VITAMIN DEFICIENCY AND FOOD RECOMMENDATION SYSTEM

¹Mrs K .Prathyusha, ²T.Ishwarya, ³N.Pravalika, ⁴V.Sowmya

¹Assistant Professor, Dept.of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad,

²BTech student, Dept.of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad,
thikariishwarya@gmail.com

³BTech student, Dept.of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad,
pravalika.nasu@gmail.com

⁴BTech student, Dept.of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad,
Sowmyavanga44@gmail.com

Abstract: The main objective of this project is to recommend a diet to different individual. The recommender system deals with a large volume of information present from the dataset. In this project own data set is prepared based on various high and low values of vitamins from (vitamin a, b, c, d, e, k) and features are divided from normal and abnormal conditions of vitamins and labels are divided in to 0 and 1 as normal and abnormal. Another dataset is prepared based on combination of various vitamins and their deficiency and food to be recommended based on which vitamin is deficient. In this project multiple classifier algorithms are used (KNN, Decision Tree, Random forest, Logistic regression, voting classifier) ensembled algorithm is used to combine multiple algorithms and train a new algorithm. Accuracy of each algorithm is calculated and best algorithm is used for prediction purpose. Prediction is shown using flask web application which will detect deficiency of vitamin and recommend type of food to be taken on various combinations.

Keywords: Vitamin Deficiency, Food Recommendation, machine learning

I. INTRODUCTION

Nowadays, a human being is suffering from various health problems such as fitness problem, inappropriate diet, mental problems etc. Various studies depict that inappropriate and inadequate intake of diet is the major reasons of various health issues and diseases. A study by WHO reports that inadequate and imbalanced intake of food causes around 9% of heart attack deaths, about 11% of ischemic heart disease deaths, and 14% of gastrointestinal cancer deaths worldwide. Moreover, around 0.25 billion children are suffering from Vitamin-A deficiency, 0.2 billion people are suffering from iron deficiency (anemia), and 0.7 billion people are suffering from iodine deficiency. The main objective of this work to recommend a diet to different individual. The recommender

system deals with a large volume of information present by filtering the most important information based on the data provided by a user and other factors that take care of the user's preference and interest. It finds out the match between user and item and imputes the similarities between users and items for recommendation based on their physical aspects (age, gender, height, weight, body fat percentage), preference (weight loss or weight gain). The recommendation process has basically three stages that are Information Collection Phase, Learning Phase and Recommendation Phase. The information is firstly collected about a particular problem and the various solutions related to that problem are categorized. After the collection of information Learning Phase comes in which various conclusions are made out of that information which is gathered and in last phase i.e., Recommendation Phase an output is given in which various recommendations are made. In our project the output of recommendation is based on user's physical aspects, preference and their Body mass Index (BMI)Balanced nutrition is important aspect of healthy lifestyle for peoples. Along with balanced diet, a regular physical exercise is crucial for healthy life.

Now a day's nutrition and health are often overlooked. The majority people suffering with diabetes, heart disease, cancer, stroke etc. The diseases are almost directly related to unhealthy eating habits. So, our body needs nutrients to stay healthy, and food supplies essential nutrients that stop us from getting sick. A healthy, balanced diet will usually include vitamins, minerals, protein, healthy fats, proteins, carbohydrates, and fiber. A healthy food pyramid is combination of plant foods, moderate amount of animal products. Which includes vegetables, grains, fruits, oils and sweets, dairy, meat and beans. Generally, a person remains unaware of major causes behind deficiency or excess of various vital substances, such as calcium, proteins, and vitamins, and how to normalize such substances through balanced diet. With the advantage of technology, the people can leave a healthier lifestyle. In this project to build a system that will aim to recommend appropriate nutrition intake to its users based on body mass index (BMI) and grocery data preferences. BMI calculate weight status categories which includes underweight, healthy weight, overweight, obese. Grocery data includes seasonal food, user's intreated food, plant foods and animal products.

