**LAB MODEL PRACTICAL EXAMINATION**

**DATA WAREHOUSING AND DATA MINING - U18CSI6203T**

**ACHUTA PRASANTH J - 18BCS020**

The main task we are going to perform here is to determine whether the person is diabetic or non-diabetic. I have downloaded the dataset from an online resource and then I have developed a decision tree classifier which predicts the state of the person based on a few sets of attributes.

**The sets of attributes are as follows:**

1. Pregnancy period - The number of months the person is pregnant
2. Glucose level - The blood sugar level, blood sugar concentration, or blood glucose level is the concentration of glucose present in the blood of humans.
3. BP level - The force of circulating blood on the walls of the arteries.
4. Skin - Skin tone
5. Insulin level - Insulin is a hormone that lowers the level of glucose (a type of sugar) in the blood.
6. BMI - Body Mass Index
7. Pedigree - A value for the pedigree chart to be made
8. Age - Age factor

**Procedure:**

1. We first import the dataset in the google colab notebook.
2. Next , we proceed with the importing the necessary packages and the modules.
3. After that, the features are classified and then the target variable has also been defined.
4. Next, we calculate the accuracy that is been given by the model.
5. After this , we create the decision tree for the values that are been obtained from the dataset with the entropy criteria.
6. Then we obtain the confusion matrix for the same.