Robert Krall

Assignment 2

SEIS 763 – Machine learning

4. What are the regression coefficients (thetas)?

[[ 0.08025966

-1.47939073

0.46962059

-0.01341834

9.67308711

-0.85650078

-1.73484051

-2.75096823

0.58637873

0.45934283]]

5. How do you interpret those numbers?

The smaller the number the less range in values also meaning they are closer to the line for linear regression.

6. If you need to identify one outlier record, which record is a potential outlier? How do you reach this conclusion?

I would start to look at the entries for smokers. They seem to have a large coefficient. It was the largest number.

7. If you need to identify one or few useless features (independent variables or predictors), which one(s) will you choose? Why do you reach this conclusion?