For people, balanced diet is a vital part of a healthy lifestyle. A healthy lifestyle requires a balanced diet as well as regular physical activity. Nutrition and health are

frequently disregarded nowadays. The majority of people have diabetes, heart disease, cancer, stroke, and other diseases. The diseases are virtually always linked to poor dietary habits. So, in order to stay healthy, our bodies require nutrition, and food provides these necessary elements. Vitamins, minerals, protein, healthy fats, proteins, carbs, and fibre are often found in a healthy, balanced diet. Plant foods with a moderate amount of animal products make up a healthy dietary pyramid. Vegetables, grains, fruits, oils and sweets, dairy, meat, and legumes are all included. In most cases, a person is unaware of the major causes of deficiency or illness. excesses of several critical nutrients, such as calcium, proteins, and vitamins, and how to restore equilibrium with a well-balanced diet People can live a healthier lifestyle thanks to technological advancements. The goal of this project is to create a system that will advise users on proper nutrition intake based on their BMI and grocery data preferences. Underweight, healthy weight, overweight, and obese are all weight status categories calculated by BMI. Seasonal foods, user-intreated foods, plant foods, and animal products are all included in grocery data

II. LITERATURE SURVEY

Raciel yera toledo proposed a food recommender system considering nutritional information and user preferences. The meal plan for the user recommended using users' preferences. This tool manages both user preferences and nutritional information. Vijay Jaiswal proposing a healthy food habit, eating patterns and calories burned count can be intake of nutrients and so on using the data mining tools. In this tool the hidden patterns and customer food taking habits are found from different data sources. In this tool decision tree learning algorithm, Random Tree algorithms are used on different datasets.

H. Jiang proposed a system to calculate the daily calorie demand. The Knapsack algorithm is used for recommended diet combinations of users. Different from other diabetic diet recommendation systems, this system can rank the recommended diet combinations using TOPSIS algorithm according to user's food nutrition. Jung-Hyun Lee proposed a customized diet recommendation service managing heart diseases. This service provides customers customized general information, family history of diseases, seasonal food intakes.

Rung-Ching Chen constructs a recipe ontology that defines some common diseases healing with verity of food recommendations and an inference engine for customer health condition and a recipe ontology can be used for proper recipe recommendations on food priorities.

FidelsonTanzil uses ABC algorithm to extract information from database according to user's requirements. K mean and SOM algorithms are used on datasets.

Mohd Afisi projected ABC algorithm in Data Mining and tested compared to six traditional

classification algorithms successfully and ABC proved as a suitable algorithm for recommendation.

Xiaoyan Gao proposed the food recommendation problem on user choice recipe recommendation factors. By using a neural network-based solution on Ordered diet Recommendation. The authors INGMAR WEBER and PALAKORN ACHANANUPARP [1] made an attempt to gain insights from machine leaned - diet success prediction which would help people trying to stay fit and healthy by keeping a track on their dietary intake. The authors used public food diaries of more than 4,000 long-term active MyFitnessPal users to study the characteristics of an unsuccessful diet. Concretely, authors trained a machine learning model to predict repeatedly being over or under self-set daily calorie goals and then look at which features contribute to the model's prediction, where research was centered around "quantified self " data. The authors observed that classification performance was sufficient and the token-based model performed better than the category-based

model and used such data feasibly for more in-depth data mining.

Nandish Shah et al. [2] presented a proposal of healthy food habits and eating system based on web data mining, to discover hidden patterns and business strategies from their customer and web data, which would track eating habits and recommend the types of food that will improve the health and avoid the types of food that raise the risk of illness. The authors used data mining algorithms like classification, clustering, association rules, etc. in the data mining process to extract useful information about people's eating habit. The nutritive structure of each kind of food was analyzed and the fat, energy, vitamin percentage in the recipe was calculated. Then they used the classification mining algorithm to process the composition data and give out the result whether the diet is healthy or not. As a result, personalized recommendations were suggested for each person.

How a coding system at the meal level might be analyzed by using data mining techniques was demonstrated by the authors Aine p. Hearty and michael j. Gibney [3] through this article. They evaluated the usability of supervised data mining methods to predict an aspect of dietary quality based on dietary intake with a food-based coding system and a novel

meal-based coding system. The authors used Food consumption databases from the North South Ireland Food Consumption Survey 1997–1999. A healthy eating index (HEI) score was developed. Quintiles of the HEI based on combinations of foods were predicted by Artificial neural networks (ANNs) and decision trees. As a result, the ANN had a slightly higher accuracy than did the decision tree in relation to its ability to predict HEI.

