Statistical Learning

Computer Homework A

106061703 藍國瑞

1.

(a)

	Coefficient	Std. Error	t-statistic	p-value
(Intercept)	-16.4411	2.78824	-5.89657	1.92E-09
Amb Temp	-0.25638	0.029344	-8.73679	1.42E-18
CH4	9.36368	3.0818	3.03838	0.001193
CO	22.9456	1.22511	18.7293	4.46E-77
NMHC	-7.44983	3.11749	-2.38969	0.008442
NO	-0.1034	0.356399	-0.29013	0.385863
NO2	0.234533	0.353019	0.664363	0.253238
NOx	-0.27637	0.353746	-0.78127	0.217333
O3	0.090552	0.009403	9.63012	3.83E-22
PM10	0.240746	0.00427	56.3784	0
RAINFALL	-0.29772	0.113022	-2.63416	0.004225
RH	-0.00498	0.014337	-0.34749	0.364117
SO2	1.36373	0.120153	11.35	5.97E-30
THC	1.10823	3.07449	0.360458	0.359257
WD HR	6.10E-05	2.96E-03	0.020592	0.491786
Wind Direc	0.005828	0.002923	1.99364	0.023111
Wind Speed	0.149399	0.286458	0.521539	0.301002
WS HR	-0.90975	0.333691	-2.72631	0.003209

RSE=11.565600660920628

R^2=0.47732791607921754

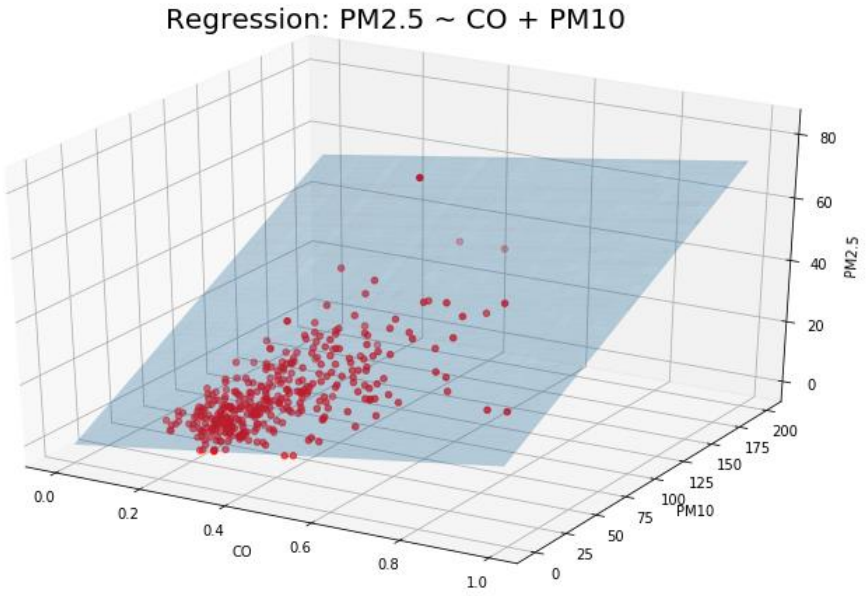
(b)

	Coefficient	Std. Error	t-statistic	p-value
CO	19.7315	0.755315	26.1235	3.16E-145
PM10	0.27185	0.004129	65.8448	0

RSE= 12.026663562708993

R^2= 0.43385485033902293

(c)



2.

(a)

The predicted value of the test data:

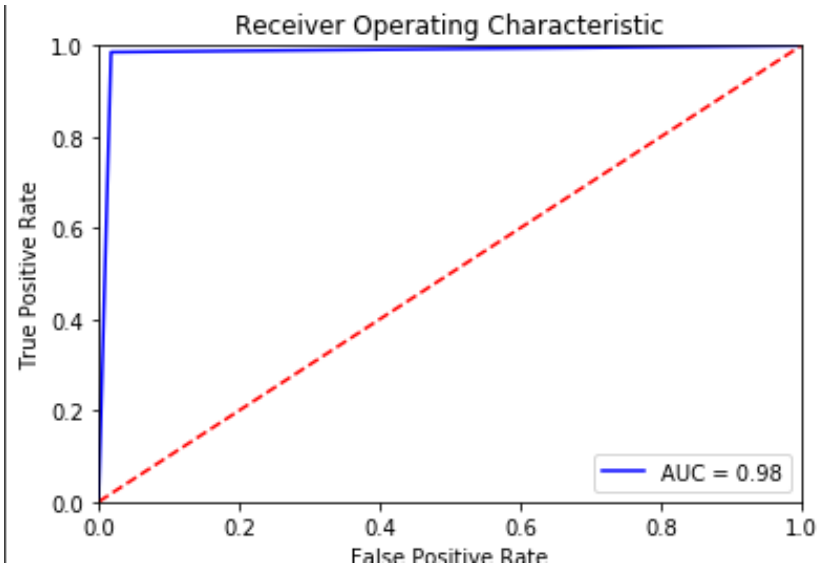
The result is in the file of “HWA_2_(a)_predictedvalue.xlsx”

Confusion matrix:

array([[666, 9],

[12 , 623]], dtype=int64)

ROC curve:



(b)

The predicted value of the test data:

The result is in the file of “HWA_2_(b)_predictedvalue.xlsx”

Confusion matrix:

array([[472, 13, 1],
[7, 407, 8],
[1, 3, 470]], dtype=int64)