DAYANANDA SAGAR UNIVERSITY

KUDLU GATE, BANGALORE – 560068

Bachelor of Technology
in
COMPUTER SCIENCE AND ENGINEERING

Major Project Phase-II Report

DIET RECOMMENDATION SYSTEM

By

L K Sabarish-ENG18CS0145
Manoj Kumar S-ENG18CS0159
Maria Jyothi S-ENG18CS0162
Juhi Bhowal-ENG18CS0167

Under the supervision of

Dr. Meenakshi Malhotra
Associate Professor, CSE

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING,
SCHOOL OF ENGINEERING
DAYANANDA SAGAR UNIVERSITY,
BANGALORE

(2021-2022)

DAYANANDA SAGAR UNIVERSITY

School of Engineering

Department of Computer Science & Engineering

Kudlu Gate, Bangalore – 560068
Karnataka, India

CERTIFICATE

This is to certify that the Phase-II project work titled **"DIET RECOMMENDATION SYSTEM"** is carried out by **L K Sabarish (ENG18CS0145), Manoj Kumar S (ENG18CS0159), Maria Jyothi S (ENG18CS0162), Juhi Bhowal (ENG18CS0167),** bonafide students of Bachelor of Technology in Computer Science and Engineering at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfillment for the award of degree in Bachelor of Technology in Computer Science and Engineering, during the year **2021-2022**.

Dr. Meenakshi Malhotra	**Dr. Girisha G S**	**Dr. A Srinivas**
Associate Professor Dept. of CSE, School of Engineering Dayananda Sagar University	Chairman CSE School of Engineering Dayananda Sagar University	Dean School of Engineering Dayananda Sagar University
Date:	Date:	Date:

Name of the Examiner	**Signature of Examiner**
1.	
2.	

DECLARATION

We, **L K Sabarish (ENG18CS0145), Manoj Kumar S (ENG18CS0159), Maria Jyothi S (ENG18CS0162), Juhi Bhowal (ENG18CS0167)** are students of the eighth semester B.Tech in **Computer Science and Engineering**, at School of Engineering, **Dayananda Sagar University**, hereby declare that the phase-II project titled **"Diet Recommendation System"** has been carried out by us and submitted in partial fulfillment for the award of degree in **Bachelor of Technology in Computer Science and Engineering** during the academic year **2021-2022**.

Student	**Signature**
L K Sabarish	
(ENG18CS0145)	
Manoj Kumar S	
(ENG18CS0159)	
Maria Jyothi S	
(ENG18CS0162)	
Juhi Bhowal	
(ENG18CS0167)	

Place : Bangalore

Date :

ACKNOWLEDGEMENT

It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this project work.

First, we take this opportunity to express our sincere gratitude to School of Engineering, Dayananda Sagar University for providing us with a great opportunity to pursue our Bachelor's degree in this institution.

We would like to thank **Dr. A Srinivas. Dean**, **School of Engineering & Technology**, **Dayananda Sagar University** for his constant encouragement and expert advice. It is a matter of immense pleasure to express our sincere thanks to **Dr. Girisha G S, Department Chairman**, **Computer Science, and Engineering**, **Dayananda Sagar University,** for providing the right academic guidance that made our task possible.

We would like to thank our guide **Dr. Meenakshi Malhotra, Associate Professor**, **Dept. of Computer Science and Engineering**, **Dayananda Sagar University**, for sparing his/her valuable time to extend help in every step of our project work, which paved the way for smooth progress and the fruitful culmination of the project.

We would like to thank our Project Coordinator **Dr. Meenakshi Malhotra, Dr.Bharanidharan** and all the staff membersof Computer Science and Engineering for their support.

We are also grateful to our family and friends who provided us with every requirement throughout the course. We would like to thank one and all who directly or indirectly helped us in the Project work.

ABSTRACT

In this modern world various people suffer from different types of diseases and illnesses. It is generally very difficult to suggest a diet as there are various number of ingredients available pertaining to a particular region in the country. The project makes use a dataset which contains recipes of various dishes with what nutrients is present in the dish, category (veg/non-veg), disease and what kind of diet. In the wake of the situation, a program is developed that recommends diet to the people based on the input collected from the user. This project uses machine learning algorithms named decision tree, random forest classification, gaussian naïve bayes, adaptive boosting, logistic regression and support vector machine to test the efficiency of the algorithm. Random Forest algorithm gives the maximum of efficiency. To predict the food items the DRS uses user inputs from a GUI including age, height, weight, vegetarian or non-vegetarian food, disease. The working prototype of the DRS lists a set of food items and recipe as per the user inputs.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

- **DRS-** DIET RECOMMENDATION SYSTEM
- **SVM**- SUPPORT VECTOR MACHINE
- **GUI**- GRAPHICAL USER INTERFACE
- **LR**- LOGISTIC REGRESSION
- **MLP**- MULTI-LAYER PERCEPTRON CLASSIFIER
- **RF**- RANDOM FOREST
- **AB**- ADA BOOST CLASSIFIER
- **GNB**- GAUSSIAN NAIVE BAYES ALGORITHM
- **DT**- DECISION TREE CLASSIFIER

CHAPTER 1 INTRODUCTION

Recommendation system for patients/dieticians is a system that monitors a user in a tailored approach towards remarkable or suitable diets or food intake in large varieties of likely selections and that results in such selections as desired output. A recommendation system for patients/dieticians is cautiously implemented for the purpose of recommending the patients to take nutritional diets and food which are considered better to meet the patients' health needs, taste and dietary preferences. Lately, in terms of life saving healthy living, recommendation systems are one of the probable solutions that will facilitate patients' choice of food intake considering the enormous amount of accessible data interrelated to foods/recipes.