However, on the basis of the meal coding system, the decision tree had higher accuracies than did the ANN. Data mining was used by Christy Samuel Raju, Sanchit [4] to develop a Fitness Advisor System. "Fitness Advisor" developed by authors was a desktop application that advised the user according to his/her problem associated with body weight by an efficient diagnosis of the same and spreading proper awareness about the health hazards. The authors considered different factors in the system such as height, weight, body type, sex, smoking, drinking, health condition, physical activity, sleeping hours etc. A combination of clustering, association and classification algorithms to effectively deliver the best possible expert advice to the user's problem

was used by authors. Apriori algorithm was used by authors for generating association rules. The final output of the system was expert's advice in terms of diet and exercise.

III. PROPOSED WORK

The System works in a Machine Learning Environment, we use multiple machine learning algorithms to check accuracy of vitamin deficiency and food recommendation and best model is used for prediction in flask web application. When user enters vitamin values algorithm will predict deficiency is vitamin and recommend food

- Automates process of vitamin deficiency detection and food recommendation
- Previous datasets are used to training and testing.
- Accuracy of model is improved compare to existing methods.

MODULES

In this project work, we have used five modules and each module has own functions, such as:

1. Data collection
2. Data pre-processing
3. Testing training
4. Initializing Multiple Algorithms
5. Predict data

Data Collection:

In this project we are using vitamin dataset and food recommendation dataset which is prepared based on min and max vitamin values from the test results and features are min and max values of vitamin a, b, c,d,e,k values and labels are deficiency and non-deficiency. Based on the vitamin deficiency food data set is prepared with various combinations. In this feature are vitamin deficient values and labels are type of food.

Data Pre-processing:

Features are extracted from data set and stored in variable as x train variable and labels are stored in y train variable. Data is pre-processing by standard scalar function and new features and labels are generated.

Testing Training:

In this stage data is sent to testing and training function and divided in to four parts x test train, and y test train. Train variables are used for passing to algorithm whereas test are used for calculating accuracy of the algorithm.

Initializing Multiple Algorithms:

In this stage machine learning algorithms are initialized and train values are given to algorithm by this information algorithm will know what are features and what are labels. Then data is modelled and stored as pickle file in the system which can be used for prediction. Data set is trained with multiple algorithms and accuracy of each

model is calculated and best model is used for prediction.

Predict data:

In this stage new data is taken as input and trained models are loaded using pickle and then values applicare pre-processed and passed to predict function to find out result which is showed on web ation

IV.   METHODOLOGY

we can see how the data is divided into different sets and then trained for different models. The dataset was first divided into training set (80%) and pre-training set (20%). The pre-training set was divided into pre-train (80%) and pre-test (20%) .The training set is further is divided into train(80%) and validation set(20%). This train set is again divided into train(80%) and test set(20%). So, now we have train validation and test sets separate which are non overlapping. The pre-train set was used to find the best models for the given dataset. I took best 4 models using pretest set. Their performance was compared based on their mean absolute errors. • Once the best 4 models were obtained, hyper parameters for these models were tuned and the best parameter was selected.

ALGORITHMS

KNN Classifier Algorithm:

K-nearest neighbor method can be used for both regression and classification predictive problems. This method helps in interpret output, calculate time and predictive power. The Machine learning techniques are used in various fields. KNN is also one of the machine learning methods. This is also called as method of sample-based learning. This will contain the data of past datasets and can be used while predicting the new datasets. This will apply function called as distance function like Manhattan or Euclidean distance. This can be used to compute distance from samples to all other training samples. It calculates the target value for new samples. The target vale will be the weighted sum of target values of the k nearest neighbours. The valve of K can be directly proportional to the prediction. Whenever the valve of K is small this indicates there is high variance and there is low bias. If the valve of the K is larger than this indicates that there is low variance and high bias. The main advantage of this KNN is it does not require any training or the optimization. This KNN uses data samples when predicting the new datasets. Hence it is having higher complexity and also more time consumption.

This work represents a review of K-NN technique for the early prediction of food recommendation. K-NN analysis is used for predicting the unknown parameter from the known parameters. In this work we are considering vitamins as input parameters which are the main parameters to be considered for a good food recommendation, although there are many other factors that can be considered. The unknown value of vitamis can be predicted from the nearest known values of the nearest neighbors by calculation of Euclidean distance between them.

