# CS6301 – MACHINE LEARNING - WEEK-1 EXPLORING WEKA TOOL 22-02-2021 MONDAY

**NAME: SRIHARI S** 

**ROLLNO:** 2018103601

COMPUTER SCIENCE ENGINEERING

BATCH - P

**DATASET USED: LABOR** 

#### ATTRIBUTES IN THE DATASET:

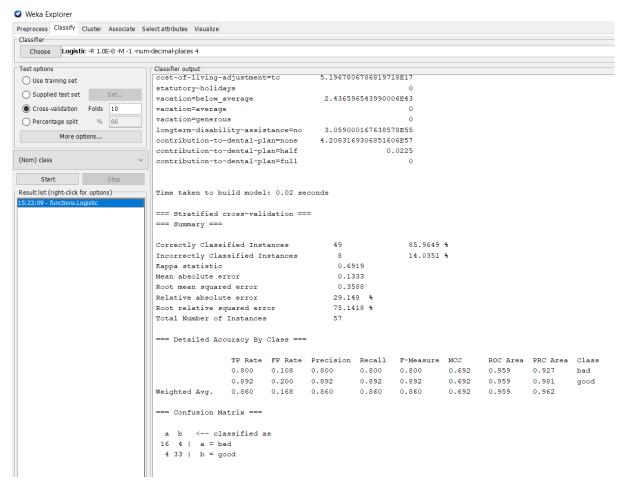
No.	Name
1	duration
2	wage-increase-first-year
3	wage-increase-second-year
4	wage-increase-third-year
5	cost-of-living-adjustment
6	working-hours
7	pension
8	standby-pay
9	shift-differential
10	education-allowance
11	statutory-holidays
12	vacation
13	longterm-disability-assistance
14	contribution-to-dental-plan
15	bereavement-assistance
16	contribution-to-health-plan
17	class

Selecting the attributes which have the major impact on the final classification using filters in WEKA.

Upon applying the filter AttributeSelection under Filters->Supervised->Attributes, we get the following attributes.

No.	Name
1	wage-increase-first-year
2	wage-increase-second-year
3	cost-of-living-adjustment
4	statutory-holidays
5	vacation
6	longterm-disability-assistance
7	contribution-to-dental-plan
8	class

### Task for Level-1: Classification on the dataset using Logistic Regression

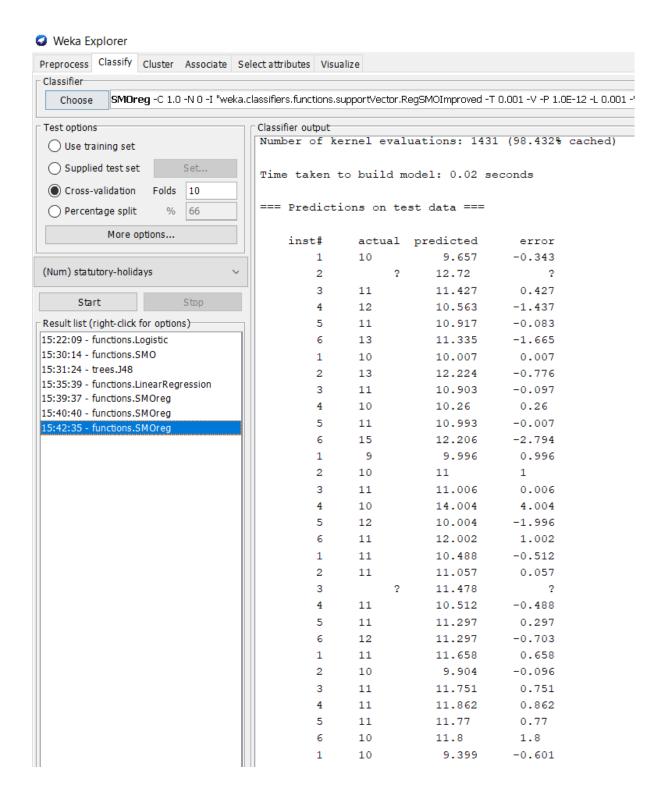


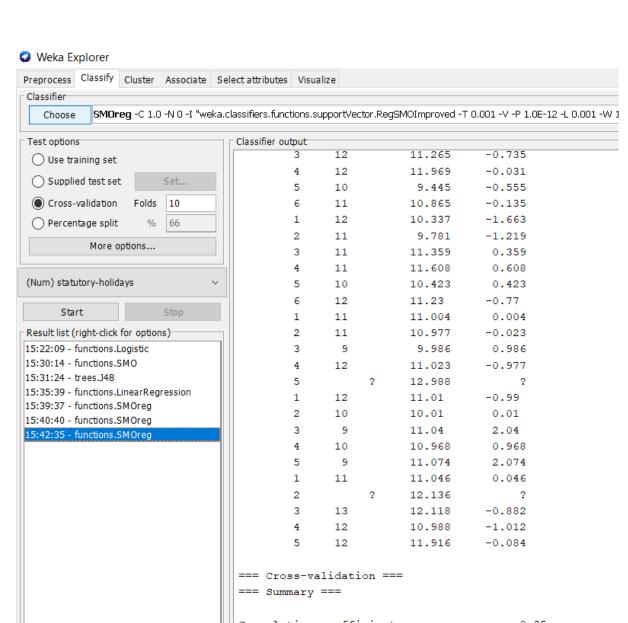
Using logistic regression, we group the employees into two classes – good and bad.

As we can see 85.9% accuracy is established.

The confusion matrix is also displayed in the bottom justifying the accuracy.

## Task for Level-2: Using SVM to the same dataset





<

Correlation coefficient 0.35

Mean absolute error 0.8526

Root mean squared error 1.2363

Relative absolute error 94.2849 %

Root relative squared error 96.7512 %

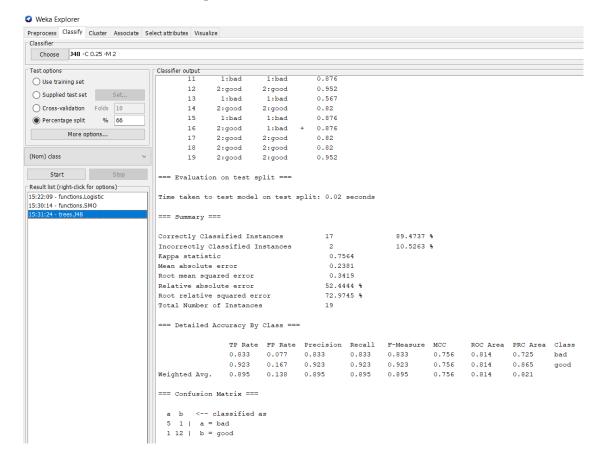
Total Number of Instances 53

Ignored Class Unknown Instances 4

Status

OK

# Task for Level-3: Using Decision Tree to the same dataset



#### Tree Visualizer