1.1 RECOMMENDER SYSTEM:

In the Diet Recommendation System, there are many factors like age, height, weight, disease, residence etc, that will use the machine learning functionalities like decision tree, random forest Classification, Gaussian naïve boost, adaptive boosting, logistic regression and support vector machine to test accuracy of the algorithms to recommend the list of food to the user. The user defined functions make use of various in-built Python's functions used in machine learning to provide the actual functionality. We are using various libraries like pandas, numpy and sklearn in general. The pandas library is used to import the data set so that the machine learning model can make useof it and the list of food items can be predicted to the user. numpy is used for values conversion to array so that it can provide the values which can be used to train the Machine Learning model. The sklearn package has implementation of many of the machine learning algorithms.

The implementation of the Diet Recommendation System begins with the user input as accepted from the Graphical User Interface (GUI). The GUI(Website) is built using Java.

1.1.1. DATASET

Data is collected from the Kaggle an internet site for creating and testing purpose. There are two datasets collected for the purpose of our project. One containing the Food details and the other containing the disease and health conditions.

1.2 SCOPE

The objective of this recommender systems is to provide recommendations based on recorded information on the users' preferences. This system uses information classifying techniques to process information and provide the user with potentially more relevant items. The input is taken via a GUI built using java. Based on the input provided by the user , the system recommends food ingredients along with recipe that a user can consume.

This project provides food for a particular input and if the user doesn’t feel better even after trying these food recommendation, he/she needs to consult a doctor since health can differ based on many other factors not included in the project.

CHAPTER 2 PROBLEM DEFINITION

Making decisions about what to eat is a major problem in our everyday lives due to a wide variety of ingredients, culinary styles, ethnicities, cultures, and personal tastes. Choosing the right dish at the right time seems to be a very difficult task. Today, many diseases that were previously thought as hereditary are now shown to be connected to biological dysfunction related to nutrition. Although being healthy and eating better issomething the vast majority of the population want, doing so usually requires great effort and organization.

This project work provides a prediction system that recommends the right food and with the right nutrition to a particular person based on their health and food preferences.

CHAPTER 3 LITERATURE REVIEW

According to the paper A DASH Diet Recommendation System for Hypertensive Patients Using Machine Learning (2019) [1], Abstract— Hypertension is becoming a serious health issue in the world. Due to poor eating habits, the rate of Non-Communicable Diseases (NCDs) such as hypertension together with the rate of death caused by such diseases are rising. In order to promote healthy eating habits in Mauritius, the paper proposes a DASH diet recommender system that recommends healthy Mauritian diet plans to hypertensive patients. The system consists of a recommendation engine that uses techniques such as content-based filtering along with machine learning algorithms to recommend personalized diet plans to hypertensive patients based on factors such as age, user preferences about food, blood pressure level. The system makes use of a mobile application which is handy and quick to use. Based on a survey carried out, the application has helped users to control and reduce their BP level.

According to the paper Improved Predictive Learning Approaches For Customized Diet Suggestion Framework In Medical Services (2020) [2], - The project makes use of a dataset which contains various nutrients in the correct amount. A program has been developed to recommend diet to the people. The items recommended are limited to three categories: Weight Loss, Weight Gain and Healthy category. The project uses Machine Learning Algorithms named K-Means Clustering for clustering the data and Random Forest Classification to classify according to the categories listed. To predict the food items the Diet Recommendation System uses user inputs from a Graphical User Interface including age, height, weight, vegetarian or non-vegetarian food and selecting the above three categories. The working prototype of the Diet Recommendation System lists a set of food items as per the user inputs. The module uses the weight and height to calculate the body Mass Index (BMI) of the user and based on the preference of the kind of diet he wants the Recommendation System predicts the list of food items. The various kinds of nutrients and their calories are taken into consideration along with other relevant details like Fats, Proteins, Iron, Calcium, Sodium, Potassium, Carbohydrates, Fiber, Vitamin D and Sugar. A diverse range of food items are considered like major Vegetarian and Non- Vegetarian foods. Carbohydras, Fats and Nutrients are the major contributors that are considered to make a food.

According to the paper Recommendation of Diet to a Patient using AHP and Fuzzy Approach (2019) [3], Marasmus is a malnutrition deficiency which can occur to anyone with severe malnutrition deficiency

but is common in children. It occurs when there is improper ingestion of carbohydrates, calories, proteins and other vital nutrients. A Marasmic patient is one who has been depleted of his stored glycogen and the present fat supplements, and the tissues generally get starved while the digestive secretions become impaired. So, for the successful recovery a diet should be followed which should comprise of all the vital nutrients in the correct proportion. The purpose of this diet recommendation system is firstly to cater the needs of the people who are suffering from Marasmus and then to recommend the best diet plan which should be followed for a finest results Analytic Hierarchy Process (AHP) is employed which helps in determining the balanced diet which has all the vital nutrients in the correct proportions and to validate result using the Fuzzy Logic.

According to the paper Micronutrient Deficiencies in the Developing World:An Evaluation of Delivery Methods (2016) [4], Malnutrition is a serious problem in developing countries, particularly among children under five and women of childbearing age. Expectant and nursing mothers are especially vulnerable to micronutrient deficiencies. Effects of micronutrient deficiencies such as weakened immune systems, overall morbidity, and stunted childhood growth are reversible in the critical age group of 0 to 5. Many solutions have been researched and implemented to increase micronutrient intake in the diets of people in developing nations. The article presents a comprehensive review of academic studies detailing different methods of delivering nutrients to individuals in developing countries. The results of this review are categorized and synthesized into a framework describing the approach, range and success of different strategies to help vulnerable groups meet their micronutrient needs. Understanding this framework will help in identifying gaps in current efforts to address micronutrient deficiencies and improve the current integration techniques into settings with insufficient resources to satisfy dietary needs of populations.

According to the paper Nourishment Recommendation Framework for Children Using Machine Learning and Matching Algorithm (2019) [5], Food plays an important role in daily life and is related to our core existence. The food that we eat supplies nutriments to the body. A balanced diet full of essential vitamins
and minerals in adequate amount will help an individual reach their full growth potential especially children because a proper diet results in an increase in the learning ability, stamina and strength. It even has a positive impact on their behaviour. To offer an efficient method regarding food intake, in this

paper, we proposed a Nourishment Recommendation Framework (NRF) where we take user input from children, analyse the data and finally an output is generated that presents an improved diet plan. It aims to provide children from 8 – 13 years with healthy meals according to their age, growth, gender and health records.