Then we would be able to predict type of food for given vitamin parameters. To measure the distance between points in a feature space, various distance functions can be used, in which the Euclidean distance function is the most widely used one. Let p and q are represented as feature vectors. To calculate the distance between p and q, the Euclidean metric is generally used by if a=(a1, a2) and b=(b1,b2) then the distance is given

Classifications Algorithms

Onto the part you've probably been waiting for all this time: training machine learning algorithms. To be able to test the performance of our algorithms, I first performed an 80/20 train-test split, splitting our balanced data set into two pieces. To avoid overfitting, I used the very common resampling technique of k-fold cross-validation. This simply means that you separate your training data into k

parts (folds) and then fit your model on k-1 folds before making predictions for the kth hold-out fold. You then repeat this process for every single fold and average the resulting predictions.

To get a better feeling of which algorithm would perform best on our data, let's quickly spot-check some of the most popular classification algorithms:

- Logistic Regression
- Linear Discriminant Analysis
- K Nearest Neighbors (KNN)

.

- Classification Trees
- Support Vector Classifier
- Random Forest Classifier

SYSTEM ARCHITECTURE

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system. Organized in a way that supports reasoning about the structures and behaviors of the system
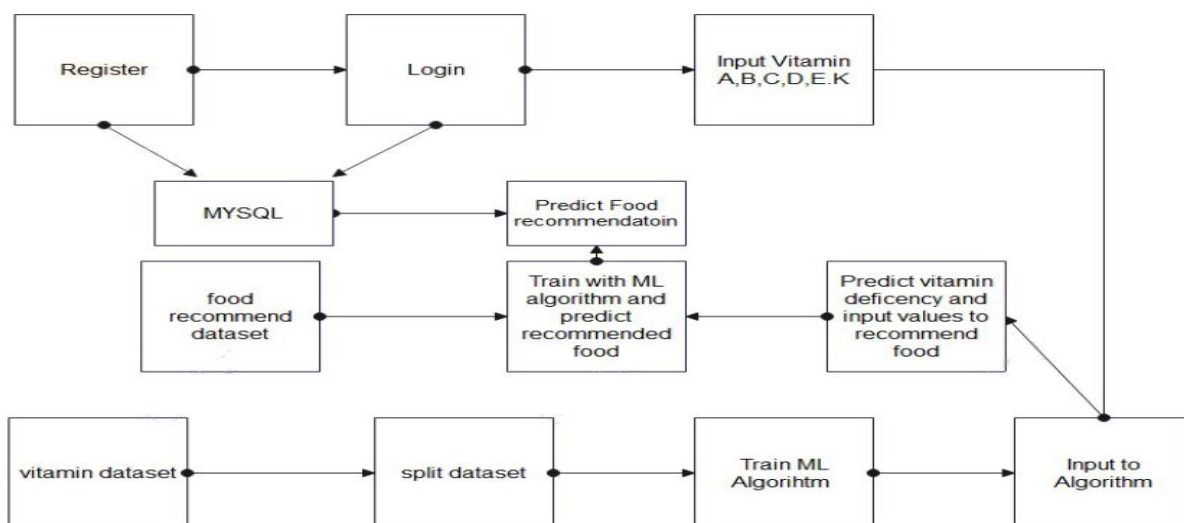
Figure.1 Proposed architecture

3-Tier Architecture:

The three-tier software architecture (a three-layer architecture) emerged in the 1990s to overcome the limitations of the two-tier architecture. The third tier (middle tier server) is between the user interface (client) and the data management (server) .

