ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management

Volume 10, Issue 4, July 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 6.551

+91 9940572462 +91 9940572462 ijarasem@gmail.com www.ijarasem.com

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 4, July 2023 |

A Study on Machine Learning Based Nutritional Recommendations for Individuals with Diabetes

Venkatesh Kondaveti, Dr. Prasadu Peddi, Dr. G. Venkata Subba Raju

Research Scholar, Dept. of Computer Science and Engineering, Shri JJT University, Rajasthan, India

Research Supervisor, Dept. of Computer Science and Engineering, Shri JJT University, Rajasthan, India

Co-Supervisor, Professor, Stanley College of Engineering and Tech for Women, Hyderabad, India

ABSTRACT: Diabetes is a serious metabolic condition that can have a significant impact on the entire body. These days, diabetes has turned into a typical sickness to humanity from youthful to old. The quantity of revealed diabetic patients is heightening step by step, because of incalculable reasons harmful or synthetic items blended in with the food, corpulence, working society and terrible eating regimen plan, surprising way of life, eating food propensities and natural elements. Diabetes is a chronic metabolic disorder affecting millions of people worldwide. Proper management of diabetes requires maintaining a well-balanced diet tailored to individual needs. In recent years, machine learning techniques have gained popularity in various domains, including healthcare. This study aims to investigate the application of machine learning algorithms to provide personalized nutritional recommendations for individuals with diabetes. The proposed approach for this management system handles the various factors that affect the health of people with diabetes by combining multiple artificial intelligence algorithms.

KEYWORDS: Machine learning, artificial intelligence, Nutritional recommendations.

I. INTRODUCTION

Diabetes is a constant sickness that happens either when the pancreas doesn't deliver sufficient insulin or when the body can't really utilize the insulin it produces. Blood sugar is controlled by a hormone called insulin. Hyperglycaemia, or raised glucose, is a typical impact of uncontrolled diabetes and, over the long haul, prompts serious harm to a large number of the body's frameworks, particularly the nerves and veins. Diabetes affected 8.5% of adults worldwide over the age of 18 in 2014. Diabetes was directly responsible for 1.6 million deaths in 2016, and high blood glucose was responsible for another 2.2 million deaths in 2012. There are two kinds of diabetes, specifically, Type 1 and Type 2. Type 1 diabetes is described by insufficient insulin creation and requires day to day organization of insulin. Type 1 diabetes has no known cause and cannot be avoided with current knowledge. Type 2 diabetes results from the body's insufficient utilization of insulin. Type 2 diabetes comprises what the majority of people with diabetes around the world have and is mostly the result of excess body weight and physical inactivity. According to the World Health Organization (WHO), to help prevent type 2 diabetes and its complications, people should achieve and maintain healthy body weight; be physically active at least 30 minutes of regular, moderate-intensity activity on most days. More activity is expected for weight control, eat a sound eating regimen, keeping away from sugar and soaked fats consumption, and keep away from tobacco utilize smoking builds the gamble of diabetes and cardiovascular sicknesses. Diet and exercise are important components of diabetes treatment, as are lowering blood glucose levels and other known risk factors for blood vessel damage.

It is evident from the WHO reviewed literature that the incidence of diabetes mellitus is ever increasing throughout the world in both developed and developing countries. A significant number of people living in both developed and developing countries are ever becoming sedentary. In addition, despite the fact that there is evidence to suggest that the complications of diabetes can be avoided, there are still diabetics who do not possess the necessary knowledge and skills to manage and control their condition by utilizing the technology that is available for healthier living and lifestyle changes.

The unpredictable learning is the learning of the descriptive form. The material is represented or outlined using this instruction. Clustering, relation law mining and so on are examples of unsupervised learning algorithms. The balance of supervised and unsupervised is semi-supervised learning. In this study author proposed a health care system for predicting and recommended a diet for the diabetes patients. In addition, the supervised learning algorithm is used to learn information about diabetes and to develop a diabetes prediction system for diabetes diagnosis. Also is used pre-

processing data set, feature selection with Machine learning by using Age, Diagnosis Duration, Diastolic Blood Pressure, Cholesterol level and Hemoglobin.

II. REVIEW OF LITERATURE

The resent reviews explore the potential applications of machine learning in diabetes care, including nutritional recommendations. It discusses the integration of machine learning algorithms with wearable devices, electronic health records, and dietary data to develop personalized diabetes management strategies. The review also highlights the importance of considering individual preferences and behaviors in providing effective nutritional recommendations.

