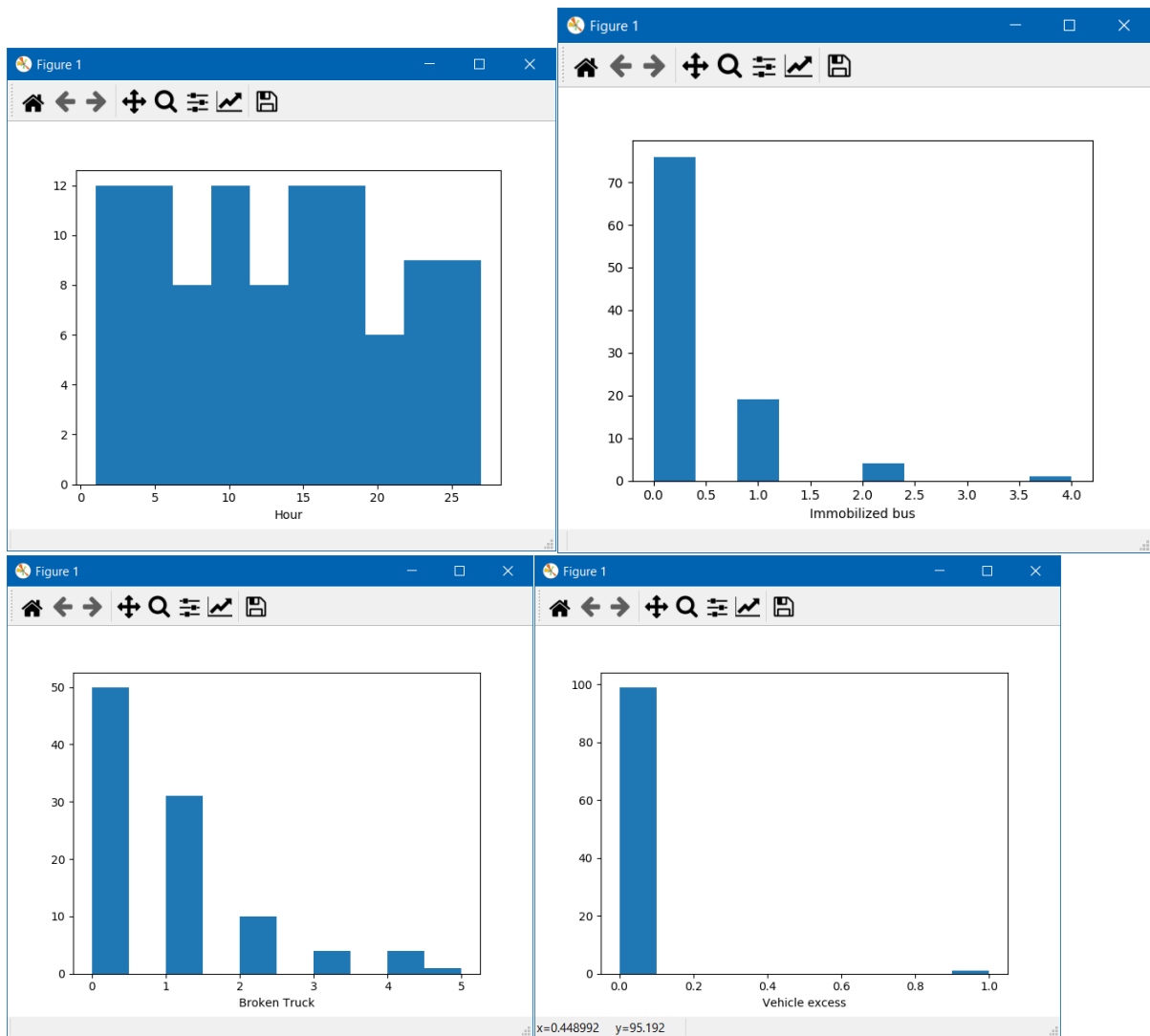


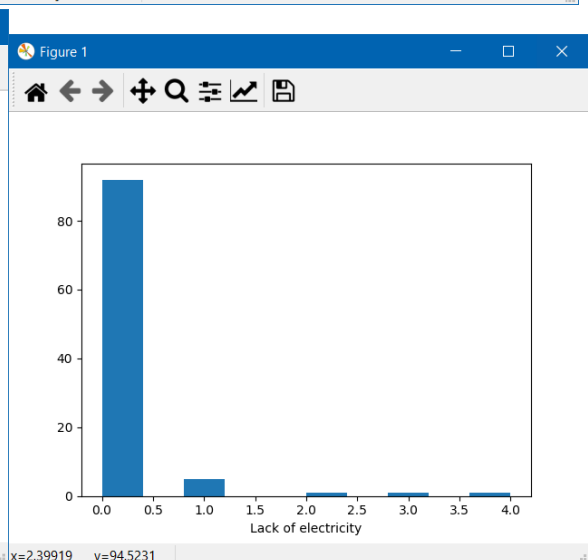
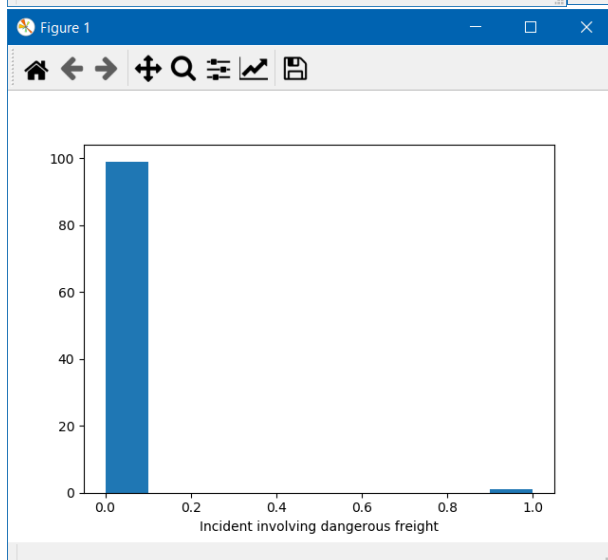
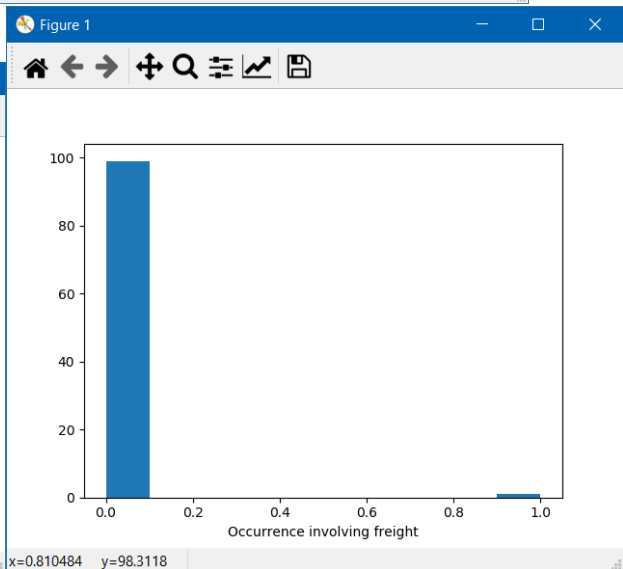
