RESEARCH ARTICLE OPEN ACCESS

VITAMIN D, B12 DEFICIENCY AND ANAEMIA DETECTION USING MACHINE LEARNING TECHNIQUES

Mrs. Dakshayini G R[1], Vijayalakshmi D[2], Prerana Kumari[3], Jeevan Reddy R[4], Nandan Kumar N[5]

Assistant- Professor, Student, Student, Student, Student

Department of Computer Science and Engineering, KSSEM, Bengaluru, Karnataka, India

[1] dakshayini@kssem.edu.in, [2]dvijayalakshmi911@gmail.com, [3]prerana99010@gmail.com, [3]jeevanani2733@gmail.com, [4]nkumarn354@gmail.com

Abstract:

Nowadays millions of people in the world suffer from vitamin deficiency which underlines health issues in people's daily lives. Health issues due to vitamin deficiency are on the rise now due to failure to acquire necessary vital minerals and nutrition because of lack of awareness of type of deficiency they may undergo without medical consultation. Identification of vitamin deficiencies helps to avoid serious causes such as anaemia, infectious disease-related deaths, and pregnancy and childbirth-related deaths. Due to advancement in technologies, there are various techniques for identifying vitamin deficiencies. This study suggests a unique machine learning algorithm-based method for identifying vitamin D and B12 deficiency. Anaemia can additionally be detected using the collected serum and MCV value in blood reports and incorporating symptoms correlated with the deficiencies, algorithm of machine learning is employed and provides a user-friendly interface. A chatbot is also included offering personalized recommendations without any medical consultation and to facilitate self assessment for individuals.

Keywords — Vitamin D deficiency, vitamin B12 Deficiency, Anaemia, Machine Learning algorithm, Blood reports, Recommendations.

I. INTRODUCTION

Vitamin deficiency is a global health concern affecting billions of individuals, with children being particularly vulnerable. According to the World Health Organization (WHO), one in three children suffers from a lack of essential vitamins. This deficiency manifests in various ways, such as night blindness and weakened immunity, which may have detrimental effects on overall health and well-being. Beyond vitamins, deficiencies in minerals like zinc, iron, and iodine also contribute to health issues across every age range. Pregnant women and children are especially susceptible to these deficiencies due to increased nutritional demands and insufficient intake. Recognizing the significance of addressing these deficiencies, recent studies have turned to innovative approaches like machine learning algorithms. Machine learning, a subset of artificial intelligence, leverages data and algorithms to mimic human learning processes and enhance accuracy over time

The ongoing research is on the basis of detecting Vitamin B12, D deficiencies and anaemia using ML algorithm. Machine learning focuses on the usage of data and algorithms to replicate the way that humans learn, and improves its accuracy. Using the collected blood reports and the MCV value, the symptoms are incorporated associated with the deficiencies.

II. LITERATURE SURVEY

Numerous recent studies have demonstrated the potential of machine learning (ML) techniques in predicting and detecting various vitamin and mineral deficiencies, furthermore related health conditions. One study conducted in 2020 focused on psychiatric patients experiencing intense episodes and found correlations between these episodes and deficiencies in Vitamin B1, B12, and folate. By employing ML models including Random Forest and Support Vector Machine, researchers were able to predict these deficiencies using blood test results and patient data. Additionally, correlations were found between deficiencies and biomarkers like alkaline phosphatase (ALP) and thyroid-stimulating hormone (TSH), suggesting the possibility of ML algorithms to facilitate faster therapy decisions in such cases.

In a separate study conducted in 2023, researchers at NEU Hospital investigated applying ML models for classifying metabolic syndrome (MtS) status based on data from 481 participants. Among the various ML models tested, Random Forest (RF) emerged as the preferred choice because of its robustness and resilience to data variability and multicollinearity. RF, along with Elastic-net Ordinal Regression (ENOR), demonstrated superior performance, particularly in smaller training sample sizes, exhibiting high accuracy, specificity, sensitivity, precision, F1-score, and Cohen's kappa.

Lastly, a 2020 study aimed to forecast a lack of Vitamin D deficiency (VDD) among university students utilizing diverse ML models. Random Forest emerged as the top-performing model, achieving a remarkable 96% accuracy rate. This predictive performance was assessed using metrics such as Precision, F1-score, and McNemar's test, demonstrating the model's effectiveness regardless of demographic and health data. These studies collectively highlight the potential of ML techniques in enhancing the detection and prediction of nutritional deficiencies and related health conditions, thereby facilitating more timely interventions and personalized healthcare strategies.

III. SYSTEM ARCHITECTURE

The system architecture has the following:

User Interface: This is the front-end of the application, where users interact with the system by inputting the mcv value, serum value, symptoms.

Database: This is the storage component, where user data and diagnosis results are stored.

Vitamin D Deficiency Diagnosis: Using the user input that is based on serum value, vitamin D is detected.

Machine Learning Model: Random forest algorithm is employed to detect vitamin B12 or anaemic patients using the user inputs mcv value and other symptoms.

R non dec

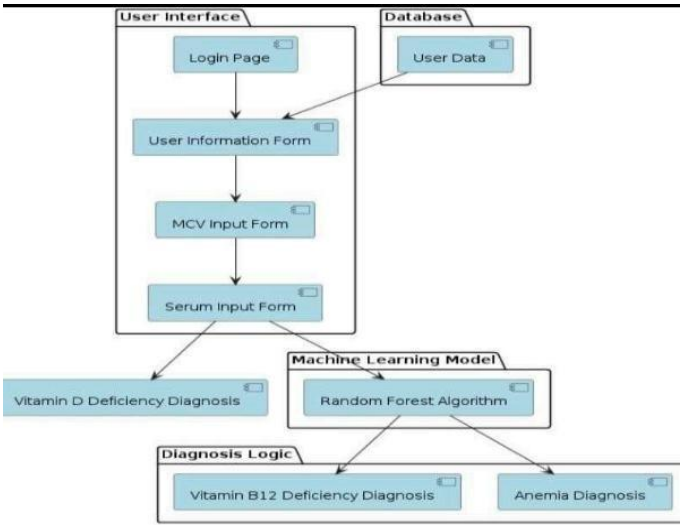

Fig 1: System Architecture

IV. PROPOSED SYSTEM

The proposed system seeks to develop a productive and user-friendly platform that utilizes modern technology to streamline the detection process. At the core of our proposed system lies the utilization of random forest classifier, for classification tasks related to B12 deficiency and anaemia. This algorithm is chosen for its ability to handle complex data and provide robust predictions making it suitable for the diverse and multidimensional nature of healthcare data.

A. Data Collection and Feature Extraction:

Visit hospitals and online sources to gather datasets including blood reports, MCV values, and symptoms. Ensure comprehensive data collection to improve the caliber of model's effectiveness and robustness. Identify relevant features such as MCV values and symptoms that contribute to the classification task. Analyze the dataset to determine which features have the greatest influence on deficiency detection.

B. Model Deployment:

Utilizing machine learning algorithm, in this particular case random forest is chosen as it can handle huge datasets with high dimensionality. The collection of data is separated into sets for training and testing. Hyperparameter tuning and cross-validation techniques may be employed to maximize the performance of the model.

C. Deployment:

Deploy the trained model in a engaging platform accessible to individuals. Users input their data, which includes name, age, MCV values, and symptoms through the frontend. The platform processes the input through the trained model classify individuals into deficient or non-deficient categories. Provide personalized feedback on deficiency status and offer recommendations for treatment and preventive measures.

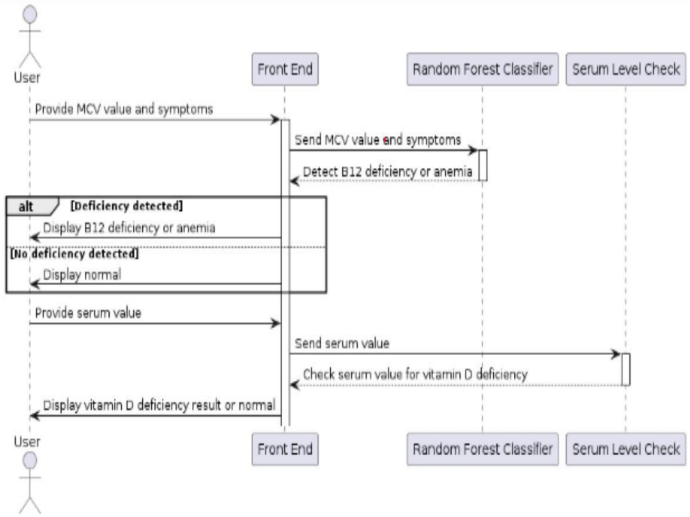

Fig 2: Sequence diagram of the model

V. RESULTS AND SNAPSHOTS

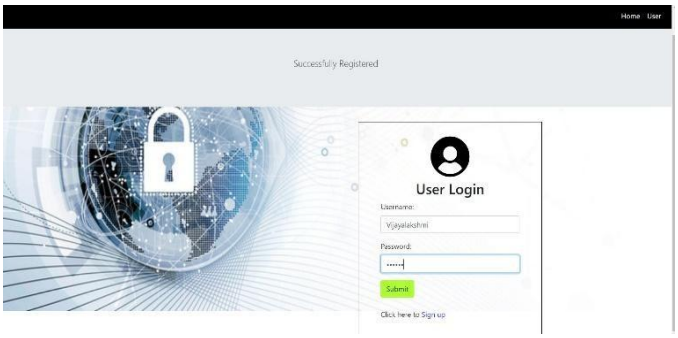

Fig 3: Login page

Fig 4: B12 and anaemia detection page

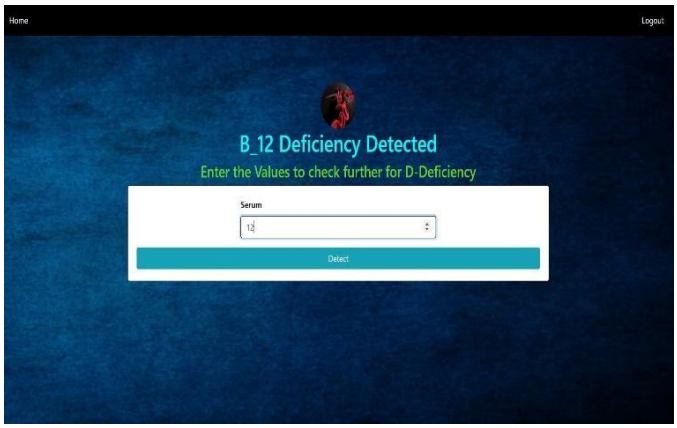

Fig 5: D detection page

VI. CONCLUSION

To sum up, our project successfully implemented an innovative method based on machine learning methods to recognize vitamin deficiency. Data collection efforts yielded diverse datasets from hospitals and online sources, containing essential information such as MCV values, symptoms, and serum levels Leveraging random forest algorithms and incorporating relevant characteristics seen in blood reports and symptoms, the developed models demonstrated high accuracy in classifying individuals into vitamin B12 deficient or anaemic categories. Additionally, people who are vitamin D deficient were identified based on serum values, allowing for early detection and intervention. The user-friendly interface facilitates easy interpretation and furthermore applying the model, enabling healthcare professionals to make informed decisions regarding deficiency detection and intervention.

ACKNOWLEDGEMENT

We express our gratitude to Mrs. Dakshayini G R, Assistant Professor, for her invaluable assistance and encouragement during the project's progress. Her expertise, encouragement, and astute critique have been tremendously helpful to us in refining our approach and streamlining our thought process. Her unwavering dedication and desire are deeply appreciated, as they have contributed significantly to the success of our project.

REFERENCES

[1] H. Tamune, J. Ukita, Y. Hamamoto, H. Tanaka, K. Narushima, and N. Yamamoto, ‘‘Efficient prediction of vitamin B deficiencies via machine learning using routine blood test results in patients with intense psychiatric episode,’’ Frontiers Psychiatry, vol. 10 (2020).

[2] Sambasivam G, Amudhavel J, Sathya G. A predictive performance analysis of vitamin D deficiency severity using machine learning methods. IEEE Access (2020).

[3] T. Merlijn, K. M. A. Swart, P. Lips, M. W. Heymans, E. Sohl, N. M. Van Schoor, C. J. Netelenbos, and P. J. M. Elders, ‘‘Prediction of insufficient serum vitamin D status in older women: A validated model,’’Osteoporosis Int., vol. 29 (2018).

[4] Shuyu Guol, Robyn M. Lucas, Anne-Louise Ponsonby, A Novel Approach for prediction of Vitamin D Status using Support Vector Regression (2019).

[5] Nuriye Sancar, and Sahar S. Tabrizi, “Machine Learning approach for the detection of vitamin D level: a comparative study, BMCMIDM (2023).

[6] Mutammimul Ula, Mursyidah, Yana Hendriana, and Richki Hardi, “An expert system for early diagnose of vitamins and minerals deficiency on the body, IEEE (2019).