Solution Architecture:
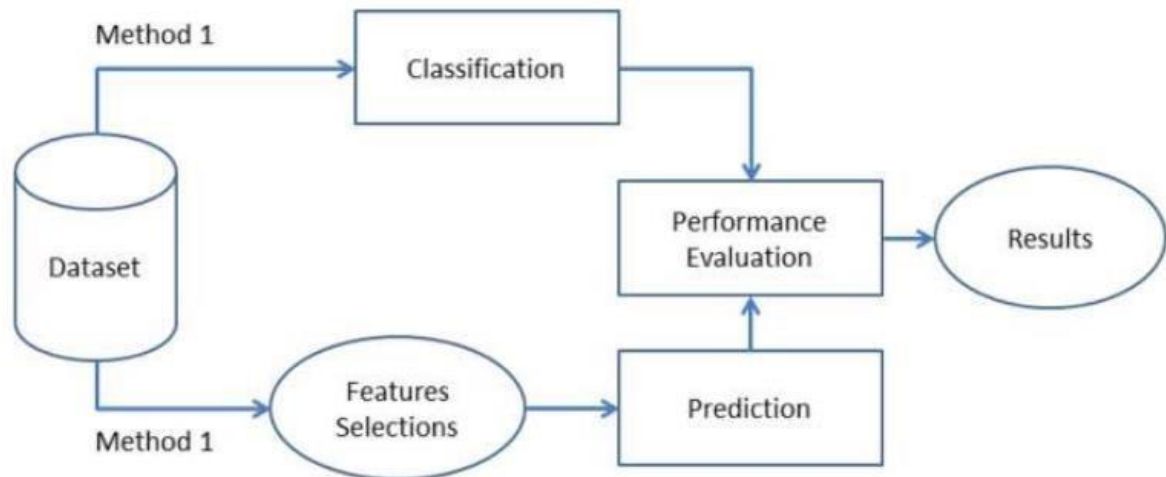
The leading cause of death in the developed world is heart disease. Therefore, there needs to be work done to help prevent the risks of having a heart attack or stroke. This dataset to predict which patients are most likely to suffer from a heart disease in the near future using the features given. Heart disease is one of the biggest causes of morbidity and mortality among the population of the world. Prediction of cardiovascular disease is regarded as one of the most important subjects in the section of clinical data analysis. The amount of data in the healthcare industry is huge. Data mining turns the large collection of raw healthcare data into information that can help to make informed decisions and predictions. The dataset consists of 303 individuals' data. There are 14 columns in the dataset, which are described below.

IMPORT DATASET

After downloading the dataset from Kaggle, we saved it to my working directory with the name dataset.csv. Next, we used read_csv () to read the dataset and save it to the dataset variable. Before any analysis, I just wanted to take a look at the data. So, we used the info () method. There are a total of 13 features and 1 target variable. Also, there are no missing values so we don't need to take care of any null values. Next, we used to describe () method.

Parameters:

1.Age

2.Sex

3.Chest pain type

4.BP

5.Cholesterol

6.FBS over 120

7.EKG results

8.Max HR

9.Exercise angina

10.ST depression

11.Slope of ST

12.Number of vessels fluro

13.Thallium

14.Heart Disease

Age: Age is the most important risk factor in developing cardiovascular or heart diseases, with approximately a tripling of risk with each decade of life. Coronary fatty streaks can begin to form in adolescence. It is estimated that 82 percent of people who die of coronary heart disease are 65 and older. Simultaneously, the risk of stroke doubles every decade after age 55.

Sex: Men are at greater risk of heart disease than pre-menopausal women. Once past menopause, it has been argued that a woman's risk is similar to a man's although more recent data from the WHO and UN disputes this. If a female has diabetes, she is more likely to develop heart disease than a male with diabetes.