According to the paper A fuzzy expert system for diagnosis of malnutrition in Children (2017) [6], The aim of this study is to design a fuzzy expert system which diagnoses whether an infant is suffering from malnutrition or not, and if he/she is in the trap of malnutrition, then what is the extent or severity of the same. For this purpose, we have taken into account 13 input variables- breast feeding, daily family income, level of mother's education, colostrum intake, energy intake, protein intake, vitamin A intake, iron intake, family size, height of infant, weight of infant, head circumference of infant and skinfold thickness of infant. The presence of malnutrition and its severity are diagnosed, keeping the following categories in view- severe under nutrition, moderate under nutrition, mild under nutrition, normally nourished, mild over nutrition and moderate over nutrition. At this very point, it is worth stressing the fact that malnutrition is not just an indicator of under nutrition. Over nutrition is as well a form of malnutrition. The designed expert system uses Mamdani inference method. With this system, one can very easily diagnose the level of nourishment in an infant by just giving appropriate values of the inputs.

CHAPTER 4 PROJECT DESCRIPTION

The goal of this project is to build a recommendation system for ingredients and recipes from the dataset based on the health conditions and factors. The factors include height, weight, age, food preference like Vegetarian or non-vegetarian, flavor (dessert, main course etc.). This project work intended to train, evaluate and test a model for given dataset to predict diet from a set of given ingredients.

The output will include a dish with all the ingredients and the recipe. Finally, this project is implemented as a Web Application created using Java as a step forward in building a recommendation system.

CHAPTER 5 REQUIREMENTS

5.1 Hardware Requirements

1.Processor : Intel i5 3.0 GHz

2.Operating system : Windows 7 and above

3.Programming Language : Python

4.Library : Matplotlib, numpy, Pandas, Sckit-learn.

5.Simulation tool : Anaconda Navigator IDE 3.7.4, (jupyter Notebook).

5.2 Software Requirements

Jupyter Notebook

Python 3

Anaconda 3

Netbeans IDE and JDK.

5.3 Functional Requirements

- The proposed project is a web-based recommender system being built using machine learning algorithms.
- The user provides inputs using the graphical Interface such as type of diet, food preferences (Veg/Non-Veg) any particular disease the user is suffering from.
- Once the input is given a profile for the user is created. Based on the input given the system provides an appropriate ingredients list and recipes from the ingredients.

The GUI is built using Java. To train the model different algorithms are used namely SVM (Support VectorMachines), ADABoost, GaussianNB, Logistic Regression, Decision Tree and Random Forest. Algorithm with highest accuracy is then selected for prediction and the user is provided with Recipe.

5.4 Non-Functional Requirements

5.4.1 Performance Requirements

- To improve the performance of the software it needs to be executed with an internet speed of 45mbps so that no issues are observed in the existing working system.
- Various other measures need to be taken care of like the storage to install JDK and Netbeans. As a result, there are no performance issues observed.

5.3.2 Security Requirements

The software is hosted on a local server and the project does not require any login credentials as a result user data is not stored to cause security issues.

CHAPTER 6 METHODOLOGY

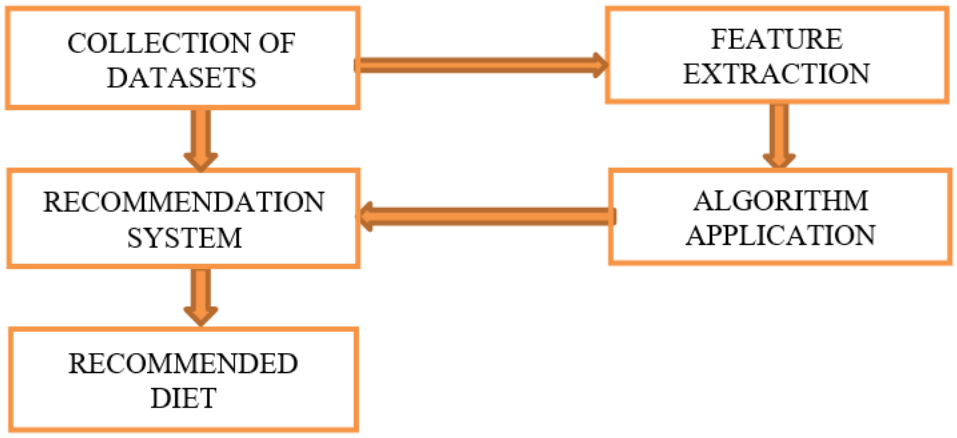


Figure 1: Process Flow Diagram

The Framework includes a procedure that is separated into various stages as depicted in figure 1. The phases are as per the following:

1) Collection of Datasets

2) Pre-processing (Noise Removal)

3) Feature Extraction

4) Applied Various Machine Learning Algorithm

5) Recommendation System

6) Recommended Diet

1) Data Collection:

There are two datasets consists of parameters like dish name, ingredient, veg/non veg, prep time, cook time, diseases etc. The datasets have been obtained from the Kaggle website. The first dataset for recipe has about 6871 instances. This dataset includes 255 different dishes. The second dataset for disease has parameters such as disease, which kind of diet etc.

