Abstract

The International Diabetes Foundation estimates 72.9 million Indians to be currently suffering from diabetes, with this number set to increase to 134.3 million by 2045. This makes India the country with the second highest number of adults living with diabetes. But, mean healthcare expenditure on diabetes per person in 2017 was only ID 426, far behind countries other countries.

The International Diabetes Foundation's 2045 conservative projections, assuming mean per capita expenditure and diabetes prevalence rate remain constant, estimate global cost of Diabetes to increase to USD 776 Billion, which represents a 7% growth.

Insurance exists to protect oneself against increasing and unforseen costs. Existing health insurance plans were unable to appropriately cover expenses of diabetes.

Only recently, have specific insurance plans for Diabetes sprung up but all seem to use age as a proxy to classify patients into premium bands and then offer adjustments based on medical state.

It is believed that doing so is convenient, but a more equitable solution exists which would not only help patients by appropriately identifying their costs, but would also help insurance companies make health classes in their diabetes insurance policy using medical indicators as well as age leading to a more accurate estimate of costs.

Using an existing dataset, the medical indicators to keep in check for non-diabetics are identified by implementing various Machine Learning Classification Algorithms. Data collected from Diabetic patients at a hospital is used to model the relationship between Annual Spending and medical indicators using Generalized Linear Models and Generalized Additive Models. Machine Learning Clustering Algorithms alongwith Decision Trees are used to split the data according to medical indicators and create homogeneous classes based on annual spending.