Identifying Disease and Diagnosis in Females Using Machine Learning

Sabyasachi Pramanik
https://orcid.org/0000-0002-9431-8751
Haldia Institute of Technology, India

Samir Kumar Bandyopadhyay
Bhawanipore Educational Society, India

INTRODUCTION

When a human body doesn't produce insulin or doesn't produce sufficient insulin or doesn't use it efficiently it can lead to dangerous complications called Diabetes (Lee et. al. 2021). This disorder occurs when the glucose of our blood gets high, also known as blood sugar. In the human body the main source of energy is blood glucose and they are getting energy from whatever we eat. For having energy, this glucose gets into the body cells to be used with the help of insulin. Possessing extra visceral fat is significantly connected to having the highest risk of developing diabetes, such as type 2 diabetes. Obese women are also more likely than males to have a healthy metabolism. The present epidemic of diabetes in India is mostly due to changes in lifestyle. Increasing prevalence may be attributed to a variety of factors, including fast changes in food habits, lack of physical activity, and increase in body weight, particularly the build-up of belly fat.

Types of Diabetes

1. Type 1 Diabetes (De Bois, M., et al. 2022): This is also known as autoimmune disorder. Here the diabetes pancreatic cells are damaged and for that the pancreas fails to produce enough body insulin.
2. Type 2 Diabetes (Yang, L., et al. 2021): It is also known as Adult-Onset Diabetes. Here the pancreas either produces excessive insulin or it resists insulin and it affects the way body processes blood sugar.
3. Pre-Diabetes (Harimoorthy, K., et al. 2021): Pre-Diabetes is nearly same as Type 2 Diabetes. Here, in the Pre-Diabetes stage the amount of blood sugar is not greater than Type 2 Diabetes.
4. Gestational Diabetes Mellitus (GDM) (Barik, S., et al. 2021): This type of diabetes consists of carbohydrate of varying intensity during pregnancy. GDM has no specific clinical features, it is diagnosed after screening.

Symptoms of Diabetes

1. Excessive thirst
2. Slow healing sores and recurrent infection
3. Feeling lazy
4. Blurred vision

DOI: 10.4018/978-1-7998-9220-5.ch187

5. Tingling in hand and feet
6. Swollen gums
7. Excessive urination
8. Weight lost

Causes of Diabetes

1. Diabetes due to obesity
2. Hereditary
3. High sugar levels during pregnancy
4. Blood vessel diseases
5. High blood pressure & high cholesterol
6. Pre-diabetes or impaired fasting glucose

BACKGROUND

The research of (Karatsiolis and Schizas, 2012) implemented a modified Support Vector Machine technique is used to solve the challenge of diagnosing Pima Indian Diabetes using data from UCI Repository of Machine Learning Databases (Akter, L., et al. 2021). The suggested algorithm's performance is compared to that of other categorization algorithms in order to show that it outperforms them. The purpose of this study is to provide readers an understanding of a technique that may be utilised to improve classification success rates attained by using traditional methodologies like Neural Networks, RBF networks (Meng, X., et al. 2021) and K-nearest neighbours. The proposed technique separates the training set into subgroups: one resulting from the combining of coherent data areas, and the other containing data that is difficult to cluster. As a result, the first subset is being used to train an RBF kernel Support Vector Machine, whereas the second subset is utilised to train a polynomial kernel Support Vector Machine. The method can determine whichever of the two Support Vector Machine algorithms to utilise during classification. The proposed approach is based on the assumption that the RBF Support Vector Machine architecture is more suited to use on data sets with varying features than the polynomial kernel (Mahindru, A., et al., 2021). The recommended technique improved the average classification success rate to 82.2 percent in this investigation, whereas the greatest performance attained in earlier experiments was 81 percent using a fine-tuned very complicated ARTMAP-IC model (Bascil, M. S., et al. 2012).

The outcome of the study by (Kaur et al., 2018) used the diabetes in a serious metabolic illness that may have a negative impact on the overall body system. Diabetic nephropathy, myocardial stroke, and other complications may all be exacerbated by undiagnosed diabetes. This illness affects millions of individuals all over the globe. Diabetes must be detected early in order to live a healthy life. Diabetes is a worldwide problem, since the number of patients is continually increasing. Machine learning (ML) (Pramanik, S. et al., 2021) is a computer technology that enhances performance by automatically learning from experience and making more accurate predictions. In this study, the authors used the R data (Li, W. et al., 2021) manipulation programme to generate trends and uncover patterns in the Pima Indian diabetes dataset using machine learning techniques. Using the R data manipulation tool, we constructed and assessed five alternative prediction models to categorise patients as diabetes or non-diabetic. We employed supervised machine learning methods such as the linear kernel support vector machine (SVM-linear), the radial basis function (RBF) kernel support vector machine, the k-nearest neighbour (K-NN),