2) Pre-Processing (Noise Removal):

For the successful application pre-processing is required. The data which is acquired from different resources is sometime in raw form. It may contain some incomplete, redundant inconsistent data. Therefore, in this step such redundant data is filtered. Data should be normalized. The prep time and cook time has -1 values which cannot be since time cannot be measured in minus values. Preprocessing of data involves replacing these values with mean value of prep time and cook time of similar dishes respectively as shown in the code snippet in figure 3.

	diet	prep_time	cook_time	flavor_profile	course	state	region
0	vegetarian	45	25	sweet	dessert	West Bengal	East
1	vegetarian	80	30	sweet	dessert	Rajasthan	West
2	vegetarian	15	60	sweet	dessert	Punjab	North
3	vegetarian	15	30	sweet	dessert	Rajasthan	West
4	vegetarian	15	40	sweet	dessert	West Bengal	East
...	...	...	...	...	...	...	...
250	vegetarian	5	30	sweet	dessert	Assam	North East
251	vegetarian	20	60	sweet	dessert	Goa	West
252	vegetarian	-1	-1	sweet	dessert	Jammu & Kashmir	North
253	vegetarian	20	45	sweet	dessert	Madhya Pradesh	Central
254	vegetarian	-1	-1	sweet	dessert	Goa	West

Figure 2: Prep time and cook time with -1

```
data[['prep_time']] = data[['prep_time']].fillna(data['prep_time'].mean())
data[['cook_time']] = data[['cook_time']].fillna(data['cook_time'].mean())
```

Figure 3: Replacing -1 with mean value

3) Feature Extraction: This step is focused on identifying and using most relevant attributefrom the dataset. Through this process irrelevant and redundant information is removed for the application of classifiers.

4) Application of Different Algorithms:

The proposed system applied different machine learning algorithms like Random Forest, AdaBoost Classifier, GaussianNB, MLP Classifier, Decision Tree Classifier, Support Vector Machine (SVM) and Logistic Regression (LR).

a) Decision Tree Classifier:

Decision tree classifiers utilize greedy methodology. It is a supervised learning algorithm where attributes and class labels are represented using a tree. The mainpurpose of using Decision Tree is to form a training prototype which we can use to foresee class or value of target variables by learning decision rules deduced from previous data (training data). The decision tree can be described by two distinct types, namely decision nodes and leaves. The leaves are the results or thefinal end results. Each node in the tree acts as a test case for some attribute, and each edge descending from that node corresponds to one of the possible answers to the test case. This process is recursive in nature and is repeated for every sub- tree rooted at the new nodes.

b) Support Vector Machine (SVM) :

Support Vector Machine (SVM) is a supervised machine learning algorithm or model which can be utilized for classification and as well as for regression challenges. However, we mainly use it in classification challenges. SVM is generally represented as training data points in space which is divided into groupsby intelligible gap which is as far as possible. In SVM algorithm, each data item is plotted as a point in n-dimensional space with each feature value being the valueof a specific coordinate. Then the classification is performed by finding the hyper-plane differentiating the two classes very well. We have applied Support Vector

c) Logistic Regression (LR) :

The Logistic Regression model is a broadly used statistical model that, in its basicform, uses a logistic function to model a binary dependent variable; many more complex extensions exist. In Regression Examination, Logistic regression is predicting the parameters of a logistic model; it is a form of Binomial regression

d) Random Forest (RF):

Random Forest is a ML algorithm. At training situation multitude decision trees are made and the output will be divided based on number of classes i.e., classification, prediction of class i.e., regression. The number of trees is proportional to accuracy in prediction. The dataset includes factors like rainfall, perception, temperature and production. These factors in dataset are used for training. Only two-third of the dataset is considered. Remaining dataset is used forexperimental basis. The algorithm random forest has 3 parameters like: n tree which describes the n number of trees which need to grow, m try - mentions howmany variables need to be taken at a node split. Node size - In terminal nodes it suggests us the number of observations need to take.

e) MLPClassifier:

MLPClassifier stands for Multi-layer Perceptron classifier which in the name itself connects to a Neural Network. Unlike other classification algorithms such as Support Vectors or Naive Bayes Classifier, MLPClassifier relies on an underlying Neural Network to perform the task of classification.

f) ADABoost:

AdaBoost also called Adaptive Boosting is a technique in Machine Learning usedas an Ensemble Method. The most common algorithm used with AdaBoost is decision trees with one level that means with Decision trees with only 1 split.

These trees are also called Decision Stumps

g) Gaussian NB

Gaussian Naive Bayes supports continuous valued features and models each as conforming to a Gaussian (normal) distribution. An approach to create a simple model is to assume that the data is described by a Gaussian distribution with no co-variance (independent dimensions) between dimensions.

These Algorithms are used in our project in the following ways:

i) Importing library

In this section all the necessary libraries are imported.

```
from sklearn.naive_bayes import GaussianNB
from sklearn.linear_model import LogisticRegression
from sklearn.neural_network import MLPClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier
```

Figure 3: Importing of libraries

ii) Next creating classifier object

For each algorithm to be executed a classifier object is created.

```
for algo in algorithms:
    clf = algorithms[algo]
```

Figure 4: Classifier Object

iii) Finally, fit the data

The training data is taken and fitted into the training model to retrieve the results.

```
algorithms = {
        "RF": ske.RandomForestClassifier(n_estimators=50),
        "AB": ske.AdaBoostClassifier(n_estimators=100),
        "GNB": GaussianNB(),
        "MLP": MLPClassifier(solver='lbfgs', alpha=1e-5,hidden_layer_sizes=(5, 2), random_state=1),
        "DT": DecisionTreeClassifier(),
        #"SVM": SVC(kernel='rbf'),
        "SVM":svm.LinearSVC(),
        "LR": LogisticRegression()
    }

results = {}
print("Algorithm Test:")
for algo in algorithms:
    clf = algorithms[algo]
    clf.fit(X_train, y_train)
    score = clf.score(X_test, y_test)
    print("%s : %f %%" % (algo, score*100))
    results[algo] = score
```

Figure 5: Accuracy Comparison

CHAPTER 7 EXPERIMENTATION

- Data was initially collected from family and friends using google forms, but enough data wasn't collected. Therefore, Data is collected from the Kaggle an internet site for creating and testing purpose.The dataset contains meal and disease and health conditions.
- The dataset had lot of missing values and inappropriate data like -1 for cook time and prep time which cannot be measured in negative value.

  These negative values were replaced with the mean of the entire column.
- Then the dataset was split into 70% training data and testing data to 30%.

