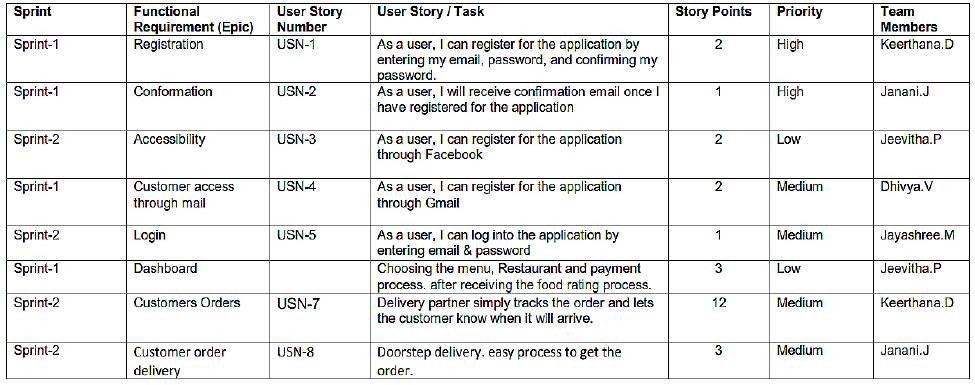
A user story is an informal, general explanation of a software feature written from the perspective of the end user or customer. The purpose of a user story is to articulate how a piece of work will deliver a particular value back to the customer.

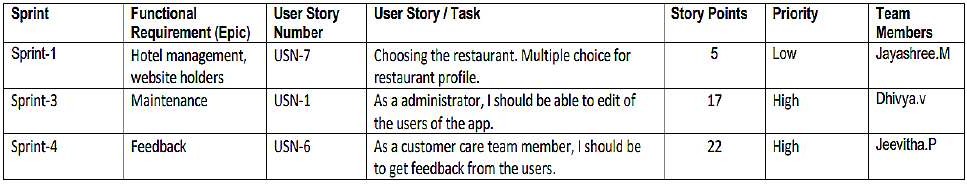


# PROJECT PLANNING & SCHEDULING

* 1. **Sprint Planning & Estimation**

In Scrum Projects, Estimation is done by the entire team during Sprint Planning Meeting. The objective of the Estimation would be to consider the User Stories for the Sprint by Priority and by the Ability of the team to deliver during the Time Box of the Sprint.





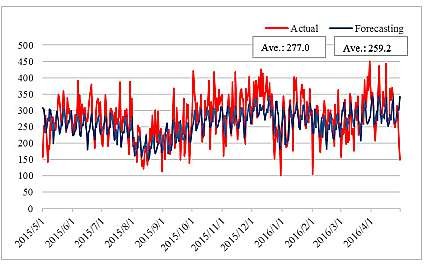
# Sprint Delivery Schedule

A sprint schedule is a document that outlines sprint planning from end to end. It's one of the ﬁrst steps in the agile sprint planning process—and something that requires adequate research, planning, and communication.



# Reports From JIRA

Jira helps teams plan, assign, track, report, and manage work and brings teams together for everything from agile software development and customer support to start- ups and enterprises. Software teams build better with Jira Software, the #1 tool for agile teams.



# CODING & SOLUTIONING

* 1. **Data Dictionary**

Our base data consists of four csv ﬁles containing information about test data, train data and other required information.

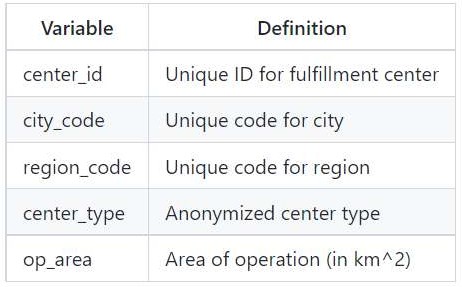
* + - train.csv: Contains information like id, week, center id, meal id, checkout price, base price, emailer for promotion, homepage featured, number of orders. This ﬁle is used for training.