Isfafuzzaman Tasin et al (2022) Diabetes is the most fatal non-communicable disease and affects 537 million people worldwide. Excessive body weight, abnormal cholesterol levels, a family history of diabetes, inactivity, unhealthy eating habits, and other factors are just a few of the many factors that can lead to diabetes. Expanded pee is one of the most well-known side effects of this illness. Diabetes can lead to a number of complications, including heart disease, kidney disease, nerve damage, diabetic retinopathy, and others. But its risk can be reduced if it is predicted early. An automated diabetes prediction system was developed using a private dataset of female Bangladeshi patients and a variety of machine learning methods in this study. The creators utilized the Pima Indian diabetes dataset and gathered extra examples from 203 people from a neighbourhood material processing plant in Bangladesh. This work makes use of the mutual information feature selection algorithm. A semi-managed model with outrageous inclination helping has been used to foresee the insulin highlights of the private dataset.

Salliah Shafi Bhat (2021) Diabetes is a serious metabolic condition that can have a significant impact on the entire body. These days, diabetes has turned into a typical sickness to humanity from youthful to old. Due to a myriad of factors, including toxic or chemical ingredients in food, obesity, a poor diet and working culture, unusual lifestyle, eating habits, and environmental factors, the number of reported diabetic patients is steadily rising. Consequently, diabetes diagnosis is crucial to life preservation. AI Strategies can be utilized to foster a proficient medical care framework to foresee different sort of diabetic sicknesses ahead of time. In this paper, a Diet Recommendation System (DRS) and Machine Learning techniques are used to diagnose diabetes and recommend the right diet for diabetic patients. The appropriate information investigation is utilized for the choice of legitimate eating routine for diabetic Patients.

Robert A. Sowah et al (2020) The food acknowledgment model was assessed with cross-entropy measurements that help approval utilizing Brain networks with a backpropagation calculation. The model accurately classified the images with given labels and corresponding accuracy, learning features of the images derived from local Ghanaian dishes that have particular nutritional value and importance for diabetic management. The model accomplished determined objectives by anticipating with high exactness, marks of new pictures. For specific calorie intakes, the food recognition and classification model achieved accuracy levels of over 95%.

Alison Gray et al (2019) Nutritional recommendations for adult diabetes patients (PWD) from a variety of scientifically based guidelines and resources are compiled in this chapter. It is intended to consolidate these guidelines into a single location and provide health care professionals who treat PWD with an overview of useful applications and advice. The segments are partitioned into parts of healthful substance, with related objectives for PWD, as well as audits of present wholesome subjects of interest, incorporating weight reduction consumes less calories in the ongoing press. The data additionally incorporates hotspots for additional consideration, and assets that can be used for PWD.

Problem statement

The problem addressed in this study is to develop a machine learning-based system that can effectively analyze various data inputs, such as demographic information, medical history, dietary patterns, and blood glucose levels, to generate personalized nutritional recommendations for individuals with diabetes. Provided the healthcare dataset, developing an intelligent framework that mines data using ML techniques and provide healthy diet recommendations is the problem considered. Similar to various recommendations pertaining to weather, products, online services and healthcare services, food is an important component of human lives. With such information, in the long run, healthy diet recommender systems play vital role to enable people to have better eating habits. There is need for a framework to generate healthy diet recommendations. NLP is used for pre-processing of data while the data is analysed using ML algorithms like KNN. A hybrid algorithm is proposed to have quality healthy diet recommendations to public.

Objectives

- Develop a machine learning model or ensemble of models that can predict personalized nutritional recommendations for individuals with diabetes.

- Compare the performance of the machine learning-based nutritional recommendation system with existing approaches, such as manual recommendations by dietitians

Naïve Bayes

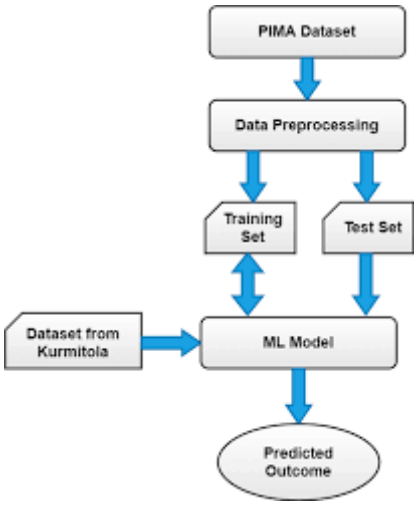
Naïve Bayes (NB) is a simple probabilistic classifier based on Bayes theorem. Te main assumption of NB is features are mutually independent. In the recent study, it is used to diagnosis of different types of disease especially diabetes disease. Te NB classifies the data using Bayes theorem as follows:

$$P((z|w_1, \ldots\ldots, w_n) = \frac{P(z)\pi_{i=1}^{n} P((w_i|z)}{\pi_{i=1}^{n} P(w_i)}$$

where, z is the response variable, W_1, W_2,…, W_n are input variables; $P(z|w_1, \ldots, w_n)$ is the conditional probability distribution of z given W_1, W_2,…, W_n; P(z) is the marginal probability distribution of z; $P(W_i|z)$ is the conditional probability distribution of W_i given z; p (W_i) is the marginal probability distribution of W_i ; π is the product symbol. We found the probability of z given the set of inputs and picked up the output with maximum probability. The corresponding classifier, a Bayes classifier is the function that assigns a class label as follows.

System Design and Development

The System's Requirements and Design Analysis. In the design and development of the architecture for the diabetes management system, the clinical requirements and design analysis of the system were based on discussions. From these discussions, the diet type of patients was determined to be an essential approach suitable for the diabetes management system. The following functionalities were mentioned: (1) Scheduling and reminding diabetic patients to take their medication and blood glucose readings, (2) recommending healthy meals for diabetics to keep their blood glucose levels in check, (3) encouraging and tracking the activity of diabetic patients, (4) providing a visual interface to help them make meaning of their readings and establishing a sufficient connection between the doctor and the diabetic patient using e-mail. Providing the diabetic patient with a data visualization tool to display the data in tables, charts, and an educational program for newly diagnosed and ongoing diabetes treatment is valuable for the treatment and management of diabetes.



The Supervised Learning/Predictive Models

Supervised learning algorithms are used to construct predictive models. A predictive model predicts missing value using other values present in the dataset. Supervised learning algorithm has a set of input data and also a set of output, and builds a model to make realistic predictions for the response to new dataset. Supervised learning includes Decision Tree, Bayesian Method, Artificial Neural Network, Instance based learning, Ensemble Method. These are booming techniques in Machine learning.

Unsupervised Learning / Descriptive Models

Descriptive models are developed using unsupervised learning method. In this model we have known set of inputs but output is unknown. Unsupervised learning is mostly used on transactional data. This method includes clustering algorithms like k-Means clustering and k-Medians clustering.

Semi-supervised Learning

Semi Supervised learning method uses both labeled and unlabeled data on training dataset. Classification, Regression techniques come under Semi Supervised Learning. Logistic Regression, Linear Regression are examples of regression techniques.

Machine Learning: Practical Advantages

There are practical advantages of ML. Since the computer learns itself to complete a task, time and effort need not be invested in instructing the computer on what to do. This not only saves time and effort that would otherwise be spent on programming but also increases adaptability to solve various problems. That is, the same algorithm can be retrained on various data sets and problems. On a similar note, ML accepts various data types as input, including structured (e.g., tabular data) and unstructured (e.g., image based). In some cases, the same ML algorithms can be applied to different problems, perform different tasks, and take as input different data types; neural networks and k-nearest neighbors (kNN) are such examples.

By predicting an outcome based on existing data, ML algorithms can save on the time and cost of having to verify such outcomes experimentally. The incorporating ML algorithms into the analysis of noninvasive and comparatively cheaper variables could avoid 81.9% of unnecessary ultrasound scans in NAFLD, which are expensive and have long waiting times for results.

