In [ ]:
```
#Decison Tree Implementation
```

In [1]:
```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier # Import Decision Tree Classifier
from sklearn.model_selection import train_test_split # Import train_test_split function
from sklearn import metrics #Import scikit-learn metrics module for accuracy calculatio
```

In [2]:
```
columns = ['pregnant', 'glucose', 'bp', 'skin', 'insulin', 'bmi', 'pedigree', 'age', 'l
# Load dataset
data = pd.read_csv("diabetes.csv", header=0, names=columns)
data.head()
```

Out[2]:

	pregnant	glucose	bp	skin	insulin	bmi	pedigree	age	label
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

In [3]:
```
#split dataset in features and target variable
feature_columns = ['pregnant', 'insulin', 'bmi', 'age','glucose','bp','pedigree']
X = data[feature_columns] # Features
y = data.label # Target variable
```

In [4]:
```
# Split dataset into training set and test set
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.15, random_state=
```

In [5]:
```
# Create Decision Tree classifer object
dtree = DecisionTreeClassifier()

# Train Decision Tree Classifer
dtree = dtree.fit(X_train,y_train)

#Predict the response for test dataset
y_pred = dtree.predict(X_test)
```

In [6]:
```
# Model Accuracy, how often is the classifier correct?
print("Accuracy of dtree:",metrics.accuracy_score(y_test, y_pred))
```

```
Accuracy of dtree: 0.7327586206896551
```