  The training data was fit into different algorithms like random forest, adaptive boosting, gaussian naïve bayes, decision tree and support vector machine to calculate the accuracy of the algorithms based on the data given.
- Random forest achieved the maximum efficiency of 87.01% of accuracy.

```
for algo in algorithms:
    clf = algorithms[algo]
    clf.fit(X_train, y_train)
    score = clf.score(X_test, y_test)
    print("%s : %f %%" % (algo, score*100))
    results[algo] = score
```

```
bestAlgo = max(results, key=results.get)
print('\nBest algorithm is %s with a %f %% success' % (bestAlgo, results[bestAlgo]*100))
clf = algorithms[bestAlgo]
res = clf.predict(X_test)
```

Figure 6: Accuracy Comparison

The code in the above figure shows a for loop that picks each algorithm and the training data is fit into the algorithm and the score for each algorithm is returned.

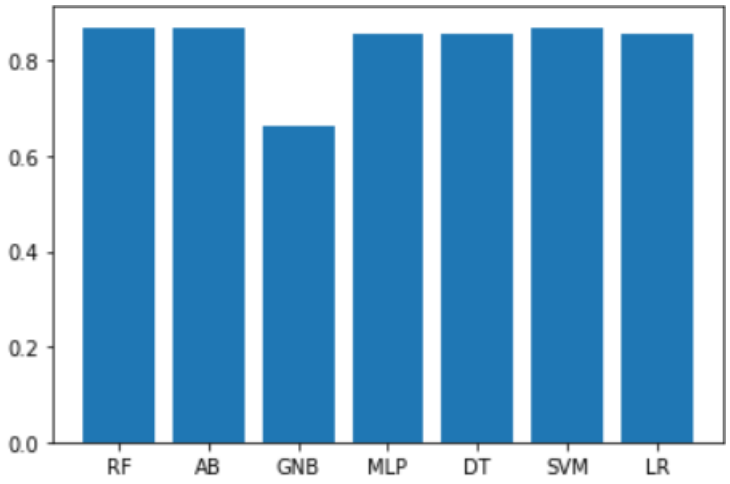


Figure 6: Accuracy Comparison

```
Testing  key  feature:
Algorithm Test:
RF : 87.012987 %
AB : 87.012987 %
GNB : 66.233766 %
MLP : 85.714286 %
DT : 85.714286 %
SVM : 87.012987 %
LR : 85.714286 %

Best algorithm is RF with a 87.012987 % success
```

Figure 7: Algorithm vice Accuracy Result

In order to understand how different recipes are predicted by the classifier a multi-dimensional confusion matrix is constructed.

```
def Patient_Diet_Food_Recommendation_System(y_true, y_pred, labels, ymap=None, figsize=(10,10)):
    if ymap != None:
        y_pred = [ymap[yi] for yi in y_pred]
        y_true = [ymap[yi] for yi in y_true]
        labels = [ymap[yi] for yi in labels]
    cm = confusion_matrix(y_true, y_pred, labels=labels)
    cm_sum = np.sum(cm, axis=1, keepdims=True)
    cm_perc = cm / cm_sum * 100
    annot = np.empty_like(cm).astype(str)
    nrows, ncols = cm.shape
    for i in range(nrows):
        for j in range(ncols):
            c = cm[i, j]
            p = cm_perc[i, j]
            if i == j:
                s = cm_sum[i]
                annot[i, j] = '%.1f%%\n%d/%d' % (p, c, s)
            elif c == 0:
                annot[i, j] = ''
            else:
                annot[i, j] = '%.1f%%\n%d' % (p, c)
    cm = pd.DataFrame(cm, index=labels, columns=labels)
    cm.index.name = 'Actual'
    cm.columns.name = 'Predicted'
    fig, ax = plt.subplots(figsize=figsize)
    sns.heatmap(cm, annot=annot, fmt='', ax=ax)
```

Figure 8: Confusion Matrix

CHAPTER 8 TESTING AND RESULTS

The result obtained says Random Forest is the best Algorithm with 87.01% success.

```
clf = algorithms[bestAlgo]
res = clf.predict(X_test)
mt = confusion_matrix(y_test, res)
Patient_Diet_Food_Recommendation_System(y_test, res,clf.classes_)
FS=f1_score(y_test, res, average='macro')
plt.show()
```

Figure 7: Model Output

In order to understand how different recipes are predicted by the classifier a multi-dimensional confusion matrix is constructed. The results obtained are shown in the figure below:

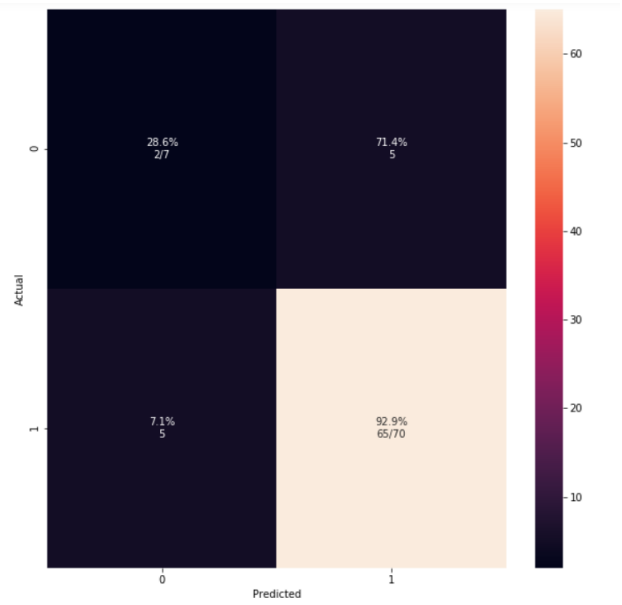


Figure 8: confusion Matrix.

The GUI built using Java.
The initial Home page of the GUI.
When clicked on enter, the system takes us to the recommender system where inputs can be given.

The input includes the type of diet the user is interested in, the disease he is suffering from and the preference veg or non-veg.