V. RESULTS

components. This middle tier provides process management where business logic and rules are executed and can accommodate hundreds of users by providing functions such as queuing, application execution, and database staging
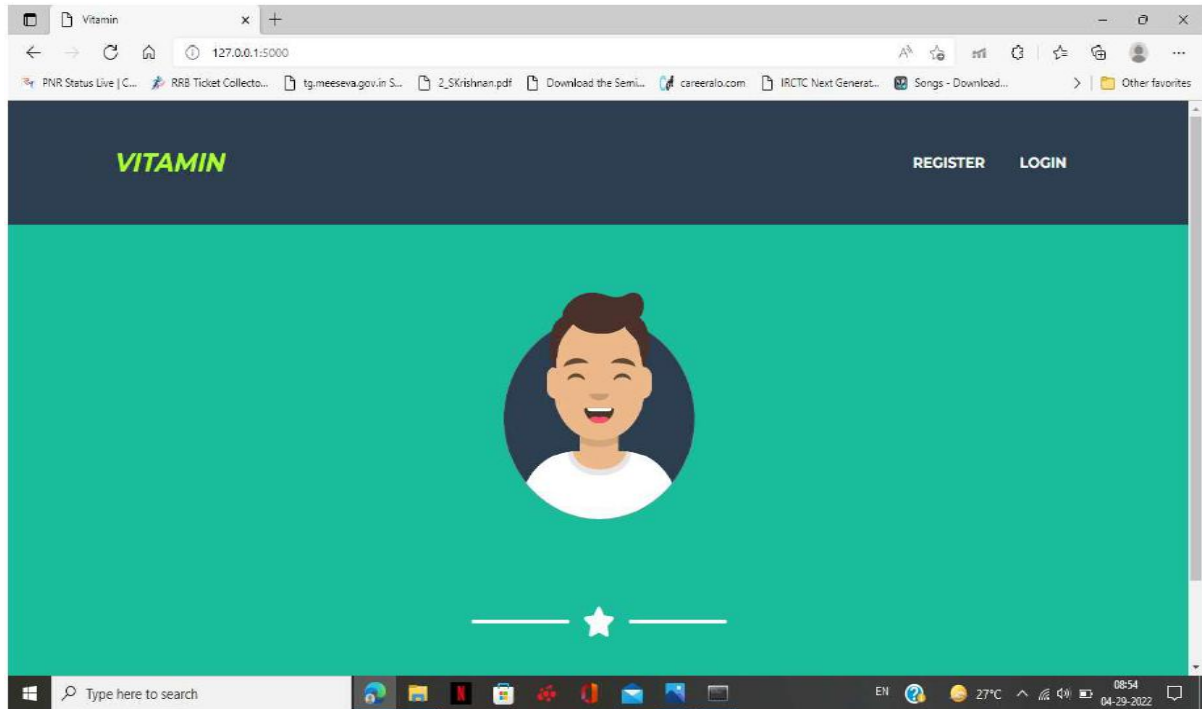
Fig.2 Home page

