**CST8390 - Lab 7 Answer Document**

**Regression**

1 a. Number of instances: 513 b. Number of attributes: 13 + 1 class attribute

c. Attribute information

CRIM: Per capita crime rate by town

ZN: Proportion of residential land zoned for 25,000 sq. ft.

INDUS: Proportion of non-retail business acres per town

CHAS: Charles River dummy variable (1 if tract bounds or 0 otherwise)

NOX: Nitric oxides concentration (parts per 10 million)

RM: Average number of rooms per dwelling

AGE: Proportion of owner-occupied units built prior to 1940

DIS: Weighted distances to five Boston employment centres

RAD: Index of accessibility to radial highways

TAX: Full-value property-tax rate per $10,000

PTRATIO: Pupil-teacher ratio by town

B: 1000 (Bk – 0.63)^2 where Bk is the proportion of blacks per town

LSTAT: % lower status of the population

MEDV A.K.A. class value: Median value of owner occupied homes in $1000’s

2 )

a. Median = ROUNDDOWN(numInstances / 2) = 21.1 \* $1000 = $21,100

b. Median = ROUNDDOWN(numInstances / 2) = 8.74

c. Median = ROUNDDOWN(numInstances / 2) = 0.33169

3)

a.

class =

-0.0914 \* CRIM +

0.0577 \* ZN +

-0.0931 \* INDUS +

2.8323 \* CHAS=1 +

-72.568 \* NOX +

2.5705 \* RM +

-1.2806 \* DIS +

0.2532 \* RAD +

-0.0132 \* TAX +

-0.7959 \* PTRATIO +

0.0094 \* B +

-0.6428 \* LSTAT +

65.9273

b.

CHAS had the highest positive influence on housing prices with a coefficient of 2.8323 and RM was second highest with a coefficient of 2.5705.

c.

NOX had by far the highest negative influence on housing prices with a coefficient of -72.568 and the second highest was DIS with a coefficient of -1.2806.