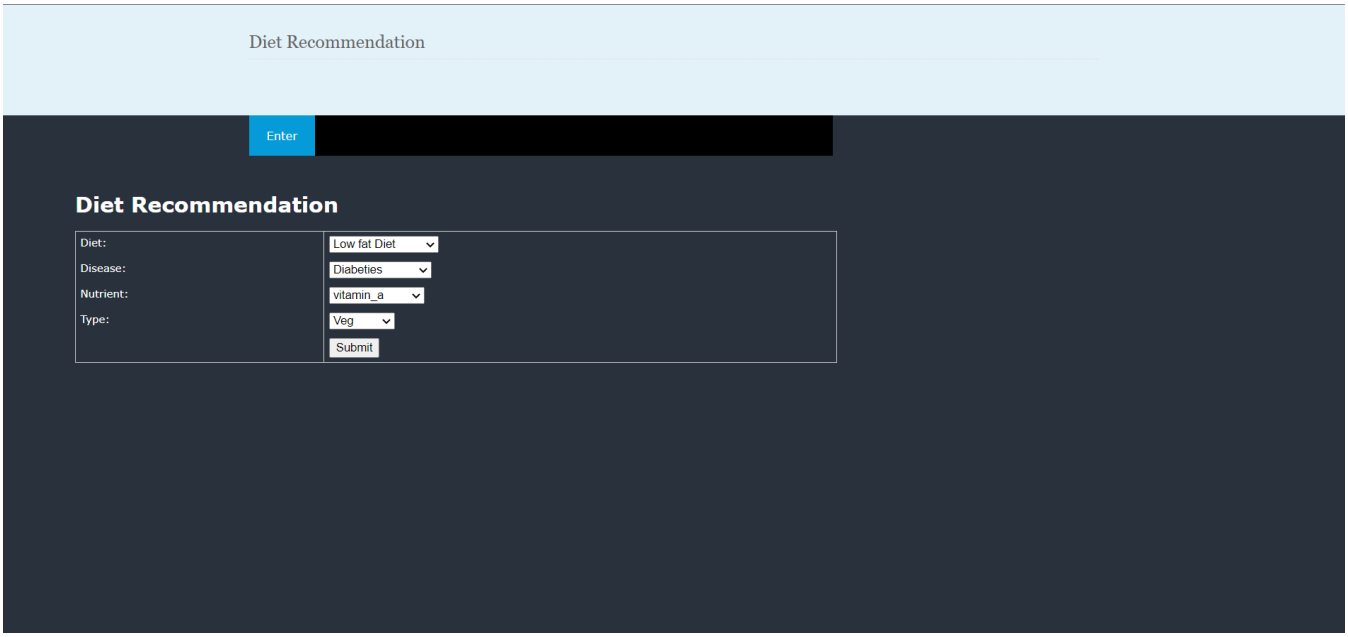
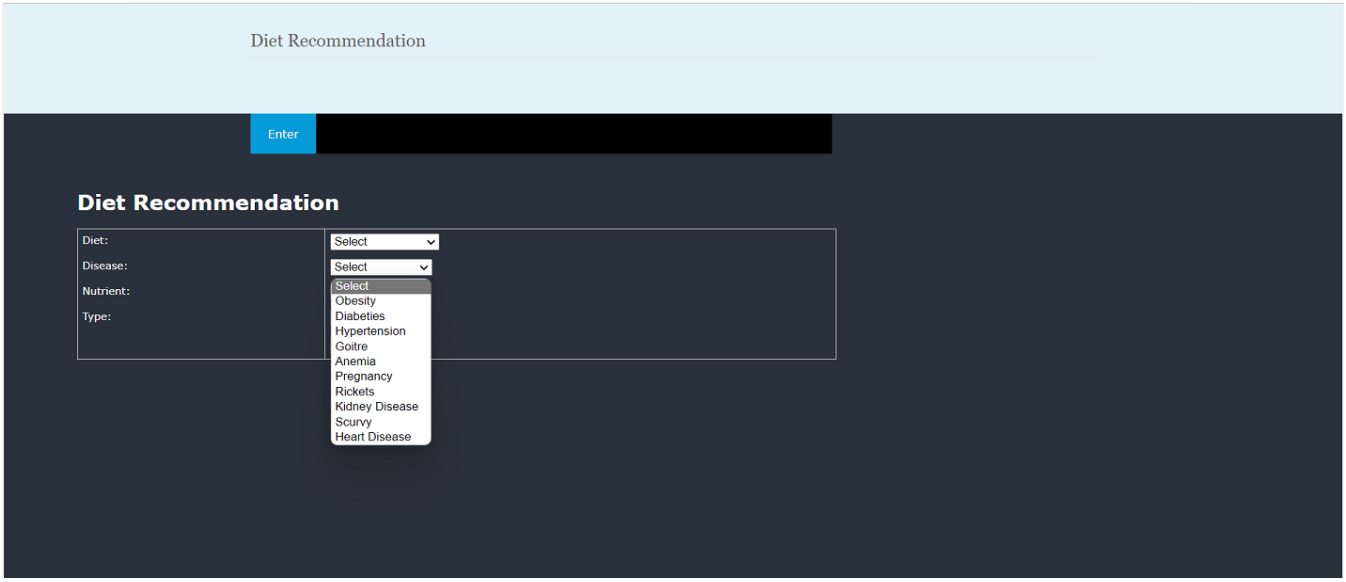


Figure 9:Input Screen

There are multiple diseases the user can select from drop down including heart disease, eye disease , scurvy, rickets etc.



The Type includes the user preference veg or non-veg.

