

Executive Summary Report

“Diabetes Risk Factors Analysis (BRFSS 2015)”

Author: Md Arman Kaishar

Date: September 11, 2024

Introduction:

Diabetes is a widespread metabolic disorder affecting millions globally, resulting in substantial economic and health burdens. The condition is characterized by the body's inability to regulate blood sugar, often due to inadequate insulin production or usage. Chronic high blood sugar can lead to complications like heart disease, stroke, and kidney failure. This report leverages data analytics to identify key risk factors predictive of diabetes, particularly focusing on type II diabetes, which is strongly influenced by lifestyle factors.

Objective:

The primary objective of this analysis is to explore and identify key factors that contribute to the risk of diabetes using a balanced dataset. By examining these risk factors, the study aims to provide insights that can aid in early diagnosis, prevention strategies and informed public health interventions.

Dataset Overview:

The dataset used for this analysis is a clean sample of 70,692 responses from the 2015 Behavioral Risk Factor Surveillance System (BRFSS), with a 50-50 split between individuals with and without diabetes (or prediabetes). The dataset includes 21 feature variables related to health indicators, such as body mass index (BMI), physical activity, blood pressure, cholesterol, and other lifestyle factors. The target variable is `Diabetes_binary`, where 0 indicates no diabetes, and 1 indicates prediabetes or diabetes.

Key features in the dataset include:

- Health Conditions: High blood pressure, cholesterol, stroke, heart disease
- Lifestyle Factors: Smoking, physical activity, fruit and vegetable consumption, alcohol use
- Socioeconomic Indicators: Income, education level
- General Health Measures: BMI, mental and physical health, general health status

Methodology:

Python code in Google Colab (Univariate Analysis, Correlation Analysis, Logistic Regression, Decision Tree Analysis)

Insights & Findings:

1. Univariate Analysis:

- High Blood Pressure: Individuals with diabetes are significantly more likely to have high blood pressure compared to non-diabetics.
- Cholesterol: A larger proportion of those with diabetes have high cholesterol levels.
- BMI: Obesity (higher BMI) is strongly associated with the presence of diabetes, indicating a direct relationship between body weight and diabetes risk.

- **Physical Activity:** Regular physical activity is less common in diabetic individuals, reinforcing its protective role against diabetes.
- **Fruit and Vegetable Consumption:** Those who consume fruits and vegetables regularly are less likely to have diabetes.
- **Socioeconomic Factors:** Lower income and education levels are associated with higher diabetes prevalence, reflecting the impact of social determinants on health.

2. Correlation Analysis:

- BMI and General Health are highly correlated with the target variable, reinforcing that obesity and poor self-reported health strongly predict diabetes risk.
- High Blood Pressure and High Cholesterol also show moderate correlations with diabetes.

3. Logistic Regression:

- BMI, General Health, and Physical Health emerged as the most important predictors of diabetes.
- High Cholesterol and High Blood Pressure also significantly impact diabetes likelihood.

4. Decision Tree Analysis:

- The decision tree highlighted BMI and General Health as the most critical factors influencing the prediction of diabetes.
- Lifestyle factors such as physical activity and fruit consumption played secondary but still notable roles in diabetes prediction.

Recommendations:

1. **Promote Healthy Lifestyle Choices:** Initiatives should emphasize regular physical activity, balanced diets with fruit and vegetable consumption, and maintaining a healthy BMI to reduce diabetes risk.
2. **Screening for Key Health Conditions:** Regular monitoring of blood pressure, cholesterol, and BMI, especially among high-risk populations, can aid in early diagnosis and effective management of diabetes.
3. **Targeted Public Health Interventions:** Public health efforts should focus on lower-income and less-educated populations, who are disproportionately affected by diabetes. Providing affordable healthcare access and education on lifestyle changes is crucial.
4. **Community Programs:** Implement community-based programs aimed at increasing awareness of diabetes risk factors and the benefits of early screening and lifestyle changes.

Conclusion:

This analysis confirms the importance of several health and lifestyle factors, particularly BMI, physical activity, and general health, in predicting diabetes risk. Addressing these factors through public health initiatives and promoting healthier lifestyles can significantly reduce the burden of diabetes and its associated complications.