

# Towards Analyzing the Prediction of Developing Cardiovascular Disease using Implementation of Machine Learning Techniques

G. Angayarkanni<sup>1</sup>, Dr. S. Hemalatha<sup>2</sup>

<sup>1</sup>PhD Scholar, <sup>2</sup>Associate Professor, Nallamuthu Gounder Mahalingam College, Pollachi, Tamilnadu

**Abstract:** Machine Learning is an emerging technique widely penetrated into the prediction of Cardiovascular Disease (CVD) events. Cardiovascular Disease is the leading global cause of death among the diabetes and non-diabetes patients. The main aim of this analysis shows the detailed literature survey of how the different machine learning algorithm is used to predict the CVD events among the type 2 diabetes and non-diabetes patients based on the major risk factors such as age, gender, hypertension, cholesterol, diabetes, smoke, alcohol intake, physical inactivity etc. and the accuracy, sensitivity, specificity of the algorithms is also compared. The early prediction have been supports the physician for medical decision making for the management of CVD for the patients and recommends for the patients to take treatment before it becomes fatal.

**Keywords:** Cardiovascular Disease, Machine Learning, Diabetes

## I. INTRODUCTION

Diabetes mellitus is a chronic disease that occurs when the pancreas is no longer able to produce insulin or when the body is unable to make use of the insulin produced. The survey done by Indian Council of Medical Research (ICMR) thirty years ago, the prevalence of diabetes mellitus was around 2% in urban and 1% in rural India. Within a span of three decades the prevalence rates have raised to 12 – 16% in Urban and 3-8% in Rural India [1]. Moreover 40 million of population in India has now been diagnosed with diabetes. This shows that India has the highest number of diabetics than any other country in the entire world. The complication of Diabetes is divided into Microvascular Complications such as retinopathy (eye damage), nephropathy (kidney damage), neuropathy (nerve damage) and Macrovascular Complications such as cerebrovascular disease (risk of stroke), cardiovascular disease (risk of heart attack), and peripheral vascular disease (poor circulation to the limbs). Individuals with diabetes most often die due to cardiovascular disease (CVD) rather than the cause uniquely related to diabetes such as hypoglycemia [2]. Cardiovascular disease is the term of all kinds of diseases that affect the heart or blood vessels, together with coronary heart disease. People with diabetes are two to four times more likely to die from heart disease than people without diabetes. American Heart Association considers diabetes as the one of seven major controllable risk factors for CVD. According to world health organization CVD account for one in every three deaths worldwide. It is very important to make the people aware about the risk of CVD.

The Health Care Industry generates large amount of data, therefore there is a need of effective technique like Machine Learning to handle effectively. Machine Learning is an application of Artificial Intelligence and is a category of algorithm that will predict the outcomes accurately. Applying Machine Learning in Cardiovascular Disease research is a best approach for diagnosis, prediction, management and other clinical related administration aspects. Machine learning can help people make a preliminary decision about CVD according to their daily physical examination data and it can be act as a reference for physicians. Maximum in the CVD finding research 14 features are used as input in order to predict CVD and the attributes are classified into continuous and categorical. CVD risks are categorized as low, medium and high. All high risk people are advised to change their diet and other life style factors and probably take medications to control the elevated features. Algorithm works by detecting some pattern in the available input data and build a model using the input data to make prediction for the new data. Regression and Classification algorithms in Machine Learning are mostly used to predict a value from the given features and most common regression algorithms such as linear regression, logistic regression, decision tree, random forest, SVM and gradient boosting are used.

The following literature review was done using 26 published papers in order to predict the CVD using machine learning algorithm in the period from 2012 to 2020.