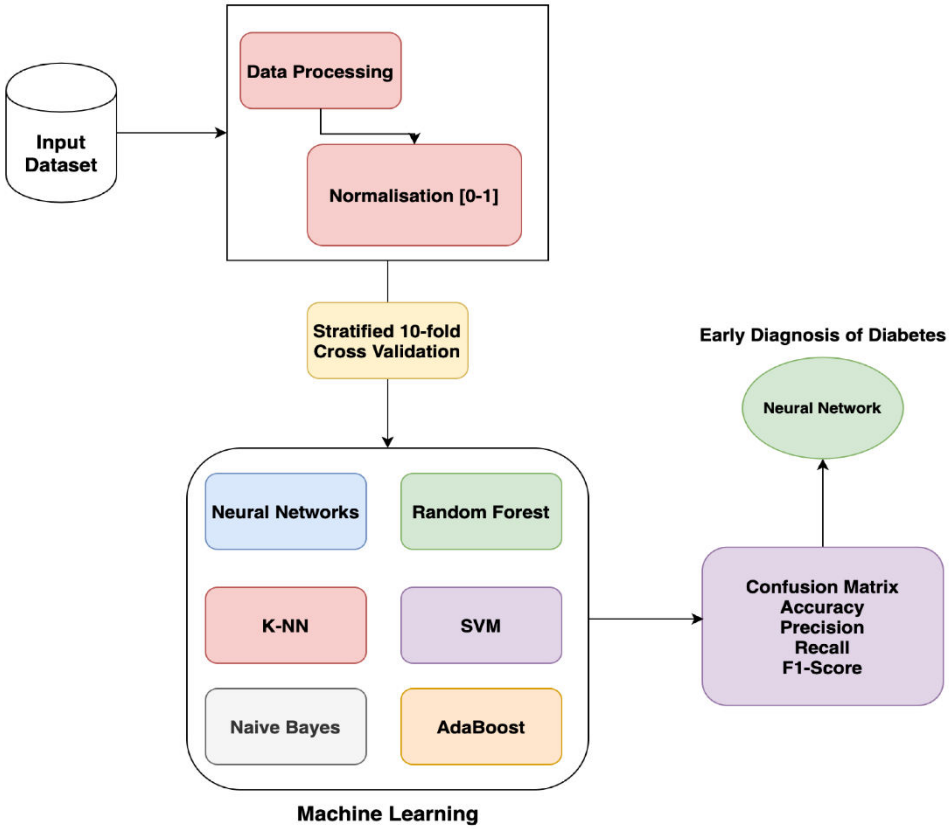


Figure 1: Activity Chart of Diabetes Prediction and Diet Plan

The proposed hybrid recommender system is based on both matrix decompositions. The two matrices aforementioned are used in order to have better recommendations and minimize RMSE. The algorithm is known as Intelligent Recommender for Healthy Diet.

Algorithm: Intelligent Recommender for Healthy Diet (IR-HD)

Inputs: User preferences P, food dataset D

Output: Healthy diet recommendations R

1. Start
2. Initialize food vector F
3. Initialize user vector U
4. For each tuple d in D
5. Populate F
6. Populate U
7. End For
8. Construct user-food matrix using Eq. 1
9. Construct ratings matrix using Eq. 2
10. Compute health score using Eq. 5
11. Compute result score using Eq. 6
12. Compute recommendations using Eq. 7 and populate R
13. Compute RMSE using Eq. 4
14. Return R
15. End

As presented in Algorithm 1, the pseudocode for the proposed recommender system is provided. Step 2 and Step 3 are used to initialize food and user vectors that are later used to construct matrices. The Step 4 through Step 7 is an iterative process to populate F and U from the food dataset. Afterwards, the user-food matrix is generated using Eq. 1 and ratings matrix is created using Eq. 2. Different computations are made before generating recommendations that are personalized. The results are the healthy diet recommendations that are stored in data structure R. The RMSE is computed for given results and the results are provided for the proposed algorithm and KNN in Section 4.

III. RESULTS

Empirical study is made with the Python data science platform based application. The results of KNN, Latent Factor Model (LFM) and the proposed hybrid algorithm are provided. RMSE is used as the metric to know the accuracy in prediction. The less in RMSE, the higher in accuracy.

