RECIPE RECOGNITION USING IBM WATSON

Developed by:Md.tajuddin,K.subash vardhan

Y.vamshi,B.shivakumar

INDUSTRIAL ORIENTED MINI PROJECT:SMART BRIDGE

1. INTRODUCTION

1.1.MOTIVATION:

Food preferences around the World are changing specially in large cities due to factors like increasing income, migration and cultural mixing, appearance of new food fads and life-styles. These dietary changes often imply changes in the quantity and/or types of specific ingredients demanded.The demand for ingredients, their processing and supplying to consumers often happens in specific locations.Understanding their relevance in recipes used in specific locations can potentially inform design of interventions to improve, among others, nutrition and health, environmental and urban planning. Inorder to characterize ingredients used in specific locations, innovative tools are being used around theWorld to generate location-specific datasets.

1.2.PROBLEM DEFINITION:

Recipe prediction is mainly used in the restaurents big hotels to know the customer choice of the recipe   and comfort to deliver the exact their choice of recipe and deliver the food item with exact recipe contain

PROJECT OBJECTIVE:

The main objective of this project is to analyze the food item like Pizza,Samosa

and Frenchfries and to know their exact recipe by recipe prediction and gain broad

understanding about image data and user can predict the food recipe by choosing

the image in web application.

1.1.LIMITATIONS OF PROJECT:

Mixed dishes or multi-ingredient foods represents the majority of items in diets worldwide.These include not only foods prepared in the home but also foods prepared in the

home but also foods prepared in restaurants,by food vendors,in institutions such as hospitals,schools,and the military,and by the food industry.To enable dietitians,nutritionists,and epidemiologists to evalute the role of these foods in the health of individuals,there is a need for composition data on these foods.Obtaining and using data on the content of multi-ingredient foods present a number of inherent difficulties,primarily because of the abundance and diversity of these kinds of foods.

ORGANIZATION OF DOCUMENTATION:

  Actually there has been many theoretical projects and several experimental projects individually done based on liver patient analysis and many different algorithms have been developed for forecasting the recipe prediction. But Machine Learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Our aim from the project is to make use of NumPy or pandas libraries in Machine Learning and predict the food recipe choosen in the web application. It has own to visualize the data by using graphs. In the end, we are predicting the food recipe like samosa,pizza and French fries. The prediction is to be done using Machine Learning algorithms and withdrawing the conclusions.

2. PROBLEM STATEMENT

The food is more important to live man. And the food contains more recipes that shows the food is more interest to man how much he love the food when he eat.but there is some problem when taking the food it contains more demand in market business. It contains more multi-ingrient and it presents a number of inherent difficulties,primarily because of the abundance and diversity of these kinds foods.

There are some statements about recipe predictions:

Samosa: a samosa or singara is a fried or baked pastry with a savory filling,including ingredients suchbas spiced potatoes,onions,and peas. It may take different forms,including triangular,cone,or half-moon shapes,depending on the region.

Health issues: No,this recipe is not healthy.The samosa are deep fried which does not work for healthy living and it has a stuffing of potatoes.your fat levels increase as deep frying increase oil absorption.

Pizza: Is a dish of Italian origin consisting of a usually round,flat base of leavened wheat-based dough topped with tomatoes,cheese,and often various other ingredients, which is then baked at a high temperature , traditionally in a wood-fired oven.A small is some times called a pizzete.

Health issue: And it isn't health.Depending on the type of crust,the amount of cheese and the toppings used,pizza can rank any where from nutrionally decent to a diet disaster. Even healthy pizzas deliver a good amount of sodium from tomato sause and cheese,so if you are watching your salt intake ,you should eat with caution.

French fries:French fries are served hot, either soft or crispy, and are generally eaten as part of lunch or dinner or by themselves as a snack, and they commonly appear on the menus of diners, fast food restaurants, pubs, and bars.

They are often salted and may be served with ketchup, vineger, mayonnaise, tomato sauce, or other local specialities. Fries can be topped more heavily, as in the dishes of poutine or chillicheese fries. French fries can be made from sweet potatoes instead of potatoes. A baked variant, oven fries, uses less or no oil.

Health issues: Given the fact that they're deep-fried in hydrogenated oils, fries come packed with a high amount of trans fat, which raises your bad cholesterol and lowers your good cholesterol. The profound effect of this is that you'll boost your heart disease risk.

3. LITERATURE SURVEY

3.1. INTRODUCTION:

Nowadays, with easy access to internet, food is delivered at our doorsteps just on the click of a button due to which people have started to consume higher amount of fast food. This has accelerated the chances of suffering from a chronic disease known as obesity. Since obesity has become such a widespread disease, various mobile e-health applications have been developed for assistive calorie measurement to help people fight against health-related problems.

