CS 6301 - Machine Learning Lab - Week 6

Date: 08.09.2023

TITLE

IMPLEMENTATION OF LINEAR REGRESSION (SINGLE AND MULTIPLE LINEAR MODEL)

TASKS

1) A simple linear regression model to express drain current I_d (in milli Ampere) as a function of ground-to-source voltage Vgs (in Volts) for a MOS transistor is hypothesized. The drain current and the ground-to-source voltage data were measured and is shown in the following table.

Drain Current I_d (mA)	**Gate-to-Source Voltage (V)**
0.734	1.1
0.886	1.2
1.04	1.3
1.19	1.4
1.35	1.5
1.50	1.6
1.66	1.7
1.81	1.8
1.97	1.9
2.12	2.0

   (a) Draw a Scatter diagram for these data. Does a straight-line relationship seem plausible?

   (b) Fit a simple linear regression model to these data.

2) Use Coronary Heart Disease (CHD) problem available at (https://axon.cs.byu.edu/Dan/478/assignments/CHD.txt)

   (a) First, use simple linear regression to build a linear model of the data.

   (b) Report the parameters of the model (w_0, w_1 and R^2).

   (c) Plot the original (probability) data and graph the linear model your program produced.

   (d) What does the model predict for the probability of someone 41 years old suffering of CHD?

3) A mechanical engineering study indicates that fluid flow through a pipe is related to pipe diameter and slope, and the relationship is power equation $Q=a_0D^{a}{}_1S^{a}{}_2$. Use multiple linear regression to analyse this data. Then use the resulting model to predict the flow for a pipe with diameter of 2.5ft and a slope of 0.025 ft/ft.

Experiment	**Diameter, D (ft)**	**Slope, S (ft/ft)**	**Flow, Q (ft^3/s)**
1	1	0.001	1.4
2	2	0.001	8.3
3	3	0.001	24.2
4	1	0.01	4.7
5	2	0.01	28.9
6	3	0.01	84.0
7	1	0.05	11.1
8	2	0.05	69.0
9	3	0.05	200.0

(a) Indicate the results of constants a_0, a_1, a_2 and the predicted flow for the given values of diameter and slope.