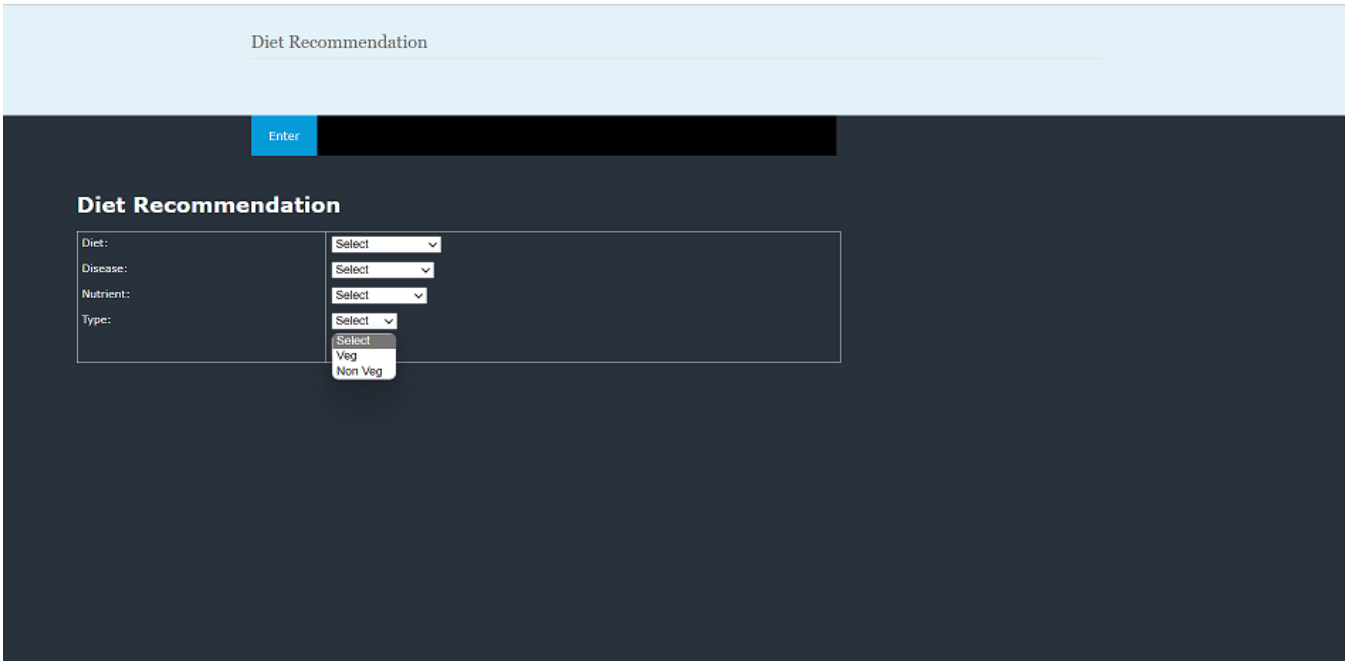
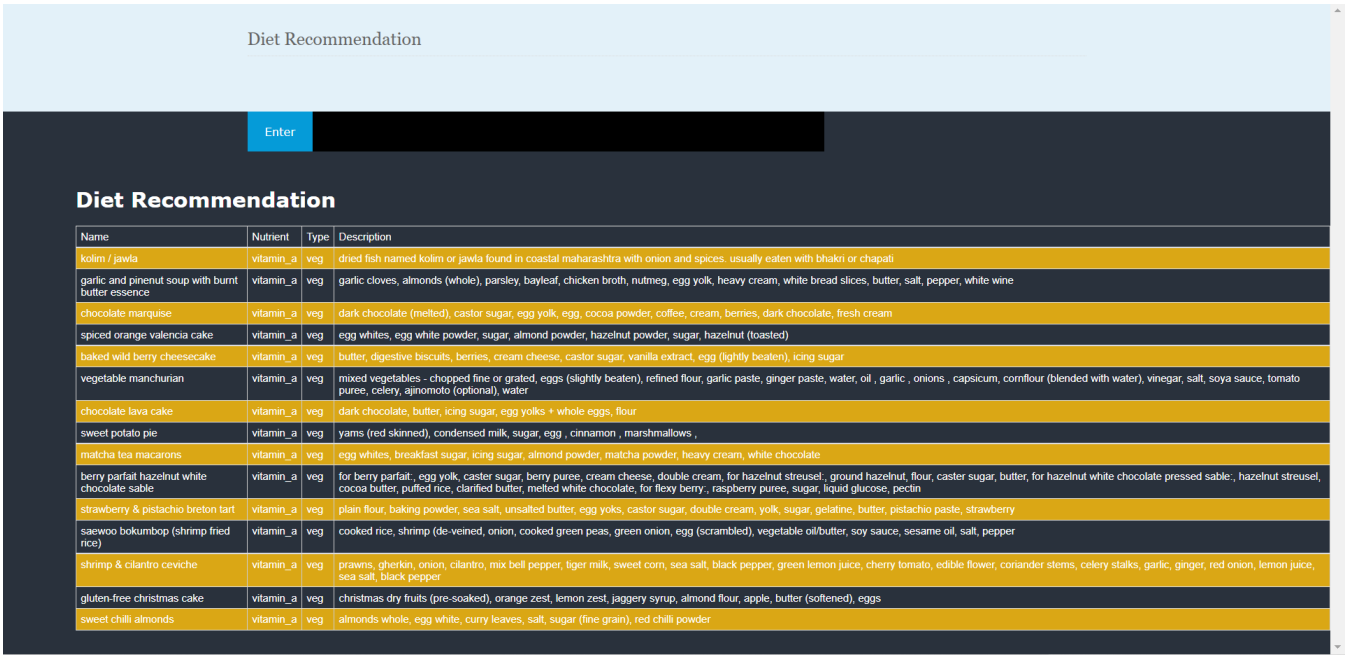


Figure 10: Output 1

Once the input is given, the output is displayed with all the nutrients and recipe.