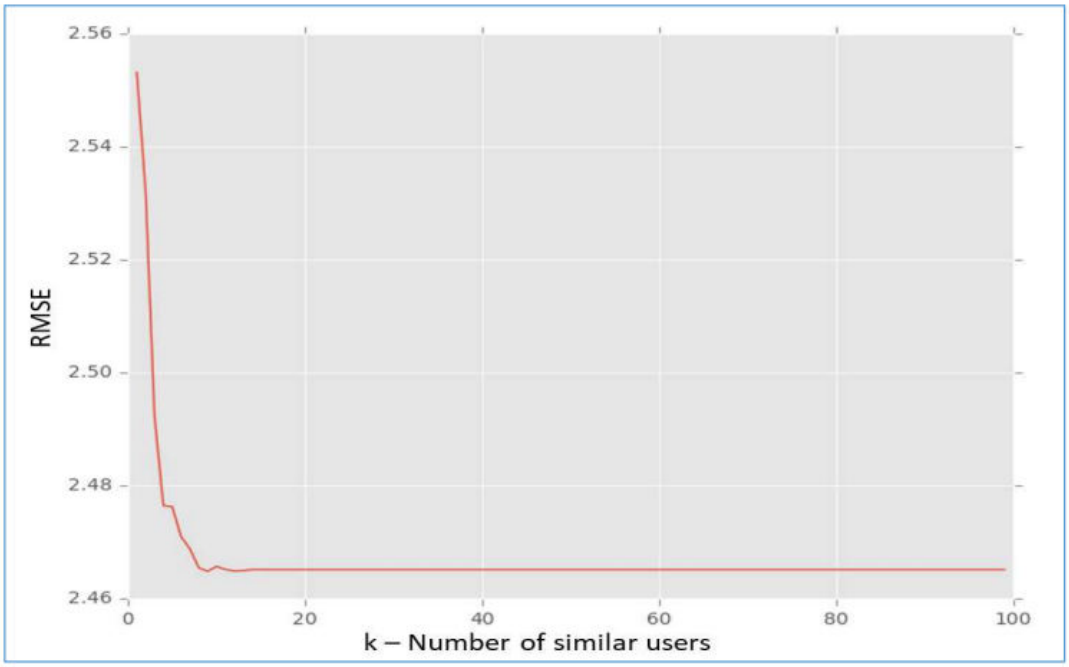


Figure 2: Cross-validation results of KNN

As presented in above Figure 2, the number of similar users is provided in horizontal axis and the vertical axis shows RMSE value. Less RMSE indicates higher in performance and the higher in RMSE indicates least performance. As the number of users is increased the RMSE is decreased indicating improvement in health score.

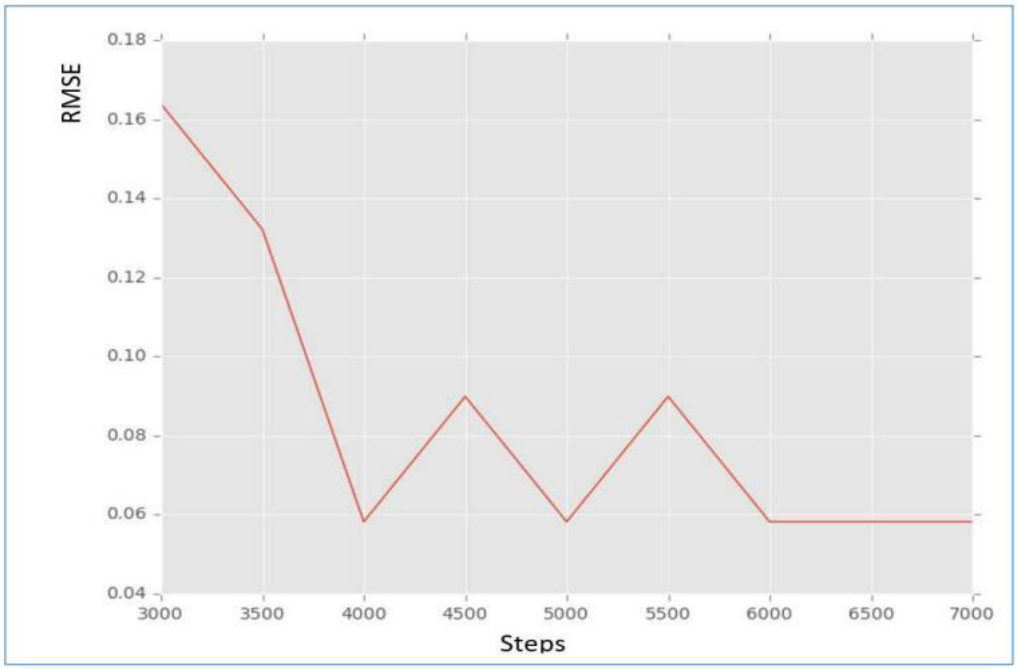


Figure 3: Cross validation results for the proposed hybrid algorithm

As presented in the above Figure 3, the steps used in the matrix factorization are shown horizontal axis. The values are taken from 3000 to 7000 incrementing by 500. RMSE on the other hand is shown in vertical axis. The results revealed that there is relationship between RMSE and steps. As the number of steps is increased, the RMSE came down showing improved performance.

Table 1: Statistical Comparison of Machine Learning Techniques

Method	Decision Tree	Naive Bayes	Random forest
Precision (%)	87	89	94
Recall (%)	77	80	86
F_Measure (%)	82	84	91
Accuracy (%)	86	89	94

Table 2: Performance comparison with different similarity metrics

Metrics	Euclidean	Cosine	Pearson
RMSE - Training Dataset	2.56594	2.59116	2.65868
RMSE - Whole Dataset	1.96741	2.39162	1.91329

