ASSIGNMENT-2

Initially the data of latitudes and longitudes are normalised by using (x-x_min)/(x_max-x_min) normalized function and new features were generated according to the degree of the polynomial to be fit.

For Gradient Descent Model:

Degree	Train Error	R2 ERROR	RMSE	Weights(W)
1	2504.96	2.50102	0.1287	[0.20848 0.09536 -0.09984]
2	2416.23	6.21248	0.1267	[0.14239 0.526028 -0.474625 -0.17928 -0.03033 0.13115]
3	2243.46	12.6526	0.1212	[0.229525 -0.1399825 1.79395354 -2.0420335 -0.5348373 -1.5242099 1.4634358 1.9033192 0.4274262 -1.57894]
4	2165.91	15.6862	0.1202	[0.200943 -0.090968 1.27206105 -0.4016413 -1.729962 -0.2121626 -1.98726 1.0842712 1.66736 0.5790535 0.026383 0.2784248 1.569726 -0.9356409 -1.40090874]
5	2124.14	17.6060	0.1179	[0.17442 0.00305 1.31990 -0.2855097 -1.46935 -0.94191 0.081287 -1.98970689 0.608159 1.321212 1.636402 -0.219968 -0.292898 0.280615 0.39492675 1.56601267 -0.10389244 -0.71290387 0.67861609 -1.26979099 -0.87991535]
6	2110.53	17.7790	0.1179	[0.17048744 -0.08962674 1.50332933 -0.02050638 -1.53132001 -1.52887083 -0.15819974 0.19751607 -1.95866819 0.57652114 0.99373625 1.036232 1.56734689 -0.37704709 -0.29113032 0.29275535 0.31045194 0.60387746 1.19626403 0.06177757 -0.36096084 -0.31454571 0.73898155 -0.59345753 -0.86432248 -0.03954899 -0.88730104 -0.15256575]

Learning rate: (1e-7)
Break criteria: (E-E ’ ≤ 5e-3)
Number of Iterations: 25000

Observations and Results:
1)As higher degree polynomials are fit to the data, the training error decreases while the R2 error increases, indicating some overfitting.
2)Degree 6 polynomial has an over fit.
3)Degree 4 polynomial has a best fit.

L1-Regularization:

Degree	Train Error	R2 ERROR	RMSE	Reg coeff	WEIGHTS
6	645.364	14.2374	0.12078	.65	[0.1845707 -0.0398244 1.7192297 -1.796607 0.4454285 -0.80041 -0.581372 -0.11608601 -0.8177302 -0.0901393 -0.87337241 0.404632 2.5543963 -0.1299462 1.07814396 0.330576 -0.42566046 -0.057612 0.68279 -0.36274973 0.723591 - 0.446528 0.108436 -0.270638 0.106849 -0.79221959 -0.637207 0.26698]

After L1-regularization:

Regularization coeff = 0.65. The accuracy after regularization is close to that of degree 4 polynomial.So we can say that degree 6 polynomial with regularization will give same result as degree 4.

L2-Regularization:

Degree	Train Error	R2 ERROR	RMSE	Reg coeff	WEIGHTS
6	648.364	11.2374	0.12078	0.25	[0.169764 0.180695 0.6551339 -0.069974 -0.586138 -0.578445 -0.16013752 -0.315269 -0.455507 -0.43127 -0.335481 0.077144 0.757779 0.553306 -0.158445 -0.19157 0.2424709 1.001346 0.63467 -0.082375 0.112650 .801614 0.15349 -0.17163259 0.36499707 -0.42624223 -0.25293941-0.89574433]

After L2-regularization:

Regularization coeff = 0.25

With low regularization coefficients we are able to get the best results and almost similar to degree 4 polynomial.