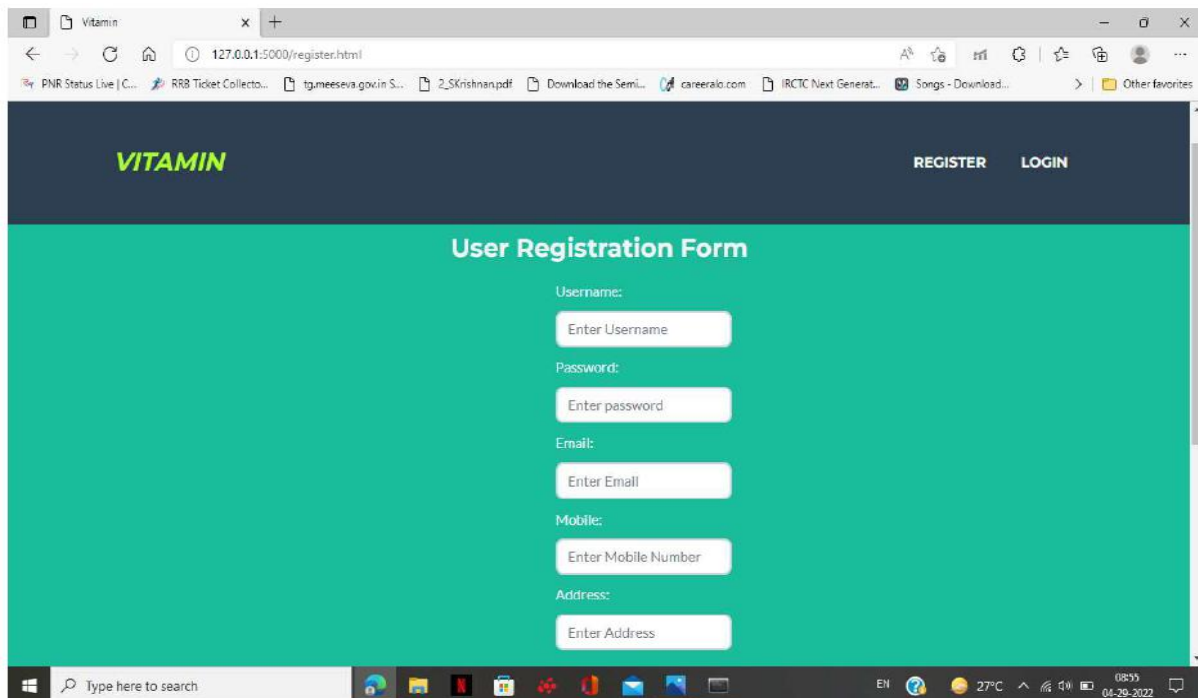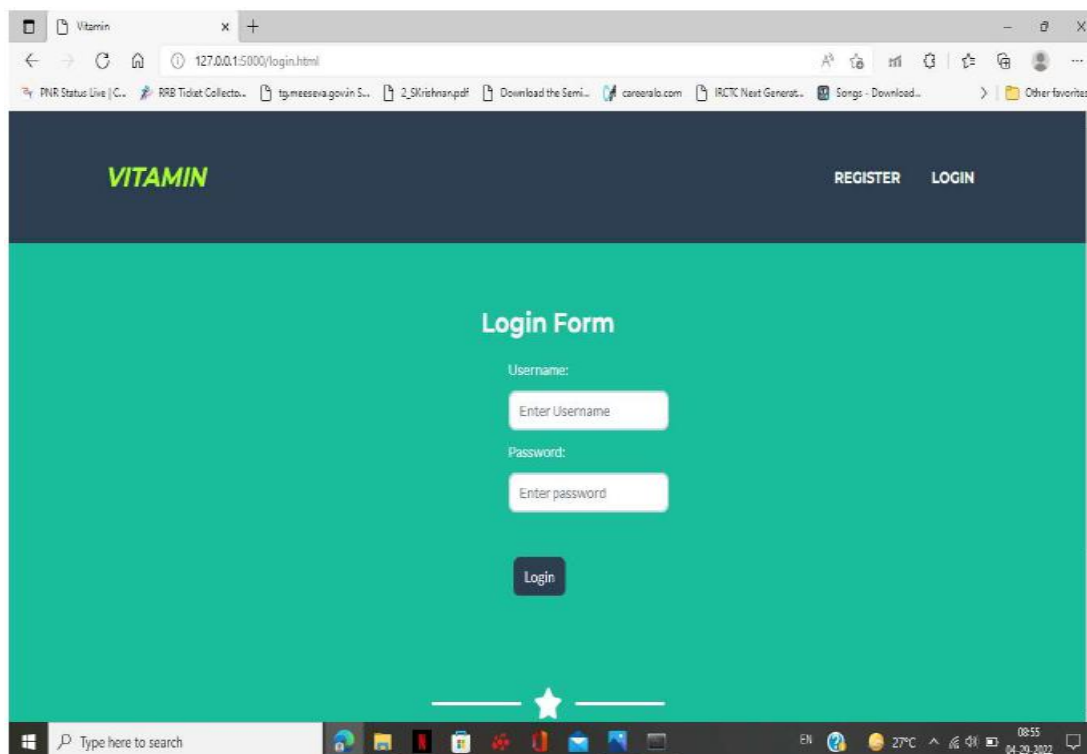Fig.3 Register page

