

# Diabetes Dataset Evaluation

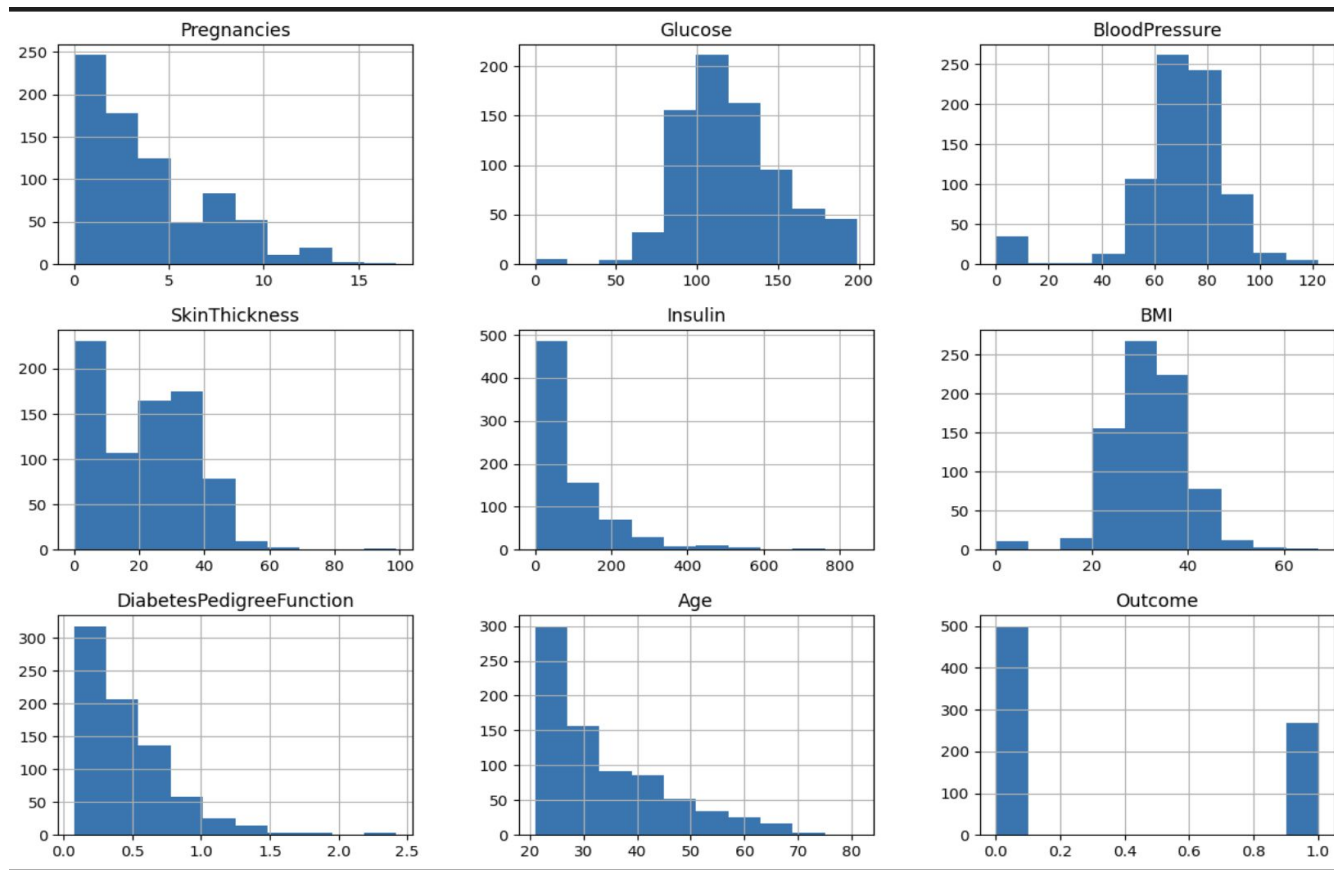
Dhruvin Andipara

# Goal: To predict whether or not a person has diabetes based on certain parameters

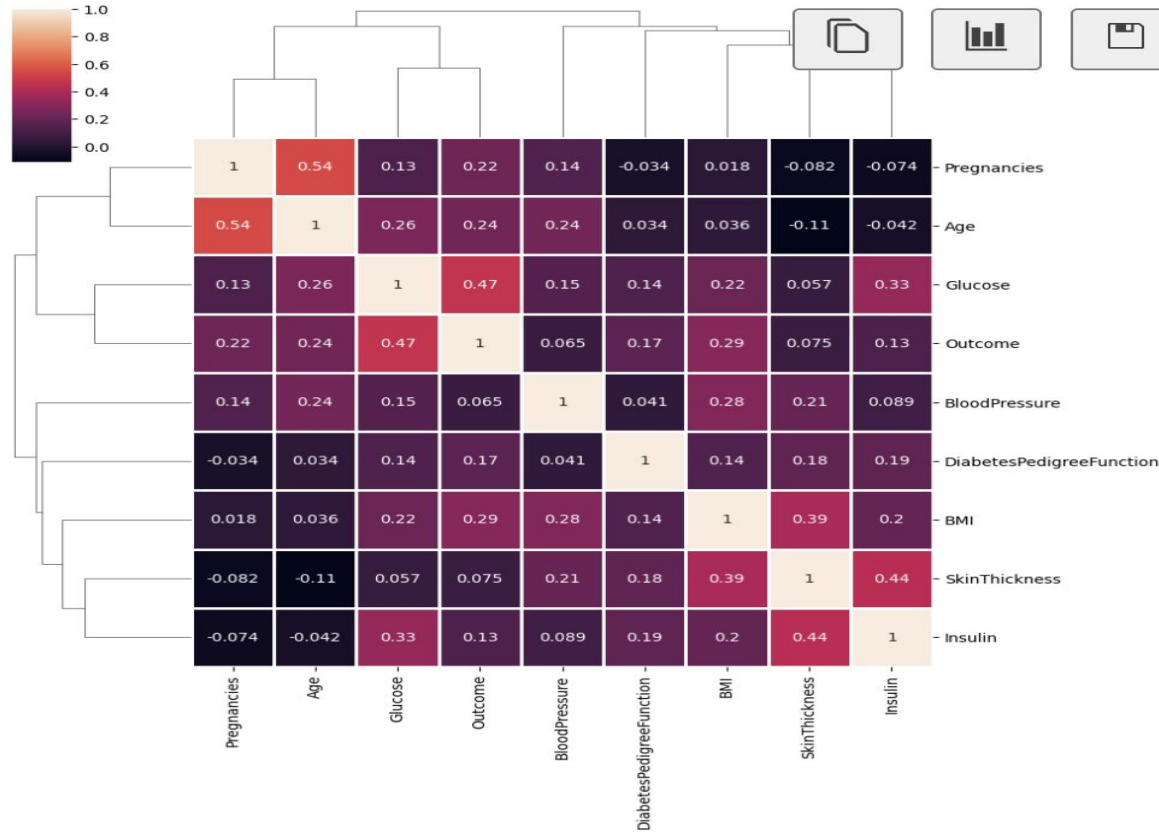
## Processes

1. Data visualisation/EDA
2. Outliers detection and investigation
3. Data pre-processing
4. Model Building and parameter tuning

# Distribution of data



# Correlation Between Features



# Model 1: Decision tree

Total accuracy of model - 0.72

Out of detected values percent of actual positives is - 57%

Percentage of positive values detected out of total positives is - 58%

## Model 2: Random Forest

Total model accuracy = 76%

Out of detected values percent of actual positives is - 71%

Percentage of positive values detected out of total positives is - 43%