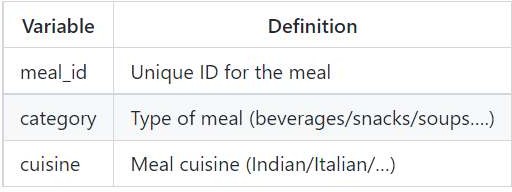
* + - test.csv: Contains information like id, week, center id, meal id, checkout price,

base price, emailer for promotion, homepage featured. This ﬁle is used for testing.

* + - fulﬁlment\_center\_info.csv: Contains information of each fulﬁlment center.



* + - meal\_info.csv: Contains information of each meal being served.



# Libraries Used

pandas, numpy, scikit learn, matplotlib, seaborn, xgboost, lightgbm, catboost

# Data Pre-Processing

* + - There are no Missing/Null Values in any of the three datasets.
    - Before proceeding with the prediction process, all the three data sheets need to be merged into a single dataset. Before performing the merging operation, primary feature for combining the datasets needs to be validated.
    - The number of Center IDs in train dataset is matching with the number of Center IDs in the Centers Dataset i.e 77 unique records. Hence, there won't be any missing values while merging the datasets together.
    - The number of Meal IDs in train dataset is matching with the number of Meal IDs in the Meals Dataset i.e 51 unique records. Hence, there won't be any missing values while merging the datasets together.
    - As checked earlier, there were no Null/Missing values even after merging the datasets.

# Feature Engineering

Feature engineering is the process of using domain knowledge of the data to create features that improves the performance of the machine learning models.

With the given data, We have derived the below features to improve our model performance.

* + - Discount Amount : This deﬁnes the difference between the “base\_Price” and “checkout\_price”.
    - Discount Percent : This deﬁnes the % discount offer to customer.
    - Discount Y/N : This deﬁnes whether Discount is provided or not - 1 if there is Discount and 0 if there is no Discount.
    - Compare Week Price : This deﬁnes the increase / decrease in price of a Meal for a particular center compared to the previous week.
    - Compare Week Price Y/N : Price increased or decreased - 1 if the Price increased and 0 if the price decreased compared to the previous week.
    - Quarter : Based on the given number of weeks, derived a new feature named as Quarter which deﬁnes the Quarter of the year.
    - Year : Based on the given number of weeks, derived a new feature named as Year which deﬁnes the Year.

# Data Transformation

* + - Logarithm transformation (or log transform) is one of the most commonly used mathematical transformations in feature engineering. It helps to handle skewed data and after transformation, the distribution becomes more approximate to normal.
    - In our data, the target variable ‘num\_orders’ is not normally distributed. Using this without applying any transformation techniques will downgrade the performance of our model.
    - Therefore, we have applied Logarithm transformation on our Target feature ‘num\_orders’ post which the data seems to be more approximate to normal distribution.
    - After Log transformation, We have observed 0% of Outlier data being present within the Target Variable – num\_orders using 3 IQR Method.

# Evaluation Metric

The evaluation metric for this competition is 100\*RMSLE where RMSLE is Root of Mean Squared Logarithmic Error across all entries in the test set.

# Initial Approach

* + - Simple Linear Regression model without any feature engineering and data transformation which gave a RMSE : 194.402
    - Without feature engineering and data transformation, the model did not perform well and could'nt give a good score.
    - Post applying feature engineering and data transformation (log and log1p transformation), Linear Regression model gave a RMSLE score of 0.634.