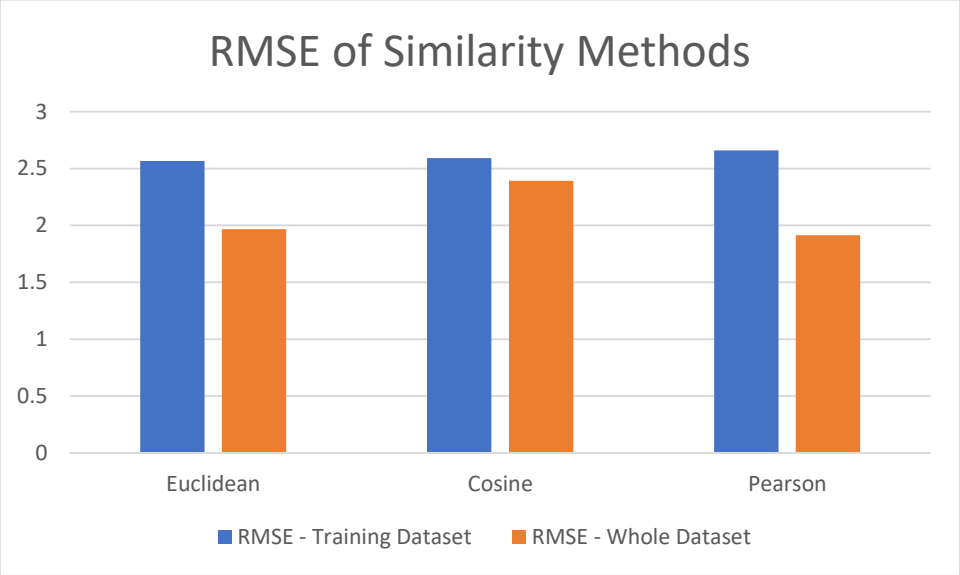


Figure 4: RMSE of similarity methods

As presented in above Figure 4, the similarity functions are provided in horizontal axis and the RMSE value is shown in vertical axis. The RMSE values are observed with both training dataset and the whole dataset. The results revealed that with training dataset, Euclidean distance metric showed highest performance as it showed least RMSE. With respect to the whole dataset the least RMSE is exhibited by Pearson metric.

Table 3: Performance comparison among different algorithms

Metric	Naïve Approach	KNN	LFM	Hybrid Approach
RMSE	2.9742	2.57622	1.11279	0.18445

As presented in Table 3, the RMSE metric is used for understanding the performance of different algorithms like Naïve approach, KNN, LFM and the proposed hybrid approach.

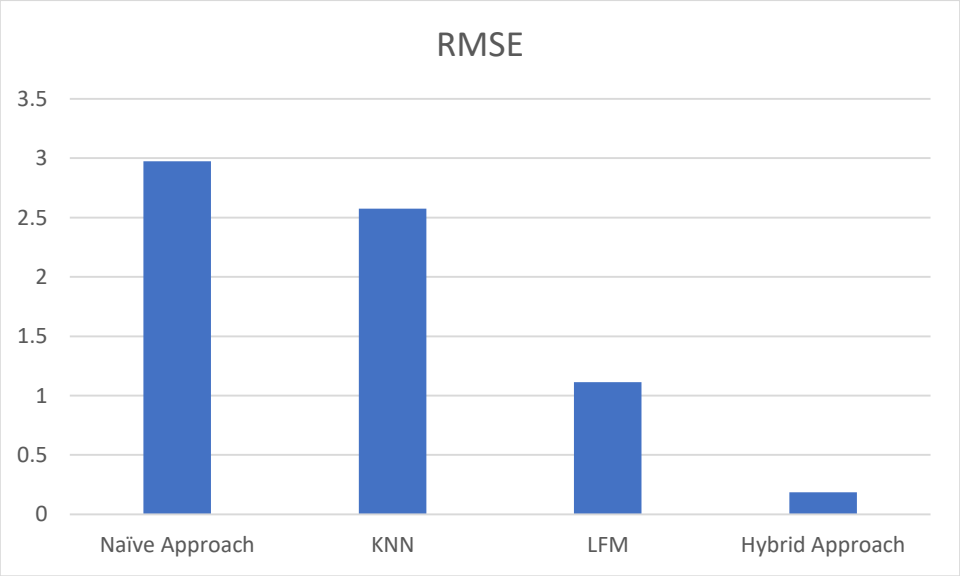


Figure 5: RMSE of different algorithms in predicting taste rating

As presented in Figure 5, the prediction algorithms are provided in horizontal axis and the RMSE value is shown in vertical axis. The RMSE values are observed for predicting taste rating. The results revealed that the hybrid approach proposed in this paper outperformed all other existing methods. It showed least RMSE indicating highest accuracy. The least performance is exhibited by naïve approach. LFM is found to be better than KNN while KNN showed better performance than naïve approach.