Fig.4 Login page

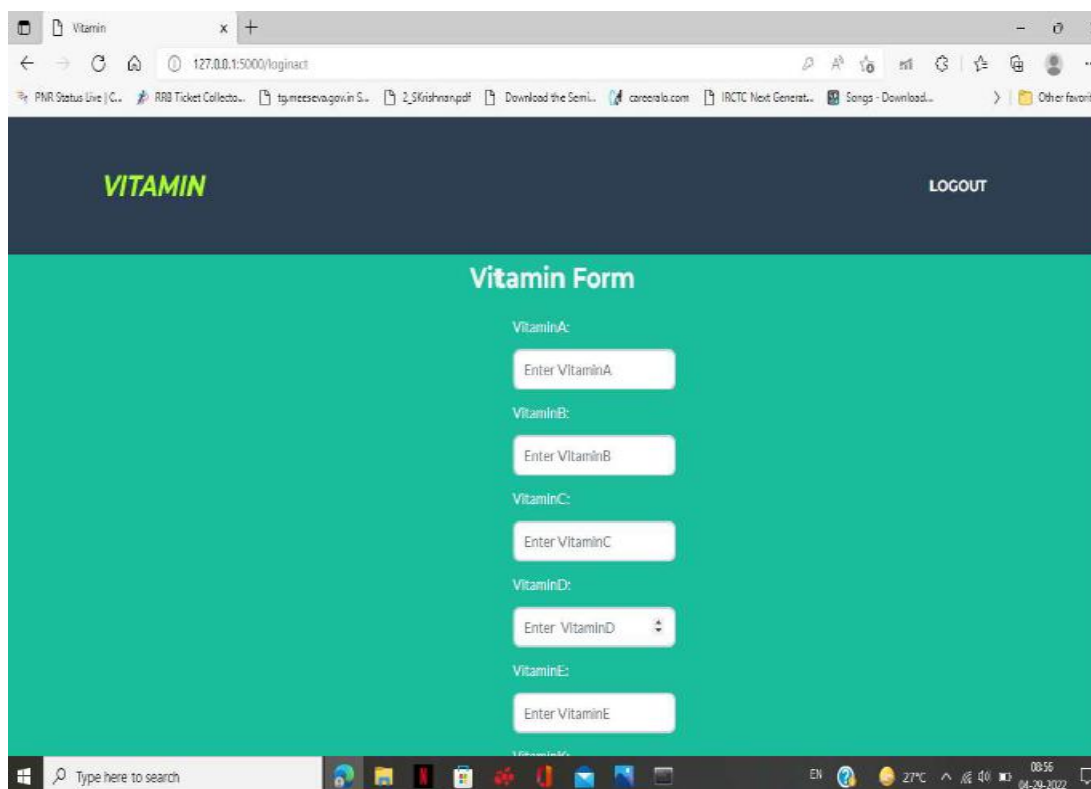Fig.5 Vitamin form

Fig.6 Result

Fig.7 Food recommendation

```
          precision    recall   f1-score    support

      0       1.00       0.90       0.95        21
      1       0.82       1.00       0.90         9

   accuracy                         0.93        30
  macro avg     0.91       0.95       0.93        30
weighted avg    0.95       0.93       0.94        30

[[19  2]
 [ 0  9]]
Accuracy of Support Vector Machine 93.33333333333333 %
```
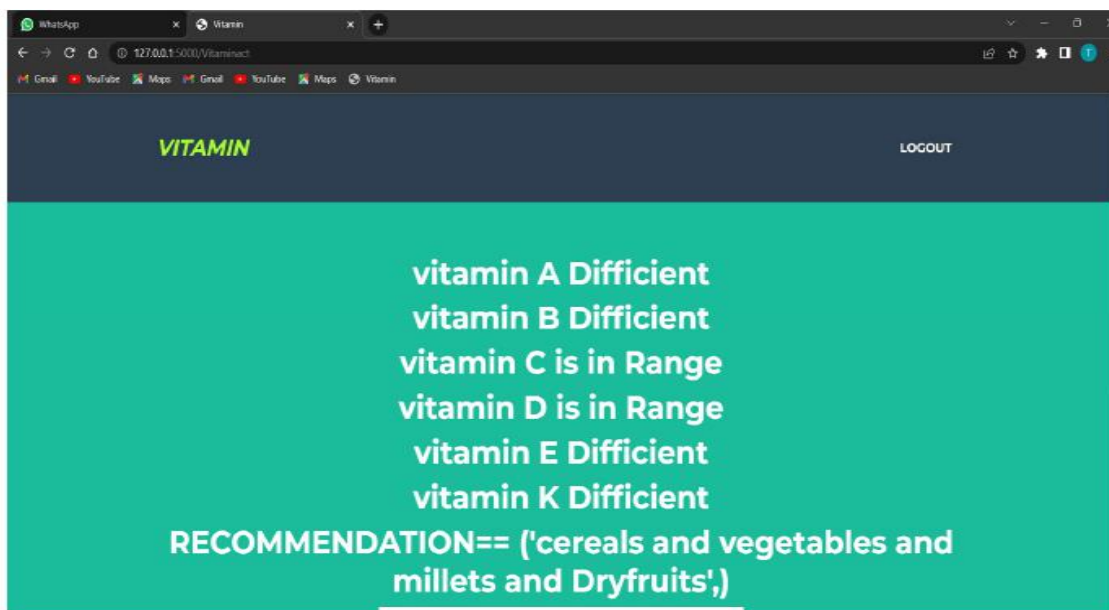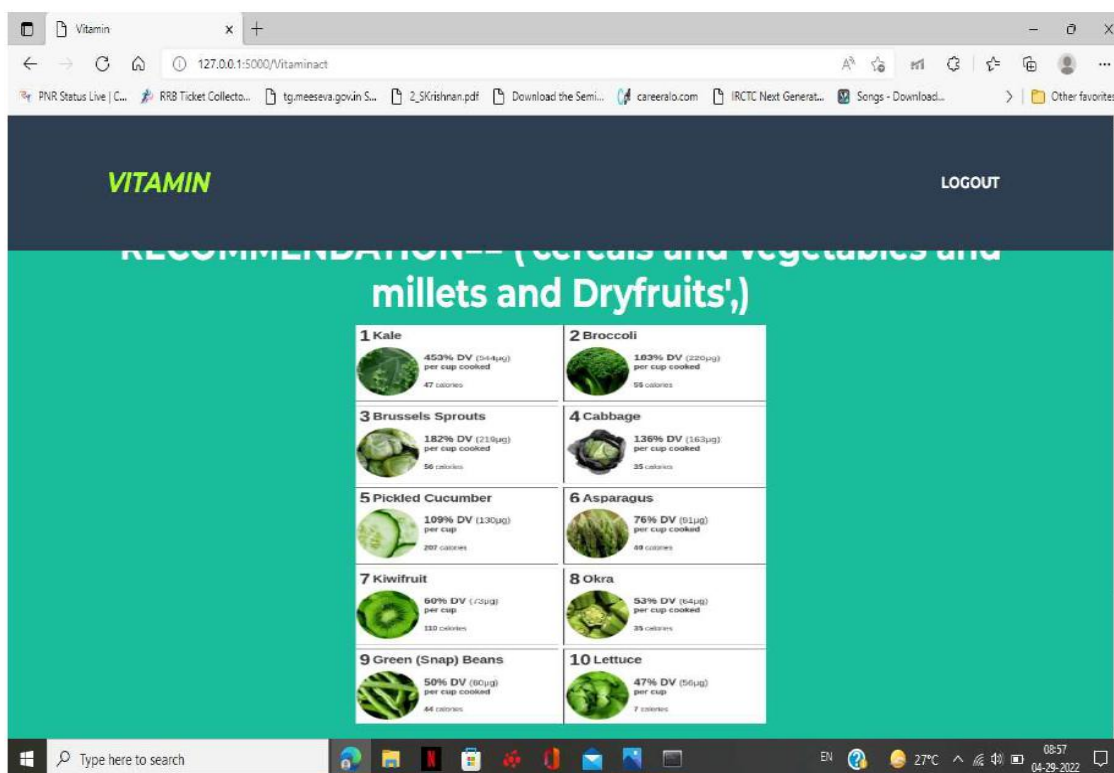
Fig.8 Final accuracy

