ABSTRACT

Nowadays, the number of diabetic patients are increasing worldwide. But there is an inability to access their diet accurately raised the development of this system. Diabetes is a long-term condition that produces high blood sugar levels. Three types of diabetics are identified in the past by experts. Mostly diabetes symptoms include frequent urination, intense thirst and hunger, weight increase, unusual weight reduction, cuts and bruises that do not heal, men with sexual dysfunction, numbness and tingling in hands leg and feet. Recent studies have shown that 70%-80% of Type2 diabetes complications can be prevented by early identification and intervention in people at risk by changing their lifestyle or by therapeutic methods. Several data mining and machine learning methods have been proposed in the past for the diagnosis and management of diabetes. In this work, a food recommendation system is proposed. For classification of the food items, four features namely size, shape, color and texture are used in this work. The system is built for food image processing and uses nutritional fact table for calorie measurement. Three techniques are adopted to extract features. From this work, it is shown that employing Computer Aided Diagnostic systems as a “second opinion” has lead to improved diagnostic decisions and Support Vector Machines (SVMs) have shown remarkable success in this area. Hence, it is necessary to predict the disease level and recommend the food suitably.

Diabetic patients are diagnosed by involving individuals at risk of developing Type2 Diabetic Mellitus. The system proposed consists of the following modules namely User Interface, Image Database Manager, Diabetic Analyzer, Food Recommender, Agent Subsystem, Decision Manager, and Fuzzy Rule Manager. Three databases have been used in this work namely

image dataset, diabetic dataset and food dataset. The decision manager monitors the overall process of this proposed system. The decision manager makes the decisions about the food analysis, disease prediction and food recommendation activities with the help of rules present in the rule base. After extracting these features, the image is classified using multiclass SVM to identify the class of provided food image. By finding the area and volume of the food samples, calorie values are calculated. The multiclass SVM methods are compared with binary SVM against the food samples.

In the disease prediction phase, intelligent and effective methodology for the automated diagnosis of diabetes is proposed in this work using Twin Support Vector Machine (TSVM) classifier algorithm. Rules are extracted from the trained TSVM model. Evaluation is carried out on these techniques to measure the accuracy and performance of the system. Results have shown that the degree of accuracy for TSVM classifier shows better results compared to other techniques. The steps of disease prediction are as follows: Food Recommendation is performed using Self-Organizing Map (SOM) and k-means clustering. The SOM composed of map units called nodes, in which SOM algorithm computes a model. These models are the representation of the input space of the training samples. Boosting is the general method which converts any weak learning algorithm into strong learner for accuracy improvement. The limitations in boosting are over fitting on the training data and filtering out the correct data in the subsequent functions since boosting concentrates on regions not predicted well by other learners. So, the approach named Cluster Based Boosting is used to address limitations in boosting. In this work, initially X-Means algorithm is used to cluster the data and those clusters are selectively boosted depending on the additional structure information that are provided by clusters and previous function’s accuracy on the data member. To apply Cluster Based Boosting to the high dimensional data, dimensionality reduction technique is performed.

In this work, we apply Global Redundancy Minimization frame work which considers the redundancy of the feature with all other features. The selected features will contribute more mutual information for prediction. This frame work can be used with any other feature selection technique. We provided experimental results on various dataset. These results demonstrate the effectiveness of Global redundancy framework and also effectiveness of Cluster Based Boosting with Global redundancy minimization framework than classifier with global redundancy minimization framework.