Ian Anderson

CS 4320

1/25/22

Linear Regression

x1 seems to have a slope of about 8.3 (500/60). The points are very spread out.

x2 has a steeper slope of about 25 (1000/40). The points are much tighter and in a linear pattern.

x3 seems to have a negative slope of about -6.25 (250/40). It's pretty difficult to see a slope.

x4 could possibly have a positive slope. If it did, I'd guess it was about 3.

x5 has a negative slope of about -10. The points are about as close together as they are in x1.

x6 doesn't look like it has a slope at all. The label axis goes from 0 to 2000, while the feature axis only goes from 0 to about 10, so there could be a significant slope without it being very noticeable.

All of the histograms are very flat with no bell curve at all.

With eta0 = 0.001:

MSE loss obtained on the training data:

3278.6987741749736

MSE loss obtained on the testing data:

3094.255816552861

Linear model coefficients:

9.34499149

31.27619539

-4.86618287

2.82016492

-10.14988273

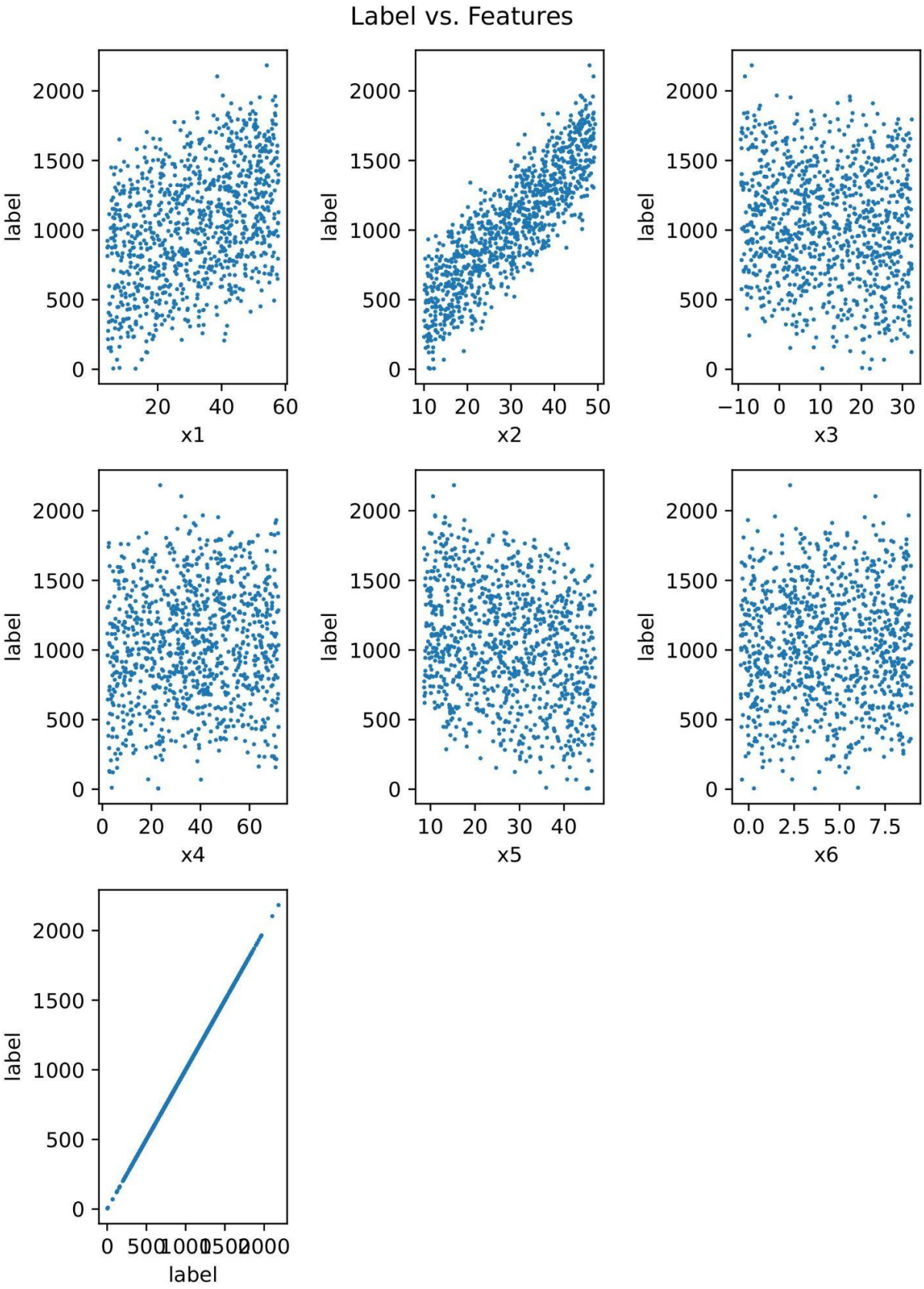
4.12915045

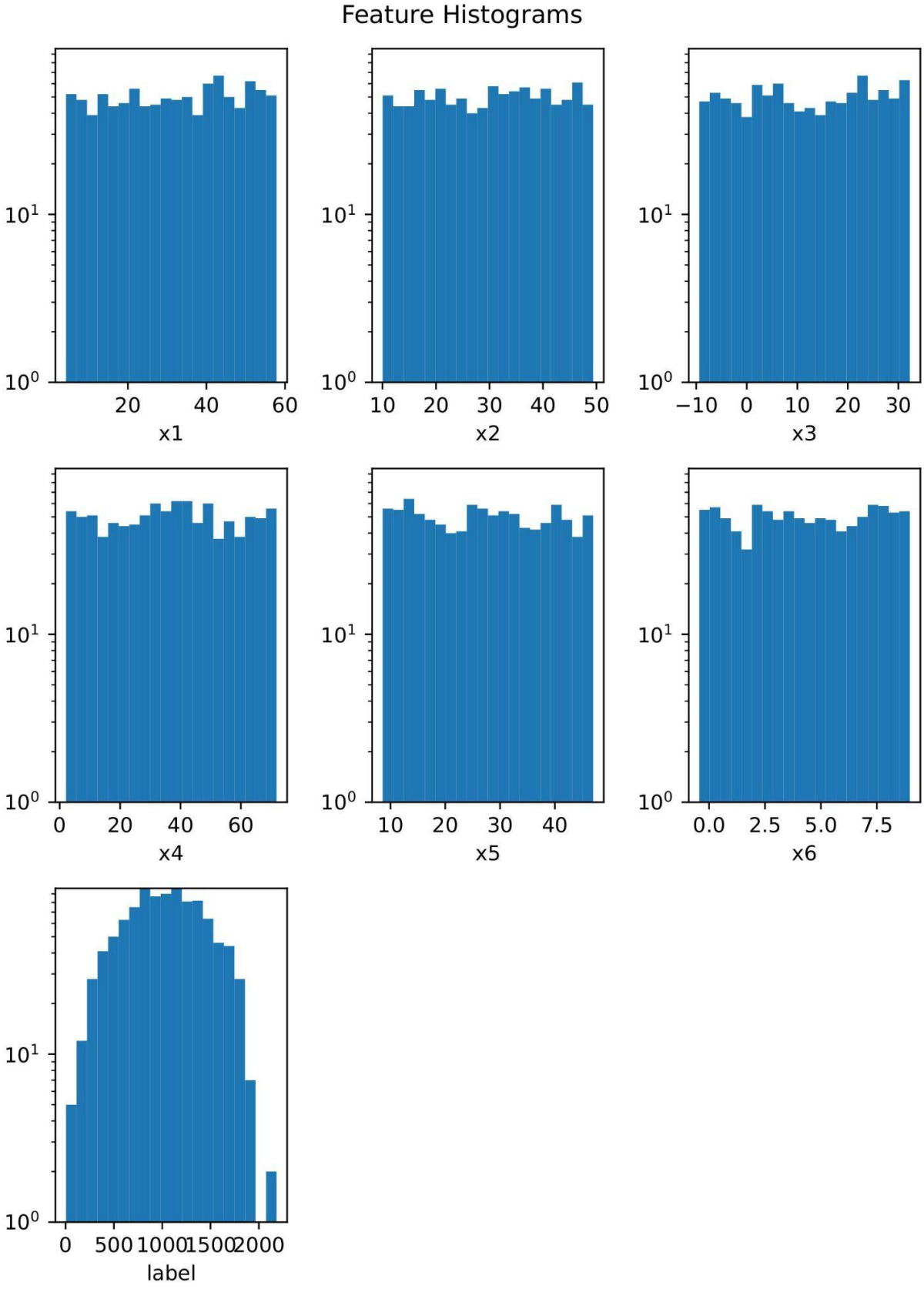
Model function:

20.71507470506372 + (9.344991488011429*x1) + (31.27619539282282*x2) +

(-4.866182865720787*x3) + (2.820164916947025*x4) +

(-10.149882727682439*x5) + (4.129150448637257*x6)


Label vs. Features
label
x1
x2
x3
x4
x5
x6
label


Feature Histograms
10^1
10^0
20
40
60
x1
10
20
30
40
50
x2
−10
0
10
20
30
x3
0
20
40
60
x4
10
20
30
40
x5
0.0
2.5
5.0
7.5
x6
0
500
1000
1500
2000
label