Name	Nutrient	Type	Description
kolim / jawla	vitamin_a	veg	dried fish named kolim or jawla found in coastal maharashtra with onion and spices. usually eaten with bhakri or chapati
garlic and pinenut soup with burnt butter essence	vitamin_a	veg	garlic cloves, almonds (whole), parsley, bayleaf, chicken broth, nutmeg, egg yolk, heavy cream, white bread slices, butter, salt, pepper, white wine
chocolate marquise	vitamin_a	veg	dark chocolate (melted), castor sugar, egg yolk, egg, cocoa powder, coffee, cream, berries, dark chocolate, fresh cream
spiced orange valencia cake	vitamin_a	veg	egg whites, egg white powder, sugar, almond powder, hazelnut powder, sugar, hazelnut (toasted)
baked wild berry cheesecake	vitamin_a	veg	butter, digestive biscuits, berries, cream cheese, castor sugar, vanilla extract, egg (lightly beaten), icing sugar
vegetable manchurian	vitamin_a	veg	mixed vegetables - chopped fine or grated, eggs (slightly beaten), refined flour, garlic paste, ginger paste, water, oil , garlic , onions , capsicum, cornflour (blended with water), vinegar, salt, soya sauce, tomato puree, celery, ajinomoto (optional), water
chocolate lava cake	vitamin_a	veg	dark chocolate, butter, icing sugar, egg yolks + whole eggs, flour
sweet potato pie	vitamin_a	veg	yams (red skinned), condensed milk, sugar, egg , cinnamon , marshmallows ,
matcha tea macarons	vitamin_a	veg	egg whites, breakfast sugar, icing sugar, almond powder, matcha powder, heavy cream, white chocolate
berry parfait hazelnut white chocolate sable	vitamin_a	veg	for berry parfait:, egg yolk, caster sugar, berry puree, cream cheese, double cream, for hazelnut streusel:, ground hazelnut, flour, caster sugar, butter, for hazelnut white chocolate pressed sable:, hazelnut streusel, cocoa butter, puffed rice, clarified butter, melted white chocolate, for flexy berry:, raspberry puree, sugar, liquid glucose, pectin
strawberry & pistachio breton tart	vitamin_a	veg	plain flour, baking powder, sea salt, unsalted butter, egg yoks, castor sugar, double cream, yolk, sugar, gelatine, butter, pistachio paste, strawberry
saewoo bokumbop (shrimp fried rice)	vitamin_a	veg	cooked rice, shrimp (de-veined, onion, cooked green peas, green onion, egg (scrambled), vegetable oil/butter, soy sauce, sesame oil, salt, pepper
shrimp & cilantro ceviche	vitamin_a	veg	prawns, gherkin, onion, cilantro, mix bell pepper, tiger milk, sweet corn, sea salt, black pepper, green lemon juice, cherry tomato, edible flower, coriander stems, celery stalks, garlic, ginger, red onion, lemon juice, sea salt, black pepper
gluten-free christmas cake	vitamin_a	veg	christmas dry fruits (pre-soaked), orange zest, lemon zest, jaggery syrup, almond flour, apple, butter (softened), eggs
sweet chilli almonds	vitamin_a	veg	almonds whole, egg white, curry leaves, salt, sugar (fine grain), red chilli powder

Figure 11: Output 2

CHAPTER 9 CONCLUSIONS

The goal of this project was to use the available collection of recipe data to build a recommendation system. Train, evaluate and test a model to predict cuisines from sets of ingredients according to the diseases and input given by the users. The proposed system recommends the diet plan by using the Random Forest algorithms. This algorithm is preferred because it provide better solutions for real world problems and accuracy. The project is intended to provide a recommender system which accepts inputs from the user such as age, height, weight, food preferences and alcoholic/smoker. The input is given using the GUI created using Java. Based on the information provided a list of ingredients is given as output.

CHAPTER 10 FUTURE WORK

- As future enhancements to the approach used in this project are the creation of a more exhaustive and accurate vocabulary for ingredients such as a conversion system between all the identified units to grams.
- Will focus on the recommendation's breakdown for different timings of the day as per prescribed medicines such as (after/before) food meals to be taken along with medicines. Moreover, we will consider the amount of nutrition in different food items as per timing and daily needs of the patients. Furthermore, group food recommendation for family/friends is another interesting research area that can be explored.

REFERENCES

[1] Romeshwar Sookrah, Jaysree Devee Dhowtal and Soulakshmee Devi Nagowah, "A DASH Diet Recommendation System for Hypertensive Patients Using Machine Learning", 7th International Conference on Information and Communication Technology-2019.

[2] V.Nallarasan,L.Anand and J.Prabakaran, "Improved Predictive Learning Approaches For Customized Diet Suggestion Framework In Medical Services", Journal of Xi'an University of Architecture & Technology,SRM Institute of Science and Technology, Chengalpattu Dt., India.

[3] Antika Thapar and Mehar Goyal, "A fuzzy expert system for diagnosis of malnutritionin children", IEEE-2017, Department of Mathematics, Faculty of Science, Agra, INDIA,

[4] Ananya Banerjee, "Nourishment Recommendation Framework for Children Using Machine Learning and Matching Algorithm", International Conference on Computer Communication and Informatics-2019.

[5] Sakshi Singh and Sanjay Kumar Dubey, "Recommendation of Diet to a Patient usingAHP and Fuzzy Approach, 9th International Conference on Cloud Computing, Data Science & Engineering-2019.

[6] Chih-Han Chen, Maria Karvela, Mohammadreza Sohbati, Thaksin Shinawatra and Christofer Toumazou, "PERSON—Personalized Expert Recommendation System for Optimized Nutrition", IEEE Transactions on Biomedical circuits and systems, VOL. 12,NO. 1, FEBRUARY 2018.

[7] Arushi Singh,Nandini Kashyap and Rakesh Garg, ”Fuzzy based approach for diet prediction”, 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence).

[8] Dhruv Rao, Connor Higgins, Hartini Margot, Taylor Lyle, Shannon McFalls, Eric Obeysekare and Khanjan Mehta, “Micronutrient Deficiencies in the Developing World:An Evaluation of Delivery Methods”, IEEE 2016 Global Humanitarian Technology Conference.

APPENDIX

The data for the disease was initially collected from family and friends. But it could not be used in the project because the data was insufficient.

The data collected was name, height, weight, age, disease and food preference(veg/non-veg).

Github Link

https://github.com/mariajyothis/Major-Project