**Pima Indians Diabetes Project**

Metadata

|  |  |
| --- | --- |
| Name | Pima Indians Diabetes |
| Data types | Multivariate |
| Data task | Classification |
| Attribute types | Integer, Real |
| Instances | 768 |
| Attributes | 8 |
| Year | 1990 |
| Area | Life |
| Description | From National Institute of Diabetes and Digestive and Kidney Diseases; Includes cost data (donated by Peter Turney) |

This dataset describes the medical records for Pima Indians and whether or not each patient will have an onset of diabetes within the years.

Features:

* Pregnancies = Number of times pregnant
* Glucose = Plasma glucose concentration 2 hours in an oral glucose tolerance test
* Blood Pressure = Diastolic blood pressure (mm Hg)
* Skin Thickness = Triceps skin fold thickness (mm)
* Insulin = 2-Hour serum insulin (mu U/ml)
* BMI = Body mass index (weight in kg/(height in m)^2)
* Diabetes Pedigree Function = Diabetes pedigree function
* Age = Age (years)

Label:

* Class (1: tested positive for diabetes, 0: tested negative for diabetes)

**Binary Classification Project**

Requirement: **Classification Accuracy Score > 90%**