# Advanced Models

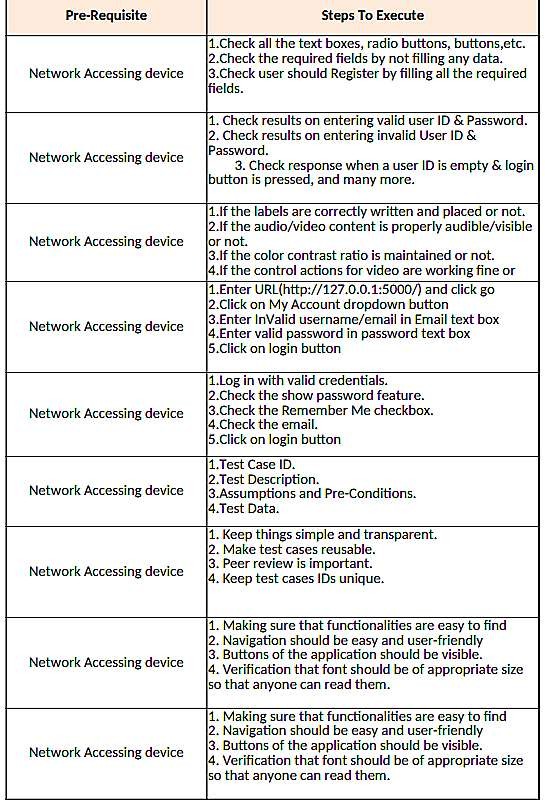
* + - With improvised feature engineering, built advanced models using Ensemble techniques and other Regressor algorithms.
    - Decision Tree Regressors performed well on the model which gave much reduced RMSLE.
    - With proper hyper-parameter tuning, Decision Tree Regressor performed well on the model and gave the lease RMSLE of 0.5237

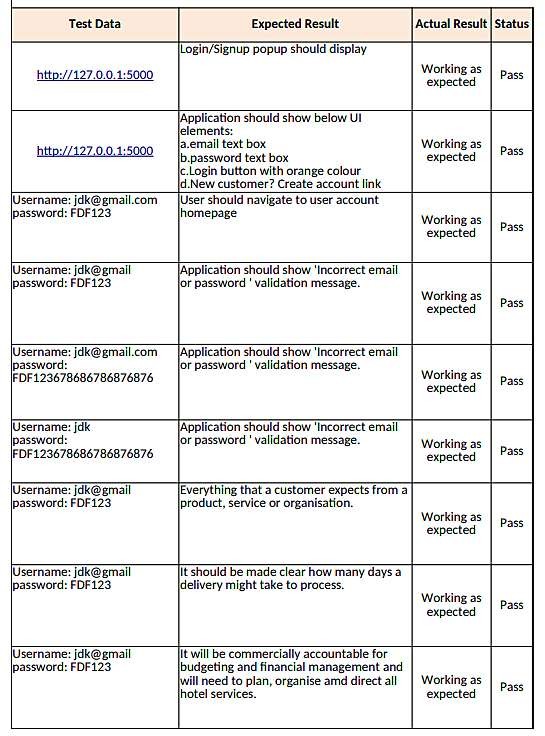
# TESTING

* 1. **Test Cases**

A test case includes information such as test steps, expected results and data while a test scenario only includes the functionality to be tested.

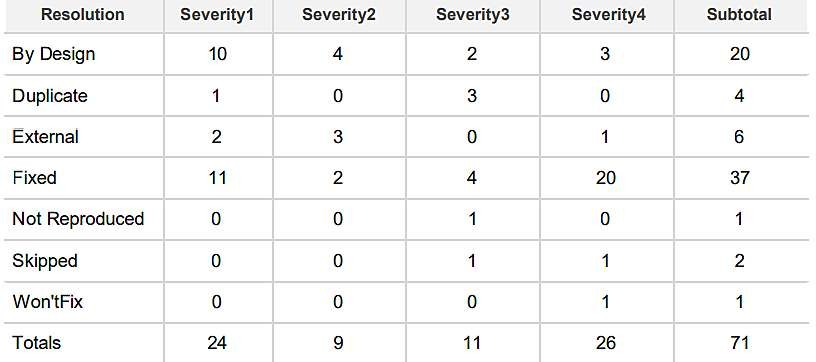




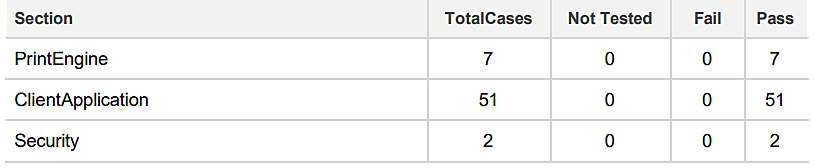
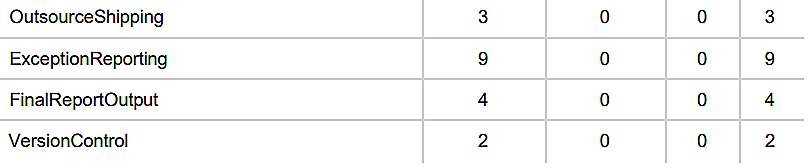


# User Acceptance Testing

User Acceptance Testing (UAT), which is performed on most UIT projects, sometimes called beta testing or end-user testing, is a phase of software development in which the software is tested in the "real world" by the intended audience or business representative.

