EXPERIMENT 4

<u>Aim:</u> Implementation of Classification technique on ARFF files using WEKA.

<u>Theory:</u> This experiment illustrates the use of j-48 classifier in Weka. The sample data set used in this experiment is "student" data available at .arff format. This document assumes that appropriate data preprocessing has been performed. Steps involved in this experiment:

Step 1: We begin the experiment by loading the data (employee1.arff)into Weka.

Step 2: Next we select the "classify" tab and click "choose" button to select the "j48" classifier.

<u>Step 3:</u> Now we specify the various parameters. These can be specified by clicking in the text box to the right of the chose button. In this example, we accept the default values. The default version does perform some pruning but does not perform error pruning.

<u>Step 4:</u> Under the "text" options in the main panel. We select the 10-fold cross validation as our evaluation approach. Since we don't have separate evaluation data set, this is necessary to get a reasonable idea of accuracy of generated model.

Step 5: We now click "start" to generate the model. The ASCII version of the tree as well as evaluation statistic will appear in the right panel when the model construction is complete.

Step 6: Note that the classification accuracy of model is about 78.57%. This indicates that we may find more work. (Either in preprocessing or in selecting current parameters for the classification)

Step 7: Now Weka also lets us a view a graphical version of the classification tree. This can be done by right clicking the last result set and selecting "visualize tree" from the pop-up menu.

Step 8: We will use our model to classify the new instances.

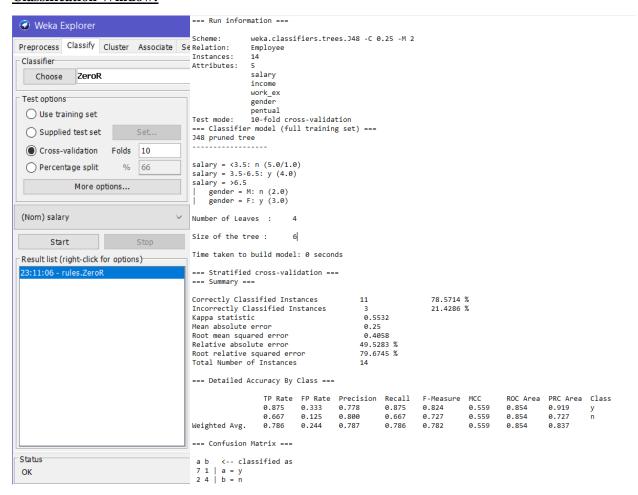
Step 9: In the main panel under "text" options click the "supplied test set" radio button and then click the "set" button. This wills pop-up a window which will allow you to open the file containing test instances.

Code:

```
@relation Employee
@attribute salary{<3.5,3.5-6.5,>6.5}
@attribute income{low,mid,high}
@attribute work_ex{y,n}
@attribute gender{M,F}
@attribute pentual{y,n}
@data
<3.5,high,y,M,n
<3.5,high,n,M,n
3.5-6.5, high, n, F, y
>6.5,mid,n,F,y
>6.5,low,y,F,y
>6.5,low,y,M, n
3.5-6.5, low, y, M, y
<3.5,mid,n,F,n
<3.5,low,y,F,n
>6.5,mid,y,F,y
<3.5,mid,y,M,y
3.5-6.5, mid, n, M, y
3.5-6.5, high, y, F, y
>6.5,mid,n,M,n
```

21/10/2020 Mohd. Aftab Alam 02013302717

Classification Window:



Classification Tree:

