

Assignment 02

Classification

The diabetes dataset is given. The objective is to predict based on diagnostic measurements whether a patient has diabetes.

Dataset description

- Pregnancies: To express the Number of pregnancies
- Glucose: To express the Glucose level in blood
- BloodPressure: To express the Blood pressure measurement
- SkinThickness: To express the thickness of the skin
- Insulin: To express the Insulin level in blood
- BMI: To express the Body mass index
- DiabetesPedigreeFunction: To express the Diabetes percentage
- Age: To express the age
- Outcome: To express the final result 1 is YES o is NO

Note: The zeros in the dataset are actually nan values. Treat zeros as nan and impute them using any suitable methods.

Dataset : [Click here](#)

Please note the following:

- Give headings to each step you are doing.
- Do the case study in Python.
- Create a repository in GitHub account as "Public".
- Upload the notebook file (. ipynb) to the repository.
- Please make sure that you are uploading the notebook file including the outputs as well.
- Share the link of this notebook from GitHub in the online text editor provided in Paatshala.