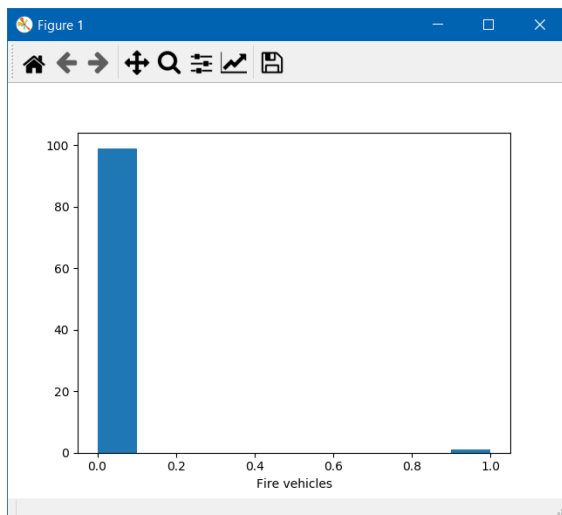
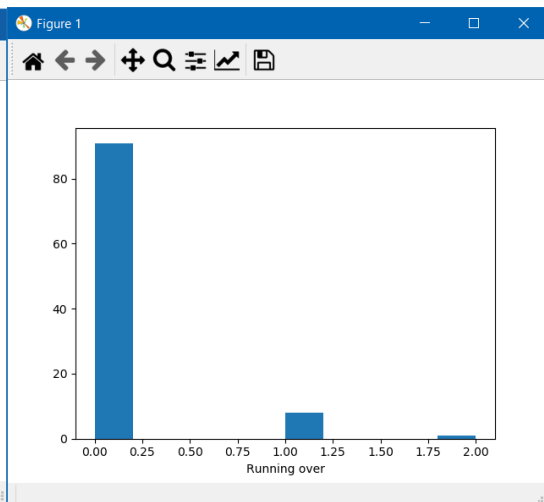
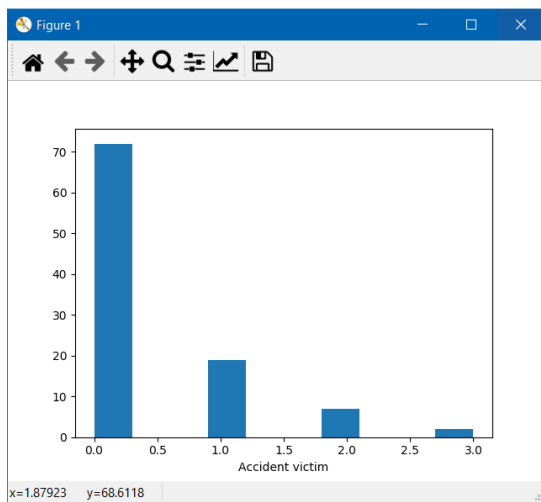
Machine Learning Assignment 2 (CSCE 633)