- Stomach pain
- Food poison
- Over fat
- Obesity
- Diabeties
- Heart stroke
- Face problem
- Loss of nutrition
- Hair issues

EXISTING SYSTEM:

The purpose of the present model is to predict the whether a the food contain coorect recipe offer by the customer.The model examines data from food recipe

concentrating on relationships between a key list of food items taste,recipe and variety using them to try and predict the likeliness of recipe. Here we are building a model by applying various machine learning algorithms find the best accurate model. And integrate to flask based web application. User can predict the disease by entering parameters in the web application.

DISADVANTAGES OF EXISTING SYSTEM:

1. Collecting large amount of data set.

2. Storing the large amount of image data sets.

3.4. PROPOSED SYSTEM:

Here we are building a model by applying various machine learning algorithms find the best accurate model. Then it will predict the exact food recipe to remove confusion in the data set.

Some of the machines learning algorithms are:

1.Linear Regression:

Linear Regression is **a supervised machine learning algorithm where the predicted output is continuous and has a constant slope**. It's used to predict values within a continuous range, (e.g. sales, price) rather than trying to classify them into categories (e.g. cat, dog).

Steps to implement Linear regression model:

    a. Initialize the parameters.

    b. Predict the value of a dependent variable by given an independent variable.

    c. Calculate the error in prediction for all data points.

    d. Calculate partial derivative w.r.t a0 and a1.

    e. Calculate the cost for each number and add them.

2. **Multiple Linear Regression:**

Multiple Linear Regression is one of the important regression algorithms which models the linear relationship between a single dependent continuous variable and more than one independent variable. Multiple regression is a broader class of regressions that encompasses linear and nonlinear regressions with multiple explanatory variables.

3.**Random Forest:**

A random forest is a machine learning technique that's used to solve regression and classification problems. It utilizes ensemble learning, which is a technique that combines many classifiers to provide solutions to complex problems.

Working of Random Forest Algorithm

Step 1 − First, start with the selection of random samples from a given dataset.

Step 2 − Next, this algorithm will construct a decision tree for every sample. Then it will get the prediction result...

Step 3 − In this step, voting will be performed for every predicted result.

Step 4 − At last, select the most voted prediction result as the final prediction result.

4.**Logistic regression:**

1. Logistic regression is **a supervised learning classification algorithm used to predict the probability of a target variable**. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.... Mathematically, a logistic regression model predicts P(Y=1) as a function of X.
2. Logistic Regression is used when the dependent variable (target) is categorical. For example,
3. To predict whether an email is a spam (1) or (0)
4. Whether the tumor is malignant (1) or not (0)

You will need to train the datasets to run smoothly and see an incremental improvement in the prediction rate.

5. <u>k-nearest neighbor algorithm:</u>

1. It is a supervised machine learning algorithm. The algorithm can be used to solve both classification and regression problem statements. The number of nearest neighbor's to a new unknown variable that has to be predicted or classified is denoted by the symbol 'K'.
2. KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.
3. K-Nearest Neighbors (KNN) is one of the simplest algorithms used in **Machine Learning for regression and classification problem**. KNN algorithms use data and classify new data points based on similarity measures (e.g. distance function). The data is assigned to the class which has the nearest neighbors.
4. It's also worth noting that the KNN algorithm is also part of a family of ─lazy learning‖ models, meaning that it only stores a training dataset versus undergoing a training stage. This also means that all the computation occurs when a classification or prediction is being made. Since it heavily relies on memory to store all its training data, it is also

referred to as an instance-based or memory-based learning method.

5. The K-NN working can be explained on the basis of the below algorithm:

    Step-1: Select the number K of the neighbors

    Step-2: Calculate the Euclidean distance of **K number of neighbors**

    Step-3: Take the K nearest neighbors as per the calculated Euclidean distance.

    Step-4: Among these k neighbors, count the number of the data points in each category.

    Step-5: Assign the new data points to that category for which the number of the neighbor is maximum.

    Step-6: Our model is ready.

4. EXPERIMENTAL ANALYSIS

Milestone 1: Data Collection

ML depends heavily on data, without data, a machine can't learn. It is the most crucial aspect that makes algorithm training possible. In Machine Learning projects, we need a training data set. It is the actual data set used to train the model for performing various actions.

You can collect datasets from different open sources like kaggle.com, data.gov; UCI machine learning repository etc. The dataset used for this project was obtained from Kaggle.

Milestone 2: Data Pre-processing

Data Pre-processing includes following

main tasks   Importing the libraries.

1. Importing the dataset.
2. Analyse the data.
3. Taking care of Missing Data.
4. Data Visualisation.
5. Splitting Data into Train and Test.

Milestone 3: Model Building

The model building process involves setting up ways of collecting data, understanding and paying attention to what is important in the data to answer the questions you are asking, finding a statistical, mathematical or a simulation model to gain understanding and make predictions. Model Building Includes:

6. Import the model building libraries.
7. Initialising the model.
8. Training the model.
9. Model Evaluation.
10. Save the Model.

Milestone 4: Application Building

11. Create an HTML File.
12. Build python code.

13. Run the app in local browser.

14. Show casting the prediction on UI.

4.1. PROJECT ARCHITECTURE:

4.2. SOFTWARE AND HARDWARE REQUIREMENTS:

Software Requirements:

1. Anaconda Environment

2. Flask

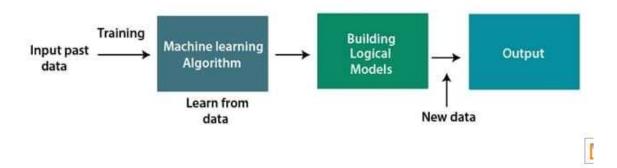3. Python 3.9

4. And other python libraries like NumPy, pandas.

Hardware Requirements:

5. Operating system

6. Processing

7. RAM

8. Operating system specifications

9. Disk space
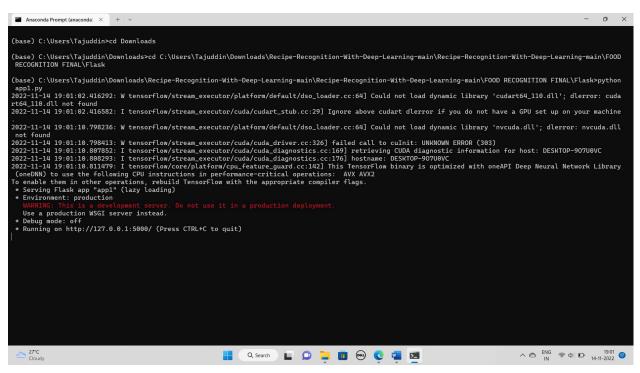
4.3. BLOCK DIAGRAM:
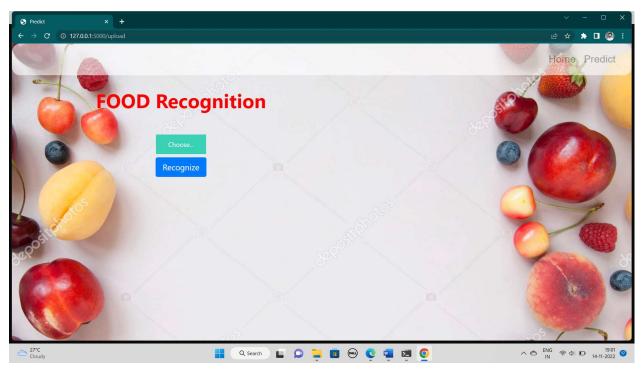
Figure 4 : Block Diagram

4.4. PROJECTFLOW:

1. User interacts with the UI (User Interface) to upload the input features.

2. Uploaded features/input is analysed by the model which is integrated.
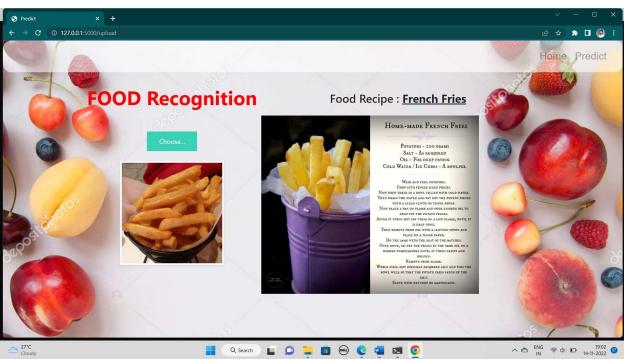
3. Once a model analyses the uploaded inputs, the prediction is showcased on the UI.

6.RESULT

7. ADVANTAGES AND DISADVANTAGES

- It uses mainly in the restaurent to know the recipe type and ingredients in the food.
- It help full for the customer to know the food variety in list of foods.
- Type of food choosing option.
- Time consumtion is less.

DISADVANTAGES

- Mention food items highly calorie contain.
- Food may Effect on the body.
- physical and mental issues.

8.APPLICATIONS

- These can be applied on the hotel management in the big restaurent.
- It can be easier to who cook the food.
- Easy to understand and quick to learn.

9.CONCLUSION

In UG Project Phase-1, we have worked on problem statement, literature survey and also done the experimental analyses which are required for the project to move forward. In experimental analysis we have discussed about the machine learning concepts and models and explained the algorithms to be used in the project. We also discussed about the flowcharts, use case diagrams, decision tree and sequence diagrams which are used in the project. Based on the experimental analysis we have designed the model for the project.

Entire designing part is involved in UG Project Phase-1.

10.FUTURE SCOPE

UG Project Phase-1 about recipe recognisation and All the implementation is done and conclusions will be retrieved in this phase. We will work on the applications, advantages, and disadvantages of the project in this phase.