Table 4: Performance comparison among different recommender models

Recommender Models	Performance			
	Accuracy	**Precision**	**Recall**	**F1-score**
Naïve Approach	0.639394	0.8756	0.6368	0.737347
k-NN	0.665263	0.88555	0.66665	0.760665
LFM	0.952527	0.9552	0.9552	0.9552
Hybrid Approach	0.957547	0.96515	0.9552	0.960149

As presented in Table 4, the performance of the recommender models is provided in terms of accuracy, precision, recall and F-score.

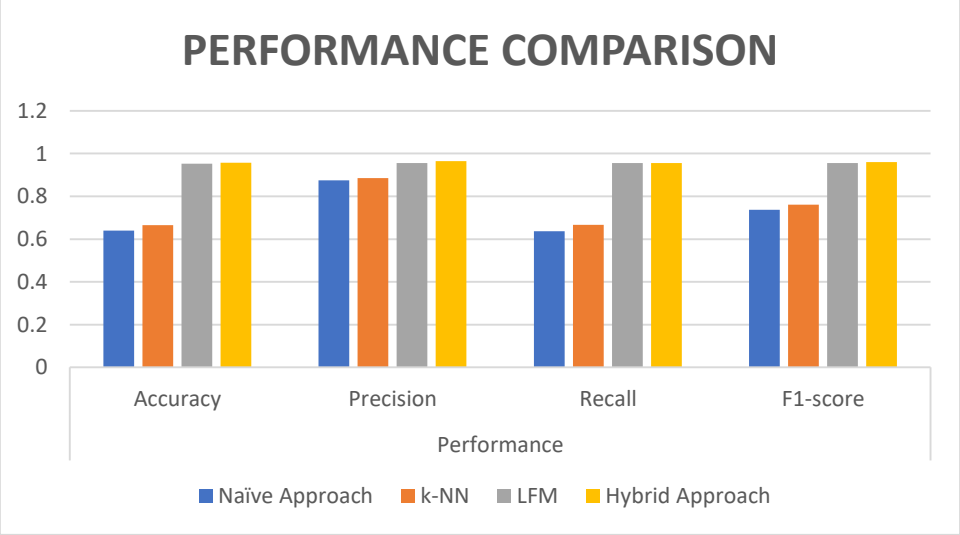


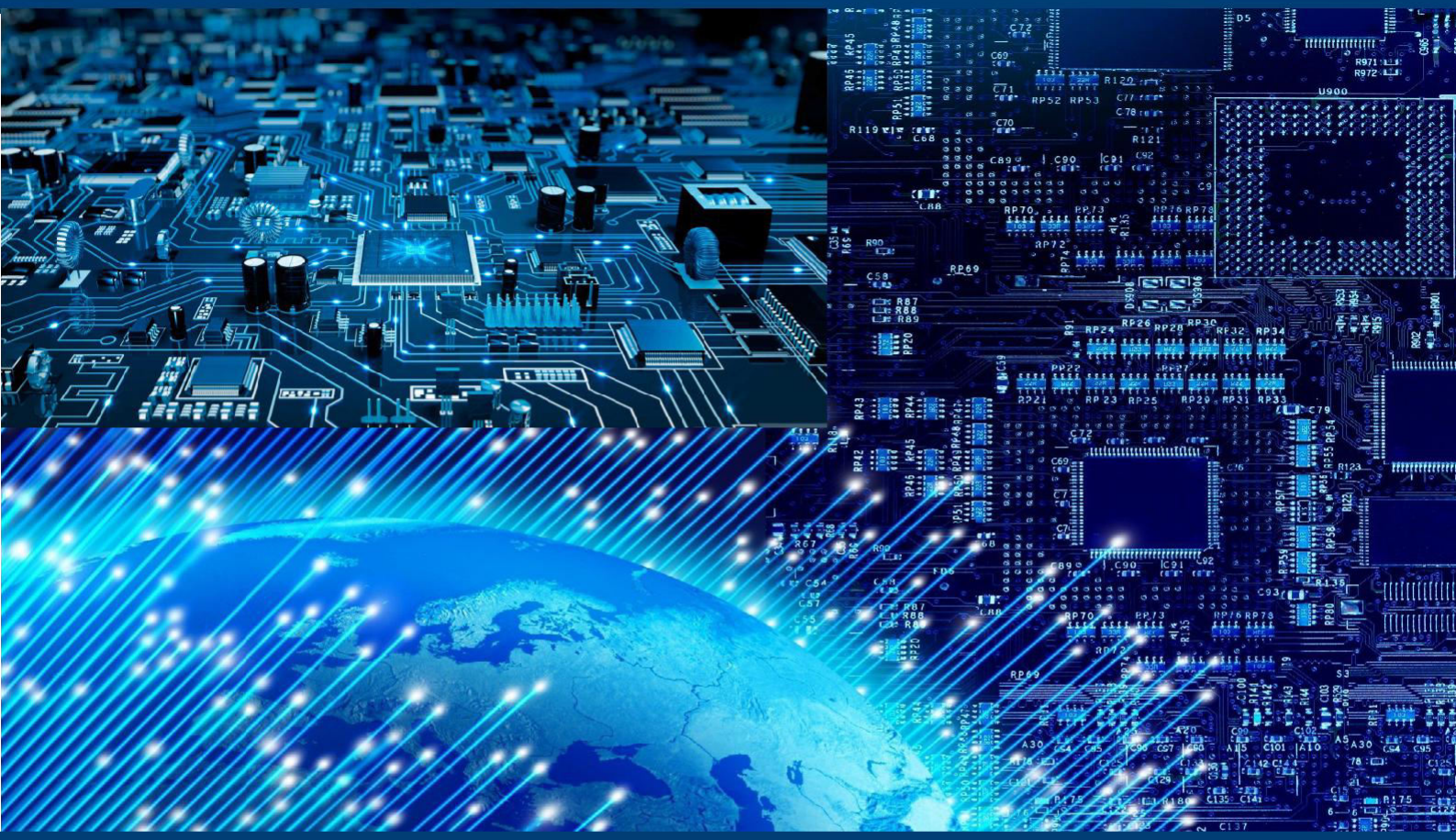
Figure 6: Performance of recommender models