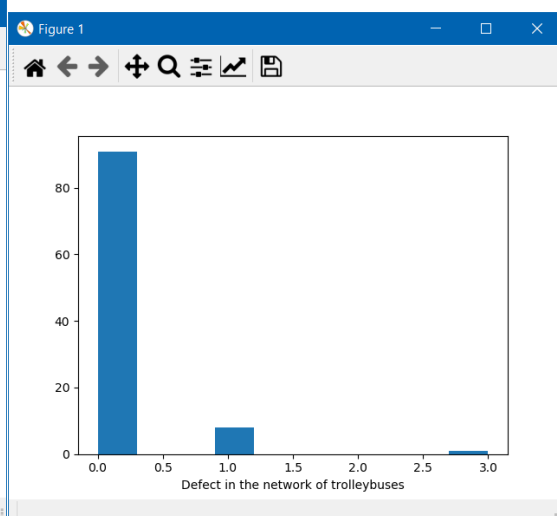
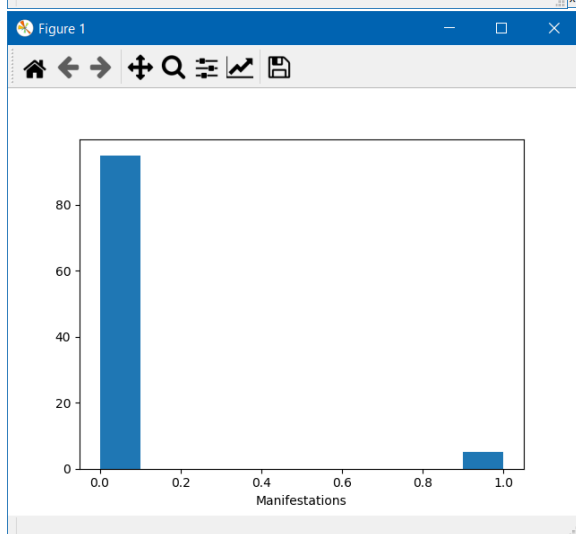
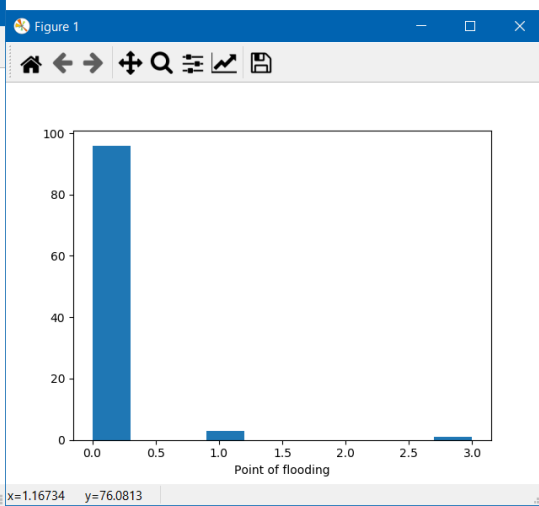
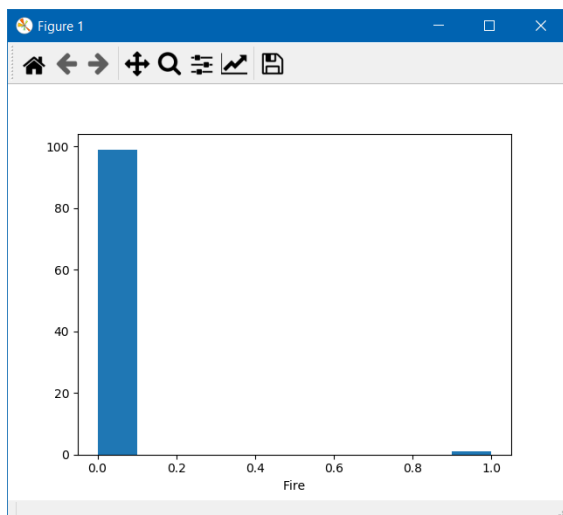
Name: Guru Sarath