VI. CONCLUSION

We have created a website which recommend the food items and predicts vitamin deficiency in which we have implemented prediction by taking input as vitamins and their deficiency. For training of the system, the initial process involves the dataset preparation of food items depending upon the vitamin deficiency. The prediction of various food recommendation depending upon which are essential for the for type of vitamin deficient. After the clustering is performed, using Random Forest classifier, the nearest food items are predicted which best suited for the appropriate diet. Our diet recommendation system allows users to basically get the desired healthy diet on the bases of vitamin deficiency.

REFERENCES

[1] Phanich, M., Pholkul, P., & Phimoltares, S., "Food recommendation system using clustering
analysis for diabetic patients," in Proc. of International Conference on Information Science and Applications, pp. 1-8, IEEE, April 2010. Article .

[2] Ge, M., Elahi, M., Fernaández-Tobías, I., Ricci, F., & Massimo, D., "Using tags and latent
factors in a food recommender system," in Proc. of the 5th International Conference on Digital
Health, pp. 105-112, ACM., May 2015.

[3] Freyne, J., & Berkovsky, S., "Evaluating recommender systems for supportive technologies,"
User Modeling and Adaptation for Daily Routines, pp. 195-217, Springer London, 2013.

[4] Prof. Prajkta Khaire, Rishikesh Suvarna, Ashraf Chaudhary, "Virtual Dietitian: An Android

based Application to Provide Diet", International Research Journal of Engineering and Technology

(IRJET),Volume: 07 Issue: 01 | Jan 2020

[5] Shivani Singh, Sonal Bait, Jayashree Rathod, Prof. Nileema Pathak," Diabetes Prediction

Using Random Forest Classifier And Intelligent Dietician " , International Research Journal of

Engineering and Technology (IRJET), Volume: 07 Issue: 01 | Jan 2020