Chest Pain: Angina is chest pain or discomfort caused when your heart muscle doesn't get enough oxygen-rich blood. It may feel like pressure or squeezing in your chest. The discomfort also can occur in your shoulders, arms, neck, jaw, or back. Angina pain may even feel like indigestion.

Blood Pressure: Over time, high blood pressure can damage arteries that feed your heart. High blood pressure that occurs with other conditions, such as obesity, high cholesterol or diabetes, increases your risk even more.

Cholesterol: A high level of low-density lipoprotein (LDL) cholesterol (the "bad" cholesterol) is most likely to narrow arteries. A high level of triglycerides, a type of blood fat related to your diet, also ups your risk of a heart attack. However, a high level of high-density lipoprotein (HDL) cholesterol (the "good" cholesterol) lowers your risk of a heart attack. **Fasting Blood Sugar:** Not producing enough of a hormone secreted by your pancreas (insulin) or not responding to insulin properly causes your body's blood sugar levels to rise, increasing your risk of a heart attack.

Resting ECG: For people at low risk of cardiovascular disease, the USPSTF concludes with moderate certainty that the potential harms of screening with resting or exercise ECG equal or exceed the potential benefits. For people at intermediate to high risk, current evidence is insufficient to assess the balance of benefits and harms of screening.

Max heart rate achieved: The increase in cardiovascular risk, associated with the acceleration of heart rate, was comparable to the increase in risk observed with high blood pressure. It has been shown that an increase in heart rate by 10 beats per minute was associated with an increase in the risk of cardiac death by at least 20%, and this increase in the risk is similar to the one observed with an increase in systolic blood pressure by 10 mm Hg.

Exercise induced angina: The pain or discomfort associated with angina usually feels tight, gripping or squeezing, and can vary from mild to severe. Angina is usually felt in the centre of your chest but may spread to either or both of your shoulders, or your back, neck, jaw or arm. It can even be felt in your hands. Types of Angina a. Stable Angina / Angina Pectoris b. Unstable Angina c. Variant (Prinz metal) Angina d. Microvascular Angina.

Peak exercise ST segment: A treadmill ECG stress test is considered abnormal when there is a horizontal or down-sloping ST-segment depression ≥ 1 mm at 60–80 ms after the J point. Exercise ECGs with up-sloping ST-segment depressions are typically reported as an 'equivocal' test. In general, the occurrence of horizontal or down-sloping ST-segment depression at a lower workload (calculated in METs) or heart rate indicates a worse prognosis and higher likelihood of multi-vessel disease. The duration of ST-segment depression is also important, as prolonged recovery after peak stress is consistent with a positive treadmill ECG stress test. Another finding that is highly indicative of significant CAD is the occurrence of ST segment elevation > 1 mm (often suggesting transmural ischemia); these patients are frequently referred urgently for coronary angiography.

Patients from age 29 to 79 have been selected in this dataset. Male patients are denoted by a gender value 1 and female patients are denoted by gender value 0. Four types of chest pain can be considered as indicative of heart disease. Type 1 angina is caused by reduced blood flow to the heart muscles because of narrowed coronary arteries. Type 1 Angina is a chest pain that occurs during mental or emotional stress. Non-angina chest pain may be caused due to various reasons and may not often be due to actual heart disease. The fourth type, Asymptomatic, may not be a symptom of heart disease. The next attribute trestbps is the reading of the resting blood pressure. Chol is the cholesterol level. Fbs is the fasting blood sugar level; the value is assigned as 1 if the fasting blood sugar is below 120 mg/dl and 0 if it is above. Restecg is the resting electrocardiographic result, thalach is the maximum heart rate, exang is the exercise induced angina which is recorded as 1 if there is pain and 0 if there is no pain, oldpeak is the ST depression induced by exercise, slope is the slope of the peak exercise ST segment, ca is the number of major vessels colored by fluoroscopy, thal is the duration of the exercise test in minutes, and num is the class attribute. The class attribute has a value of 0 for normal and 1 for patients diagnosed with heart disease.