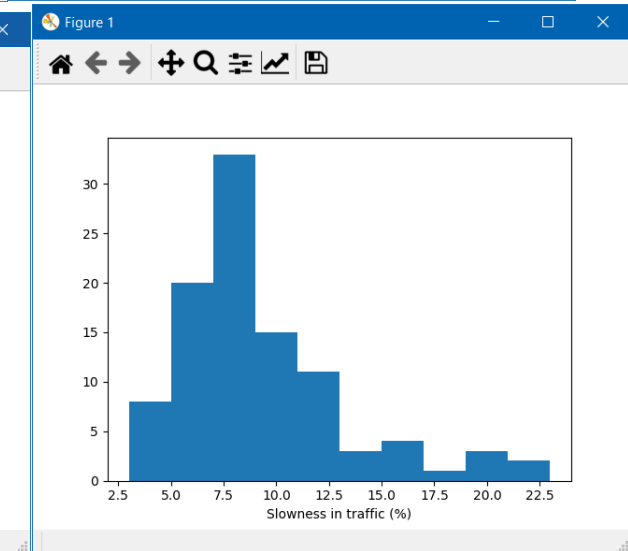
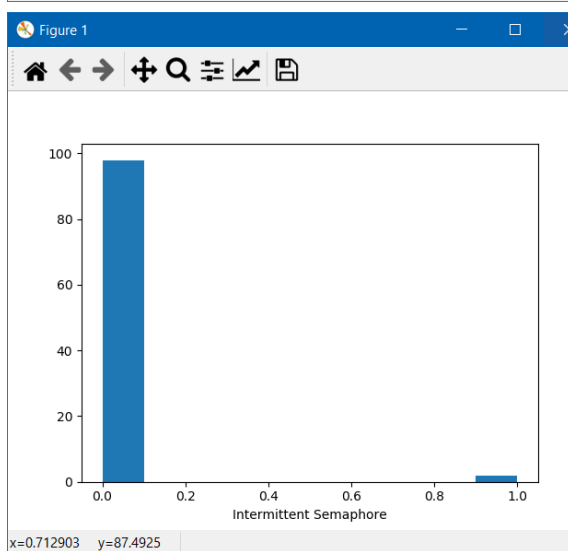
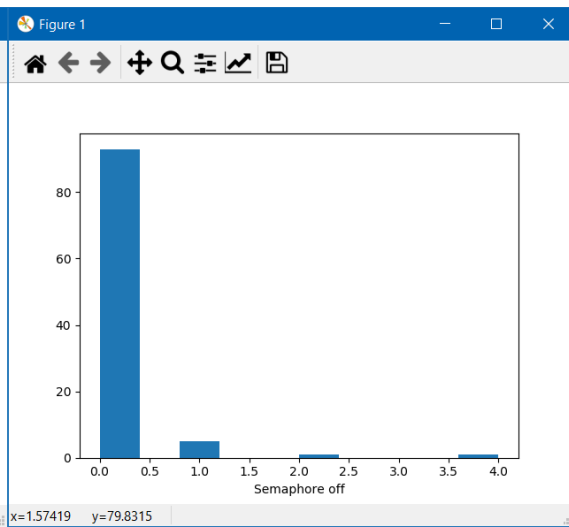
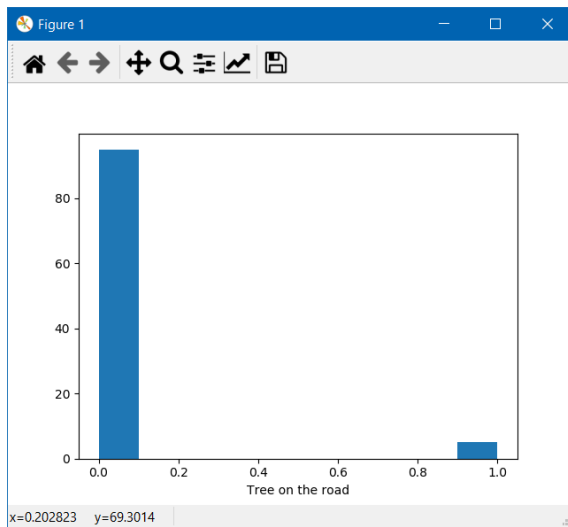
UIN: 829009551

