Foundations of Data Science (CS F320)
Assignment - 1

TEAM MEMBERS:

1. **Shreeya Nelakar 2017A7PS0093H**
2. **Prateek Agarwal 2017A7PS0075H**
3. **Shriya Choudhary 2017AAPS0409H**

1. Dataset Information:

The dataset used on which our model was fit was the 3-d Road Network, where the latitude, longitude and altitude values were recorded from 434874 different points. Models were built using linear regression based on different types of descent algorithms.

Attributes:

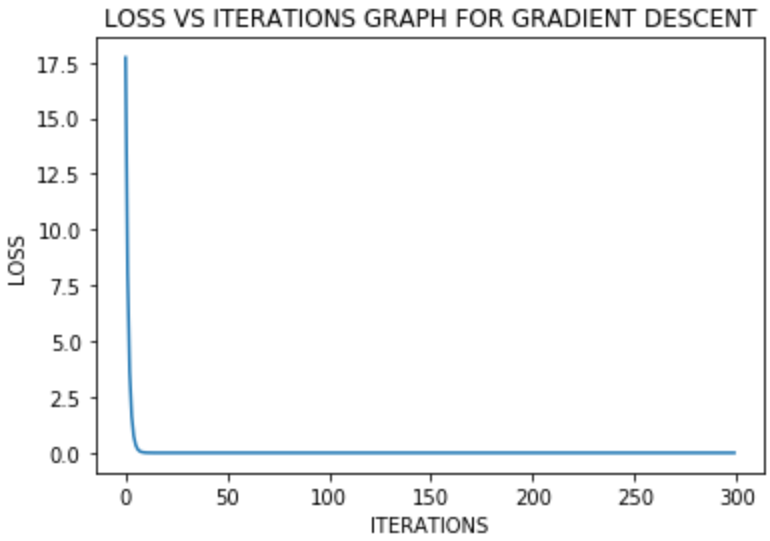
1. Latitude
2. Longitude
3. Altitude

Dataset characteristics	Sequential	**Number of Instances:**	434874
Attribute Characteristics:	Real	**Number of ATtributes:**	3

Gradient Descent Method:

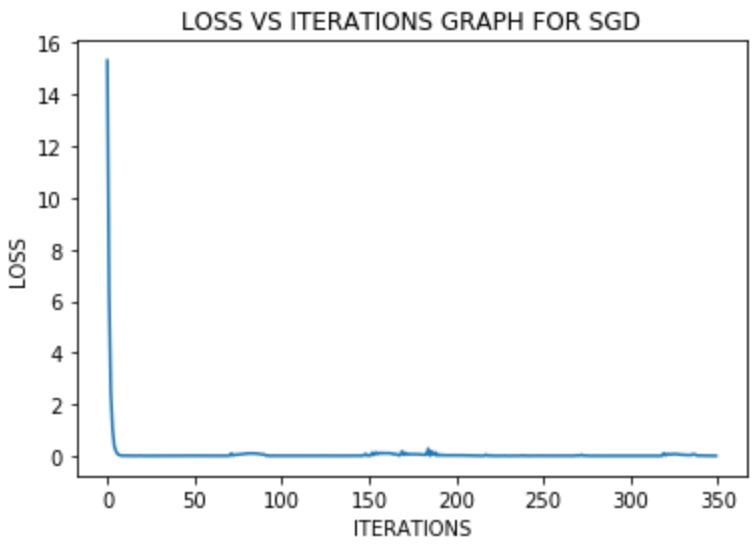
1. Initialization for the weights: [0,0,0]
2. Learning rate: 0.1

3. Stopping criteria:Iteration number =300
4. Final Weights: [0.20348837 -0.05331657 0.06546252]
5. Cost: 0.008291408541182684
6. RMS ERROR: 0.128774287349477
7. R squared value: 0.7379359928576632