Defect Analysis:

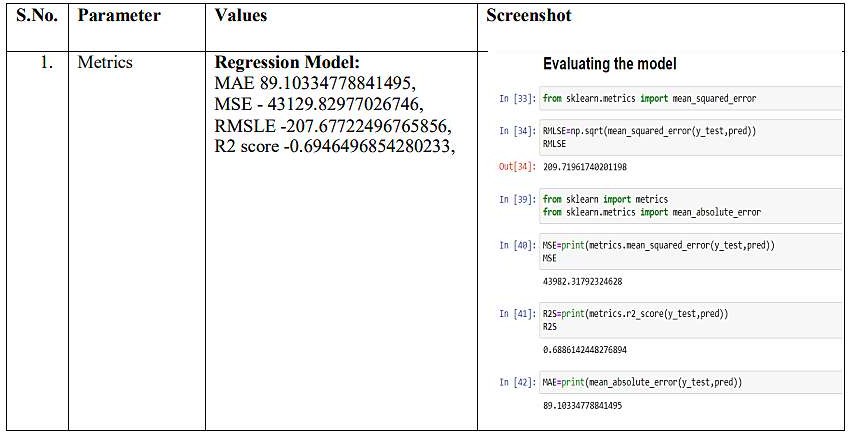
Test Case Analysis:

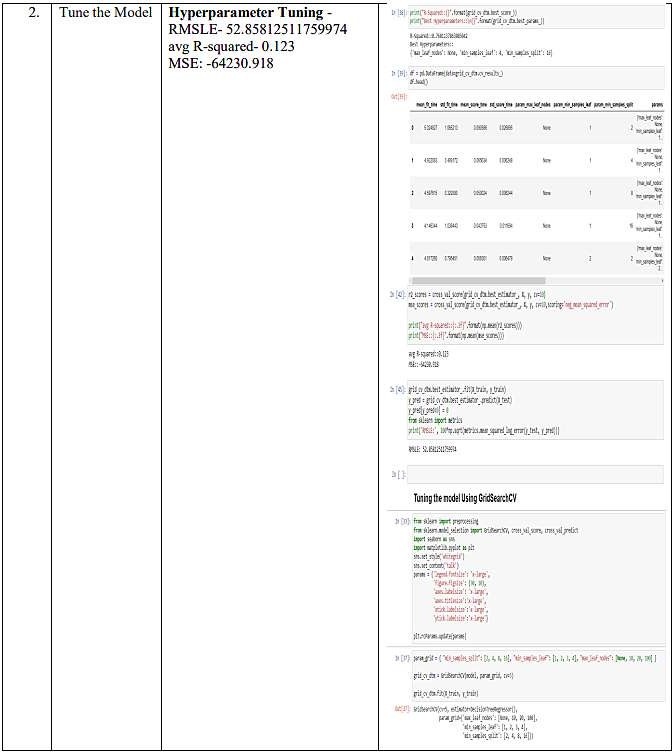


# RESULTS

* 1. **Performance Metrics**

Performance testing is the practice of evaluating how a system performs in terms of responsiveness and stability under a particular workload. Performance tests are typically executed to examine speed, robustness, reliability, and application size.





# ADVANTAGES & DISADVANTAGES

**Advantages:**

1. Food wastage will be minimized.
2. Simple and easy to use framework.

# Disadvantages:

1. The output obtained may not be precised, due to the use of limited datasets.

# APPLICATIONS

This project focuses on one food delivery client, which delivers food in many different cities through distribution networks and fulﬁllment centers.

# CONCLUSION

The main moto behind this project is to reduce food wastage.The availability of the food items makes the society better. Our purposed model would deﬁnitely come handy to a company for predicting then number of food orders and help them to serve their customers better.

# FUTURE SCOPE

* 1. Working on the frontend to make the framework more dynamic.
  2. In the future, we also plan to improve forecasting accuracy and research on the eﬃciency of store management.

# APPENDIX

**SOURCE CODE**:

# home.html

<!DOCTYPE html>

<html>

<head>

<title>Home</title>

<style>

.navbar

{

margin: 0px; padding:20px; background-color:white; opacity:0.6;

color:black;

font-family:'Roboto',sans-serif; font-style: italic;

border-radius:20px; font-size:25px;

}

a

{

color:grey; ﬂoat:right;

text-decoration:none; font-style:normal; padding-right:20px;

}

a:hover{

background-color:black; color:white;

border-radius:15px;0 font-size:30px; padding-left:10px;

}

p

{

color:white; font-style:italic; font-size:30px;

}

body

{

background-image: url("https://img.freepik.com/free-photo/grilled-chicken-rice-spicy- chickpeas-avocado-cabbage-pepper-buddha-bowl-dark-top-view\_127032- 1966.jpg?w=2000");

background-size: cover;

}

</style>

</head>

<body>

<div class="navbar">

<a href="/pred">Predict</a>

<a href="/home">Home</a>

<br>

</div>

<br>

<center><b><font color="yellow" size="15" font-family="Comic Sans MS" >Food Demand Forecasting</font></b></center>

<div>

<br>

<center>

<p>A food delivery service has to deal with a lot of perishable raw materials which makes it all, the most important factor for such a company is to accurately forecast daily and weekly demand. Too much inventory in the warehouse means more risk of wastage, and not enough could lead to out-of-stocks - and push customers to seek solutions from your competitors. The replenishment of majority of raw materials is done on weekly basis and since the raw material is perishable, the procurement planning is of utmost importance, the task is to predict the demand for the next 10 weeks.</p>

</center>

</div>

</body>

</html>

# upload.html

<html lang="en">

<head>

<title>Predict</title>

<link href="https://cdn.bootcss.com/bootstrap/4.0.0/css/bootstrap.min.css" rel="stylesheet">

<style>

.bar

{

margin: 0px; padding:20px; background-color:white; opacity:0.6;

color:black;

font-family:'Roboto',sans-serif; font-style: italic;

border-radius:20px; font-size:25px;

}

a

{

color:red; ﬂoat:right;

text-decoration:none; font-style:normal;

padding-right:20px;

}

a:hover{

background-color:black; color:white;

border-radius:15px;0 font-size:30px; padding-left:10px;

}

body

{

background-image: url("https://images.pexels.com/photos/1640777/pexels-photo- 1640777.jpeg?cs=srgb&dl=pexels-ella-olsson-1640777.jpg&fm=jpg");

background-size: cover;

}

p

{

color:white; font-style:italic; font-size:30px;

}

h1,h2

{

color:0101DF;

}

</style>

</head>

<body>

<div class="bar">

<a href="/pred">Predict</a>

<a href="/home">Home</a>

<br>

</div>

<div class="container">

<center> <div id="content" style="margin-top:2em">

<h2><center>Food Demand Forecasting</center></h2>

<form action="{{ url\_for('predict') }}" method="POST">

<select id="homepage\_featured" name="homepage\_featured">

<option value="">homepage\_featured</option>

<option value="0">No</option>

<option value="1">Yes</option>

</select><br><br>

<select id="emailer\_for\_promotion" name="emailer\_for\_promotion">

<option value="">emailer\_for\_promotion</option>

<option value="0">No</option>

<option value="1">Yes</option>

</select><br><br>

<input class="form-input" type="text" name="op\_area" placeholder="Enter the op\_area(2-7)"><br><br>