(i)









(ii)

Hour : 0.6707105739042725
Immobilized bus : 0.15510125658984802
Broken Truck : 0.1471936844073055
Vehicle excess : -0.1464640400232247
Accident victim : 0.12678483696305048
Running over : -0.012205336668602063
Fire vehicles : 0.18409716141808105
Occurrence involving freight : 0.05695823778680961
Incident involving dangerous freight : 0.03153045306055532
Lack of electricity : 0.5737326666240945
Fire : -0.04475290111820755
Point of flooding : 0.4556179652155694
Manifestations : -0.0557212122061783
Defect in the network of trolleybuses : -0.16783420054885045
Tree on the road : -0.07893838395875258
Semaphore off : 0.42866617010993074
Intermittent Semaphore : -0.13589889585548648
Slowness in traffic (%) : 1.0

(iii)

Number of parameters = 18

<See the output in the jupyter notebook>

(iv)

RSS = 501.8642602243618

Correlation with actual output = 0.8197758586024668

<See the output in the jupyter notebook>

(v)

Correlation between most features is mostly less than 0.5, this means that most of the features are uncorrelated. Hence, all features are important for the prediction process.

<See the output in the jupyter notebook>

(vi)

The threshold was obtained by taking mean of output value both test data and train data.

Accuracy of classifier = 62.8%

<See the output in the jupyter notebook>

(vii)

Best accuracy (85%) is achieved in train dataset is when C = 1 or 0.5

Accuracy in test dataset for C=0.5 is 65.7%

<See the output in the jupyter notebook>