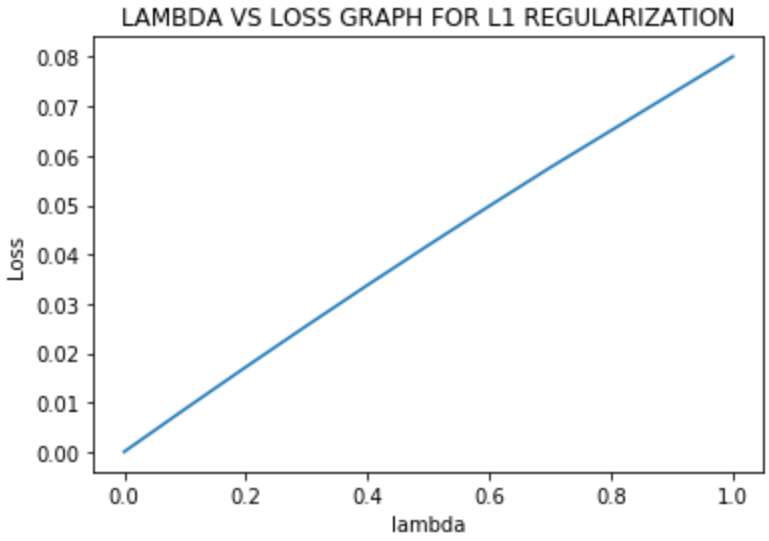
Stochastic Gradient Descent Method:

1. Initialization for the weights: [0,0,0]
2. Learning rate: 0.1
3. Stopping criteria:Iteration number =50
4. Final Weights: [0.00270535 ,-0.15307194, -0.02170379]
5. Cost: 0.009586743137742385
6. RMS ERROR: 0.1323762527066123
7. R squared value: 0.7230704926609286



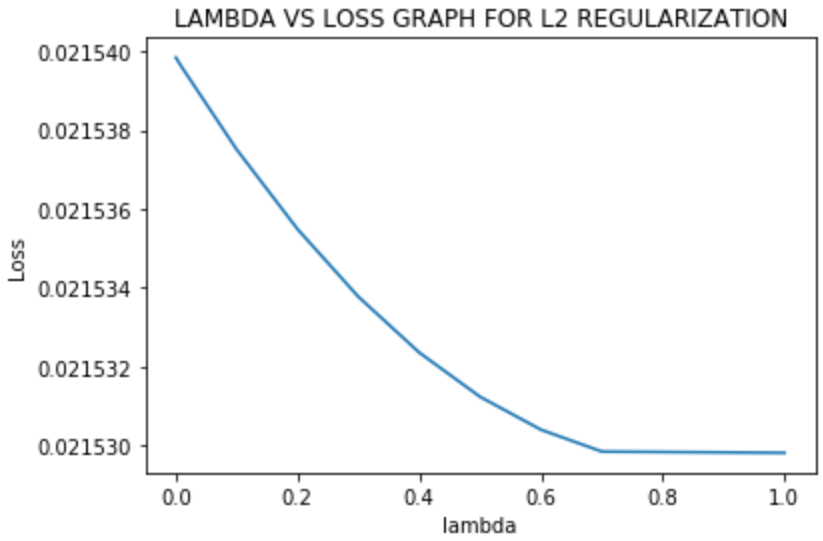
Gradient Descent Method with L1 Regularization:

1. Initialization for the weights: [0,0,0]
2. Learning rate: 0.1
3. Stopping criteria:Iteration number =300
4. Final Weights: [0.0027061, -0.15309628 , -0.02164847]
5. Cost: 0.009776319171118564
6. RMS ERROR: 0.13029686136888752
7. R squared value: 0.7317022837429292



Gradient Descent Method with L2 Regularization:

1. Initialization for the weights: [0,0,0]
2. Learning rate: 0.1
3. Stopping criteria:Iteration number =300
4. Final Weights: [0.00777788 -0.15354583 -0.01744042]
5. Cost:0.021529806897285814
6. RMS ERROR: 0.1678969495243312
7. R squared value: 0.5545133659657244



Normal Equations Method:

The process uses various matrix manipulation techniques that solve the normal equations.
Obtained W matrix:

1. Final Weights: [-0.20875515 0.09525453 -0.10010714]
2. RMSE: 0.12841694924590594
3. R2 value: 0.7393883872242046

Comparative study of all models with measures like R2 and RMSE:

Method	RMS Error	R squared value
Normal Equations	0.12841694924590594	0.7393883872242046
Gradient Descent	0.128774287349477	0.7379359928576632
Stochastic Gradient Descent	0.1323762527066123	0.7230704926609286
GD with L1 regularization	0.13029686136888752	0.7317022837429292
GD with L2 regularization	0.1678969495243312	0.5545133659657244