IV. CONCLUSION

Many healthcare applications came into existence. However, with the big data analytics and NLP there is room for more in the industry. In this paper, we proposed a framework with hybrid approach based on ML for big data analytics and NLP for pre-processing of data for generating diet recommendations. It leverages health diet in the form of recommendations to people. In conclusion, there is much potential for ML to make progress in nutrition science. ML can capture the complex interactions that exist and are increasingly generated with modern technologies in nutrition and health data. The failure to be able to use techniques that can analyze complex data, such as ML, represents an unnecessary barrier to scientific progress. Although still relatively new, it is evident that AI approaches have much potential to supplant traditional and domain-specific methods in predictive capabilities, efficiency, costs, and convenience. ML can also be helpful in the data collection and data preprocessing stages in various fields of nutrition.

REFERENCES

1. Robert A. Sowah et al (2020) Design and Development of Diabetes Management System Using Machine Learning, Int J Telemed Appl. 2020; 2020: 8870141. Published online 2020 Jul 16. doi: 10.1155/2020/8870141.
2. Alison Gray et al (2019) Nutritional Recommendations for Individuals with Diabetes, National Library of medicine.
3. Isfafuzzaman Tasin et al (2022) Diabetes prediction using machine learning and explainable AI techniques, Healthcare Technology Letters, Volume10, Issue1-2, https://doi.org/10.1049/htl2.12039.
4. Salliah Shafi Bhat (2021) Predictions of Diabetes and Diet Recommendation System for Diabetic Patients using Machine Learning Techniques, International Conference for Emerging Technology, 978-1-7281-7029-9/21.
5. Colmenarejo G. Machine learning models to predict childhood and adolescent obesity: a review. Nutrients. 2020;12(8):2466.
6. Rigdon J, Basu S. Machine learning with sparse nutrition data to improve cardiovascular mortality risk prediction in the USA using nationally randomly sampled data. BMJ Open. 2019;9(11):e032703.
7. Abbas, A., Bilal, K., Zhang, L., & Khan, S. U. (2015). A cloud-based health insurance plan recommendation system: A user centered approach. Future Generation Computer Systems, p 99–109.
8. Shakya, Subarna, and Lalitpur Nepal. "Computational Enhancements of Wearable Healthcare Devices on Pervasive Computing System." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 2, no. 02 (2020): 98-108.
9. Aljumah, A.A., Ahamad, M.G., Siddiqui, M.K., 2013. Application of data mining: Diabetes health care in young and old patients. Journal of King Saud University - Computer and Information Sciences 25, 127–136. doi:10.1016/j.jksuci.2012.10.003.
10. Sharief, A.A., Sheta, A., 2014. Developing a Mathematical Model to Detect Diabetes Using Multigene Genetic Programming. International Journal of Advanced Research in Artificial Intelligence (IJARAI) 3, 54–59. doi:doi:10.14569/IJARAI.2014.031007.



International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |