Chapter 13
Applications of Machine Learning Models With Medical Images and Omics Technologies in Diabetes Detection

Chakresh Kumar Jain
https://orcid.org/0000-0002-9226-7719
Jaypee Institute of Information Technology, India

Aishani Kulshreshtha
Jaypee Institute of Information Technology, India

Avinav Agarwal
Jaypee Institute of Information Technology, India

Harshita Saxena
Jaypee Institute of Information Technology, India

Pankaj Kumar Tripathi
https://orcid.org/0000-0002-9929-359X
Jaypee Institute of Information Technology, India

Prashant Kaushik
Jaypee Institute of Information Technology, India

ABSTRACT

Diabetes mellitus is a long-term condition characterized by hyperglycaemia resulting in the emergence of a variety of health problems, such as diabetic retinopathy, kidney failure, dental problems, heart disease, nerve damage, etc.; and is governed by several factors, i.e. biological, genetics, food habits, sedentary lifestyle choices, poor diets and environments, etc. According to the recent morbidity figures, the global diabetic patient population is anticipated to reach 642 million by 2040, implying that one out of every ten people will be diabetic. The data generation and AI based methods—i.e., SVM, kNN, decision tree, Baysian method in medical health –have facilitated the effective prediction and classification of voluminous size of biological data of different types of BMI, skin thickness, glucose, age, tongue and retinal images apart from Omics data, for early diagnostics. The chapter summarizes the basic methods and applications of machine learning and soft computing techniques for diabetes diagnosis and prediction with limitations of integrative approaches.

DOI: 10.4018/979-8-3693-3026-5.ch013

INTRODUCTION

Healthcare is one of the most major concerns for the world in this day and age. Health is a measure of an individual's well-being. A healthy lifestyle is one of the crucial goals of any person, and access to proper healthcare forms an essential aspect of every individual's life. Due to increasing sedentary lifestyle choices and poor diets, people are getting diagnosed with Diabetes at an alarming rate. A whopping figure of 422 million people, according to the World Health Organisation (WHO) have been diagnosed with diabetes, which adds up to about 1.5 million deaths annually. It is estimated that this number will increase rapidly in the coming years, affecting primarily India, China and the USA (World Health Organization, 2022).

Diabetes is an incurable condition wherein the glucose level in the blood is more elevated than usual. One of the reasons for this can be inadequate secretion of insulin by the body, or it's improper response to it (Thakkar et al., 2021). Diabetes can both be due to inadequate lifestyle as well as certain genetic factors. There are primarily two types of diabetes; Type 1 and Type 2 diabetes as depicted in figure I. Typically type 1 diabetes patients are younger and have to be treated with insulin therapy. Due to this, this type is commonly known as insulin-dependent diabetes. Comparatively type 2 diabetes patients are usually middle-aged and above, and can be effectively treated with oral medication. This type is often associated with other complications such as obesity and hypertension amongst others (Robertson et al., 2011). Apart from these two, another type of diabetes known as gestational diabetes primarily affects pregnant women. This type of diabetes is resolved after birth, but results in greater chances of patients developing type 2 diabetes later on in life. Apart from this, Diabetes can lead to more dangerous medical conditions such as eye, kidney and dental problems, heart disease, and nerve damage (National Institute of Diabetes and Digestive and Kidney Diseases, 2016). Such a disease therefore demands efficient treatment and control, especially since a significant portion of patients still remain undiagnosed (Kaul et al., 2012).

Traditionally, diabetes is diagnosed on the basis of laboratory test reports and the assessment of the doctor. This constitutes clinical data. While it is still a common and effective diagnostic technique, it is time-consuming and bias prone. The diagnosis is entirely dependent upon the experience of the medical practitioner, and quality of the laboratory tests, which may fluctuate (Choubey et al., 2018). Clearly, there is a need for early detection of this disease for effective treatment. With the advent of soft computing techniques, the diagnosis and prediction of diabetes has become less time-consuming. Soft computing enables the incorporation of methods that aim to quickly design answers to challenging real-world problems. A variety of strategies can be combined to create soft computing techniques, which are used to solve complicated real-world problems (ex. in management, agriculture, economics, and other fields) that were otherwise insurmountable (Agarwal & Mehta, 2014).

This chapter aims to provide a comprehensive review on the current updates in various advanced soft computing techniques that have been employed for diabetes classification and prognosis. Various techniques such as logistic regression, naïve bayes, and decision tree have been discussed.

RELEVANCE OF BIOLOGICAL DATA

Before discussing various integrative biology and soft computing techniques used for diabetes detection and classification, it is important to put emphasis on the type and quality of biological data used for these studies. Despite being so prevalent, unfortunately only a minimal number of datasets are available and utilised for prediction of diabetes, and amongst those only one or two datasets are available for public