<select id="cuisine" name="cuisine">

<option value="">Cuisine</option>

<option value="0">Continental</option>

<option value="1">Indian</option>

<option value="2">Italian</option>

<option value="3">Thai</option>

</select><br><br>

<select id="city\_code" name="city\_code">

<option value="">City Code</option>

<option value="590">590</option>

<option value="526">526</option>

<option value="638">638</option>

<option value="others">Others</option>

</select><br><br>

<select id="region\_code" name="region\_code">

<option value="">Region Code</option>

<option value="23">23</option>

<option value="34">34</option>

<option value="35">35</option>

<option value="56">56</option>

<option value="71">71</option>

<option value="77">77</option>

<option value="85">85</option>

<option value="93">93</option>

</select><br><br>

<select id="category" name="category">

<option value="">Category</option>

<option value="0">Beverages</option>

<option value="1">Biryani</option>

<option value="2">Desert</option>

<option value="3">Extras</option>

<option value="4">Fish</option>

<option value="5">Other Snacks</option>

<option value="6">Pasta</option>

<option value="7">Pizza</option>

<option value="8">Rice Bowl</option>

<option value="9">Salad</option>

<option value="10">Sandwich</option>

<option value="11">Seafood</option>

<option value="12">Soup</option>

<option value="13">Starters</option>

</select><br><br>

<input type="submit" class="my-cta-button" value="Predict">

</form>

</center>

<br>

<h1 class="predict">Number of orders: {{ prediction\_text }}</h1>

</div>

</div>

</body>

</body>

# app.py

# import the necessary packages import pandas as pd

import numpy as np import pickle import os

from ﬂask import Flask,request, render\_template app=Flask( name ,template\_folder="templates") @app.route('/', methods=['GET'])

def index():

return render\_template('home.html') @app.route('/home', methods=['GET']) def about():

return render\_template('home.html') @app.route('/pred',methods=['GET'])

def page():

return render\_template('upload.html') @app.route('/predict', methods=['GET', 'POST']) def predict():

print("[INFO] loading model...")

model = pickle.load(open('fdemand.pkl', 'rb')) input\_features = [ﬂoat(x) for x in request.form.values()] features\_value = [np.array(input\_features)] print(features\_value)

features\_name = ['homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category']

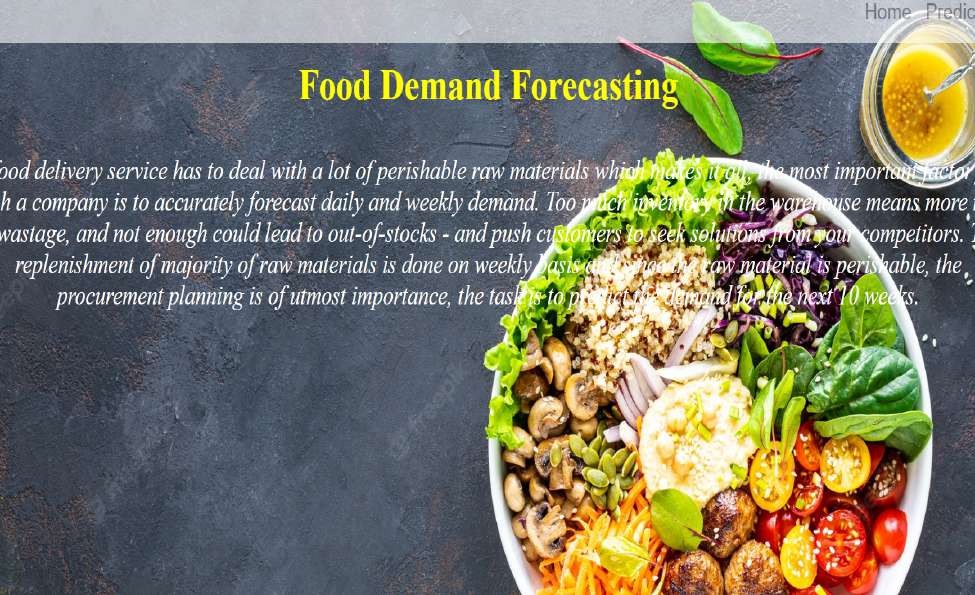
prediction = model.predict(features\_value) output=prediction[0]

print(output)

return render\_template('upload.html', prediction\_text=output)

if name == ' main ': app.run(debug=False)

# OUTPUT SCREENSHOTS:





**GITHUB LINK:**

https://github.com/IBM-EPBL/IBM-Project-30037-1660138781.git

# PROJECT DEMO LINK:

https://kapwi.ng/w/UtwvaoGVGU