Machine Learning (60050) : Assignment 1

K Rahul Dev - 15CS10021

Part c :

Experimenting with Combination of Features

- **Features** : Linear Combinations,Quadratic Combinations,Cubic Combinations
- **Error Function** : Mean Square Error
- **Optimization Function** : Gradient Descent
- **Learning Rate** : 0.05

Procedure :

1.Split the Dataset into Train Set and Test Set [80:20]
2.Theta Values (Learning Parameters) are initialized randomly.

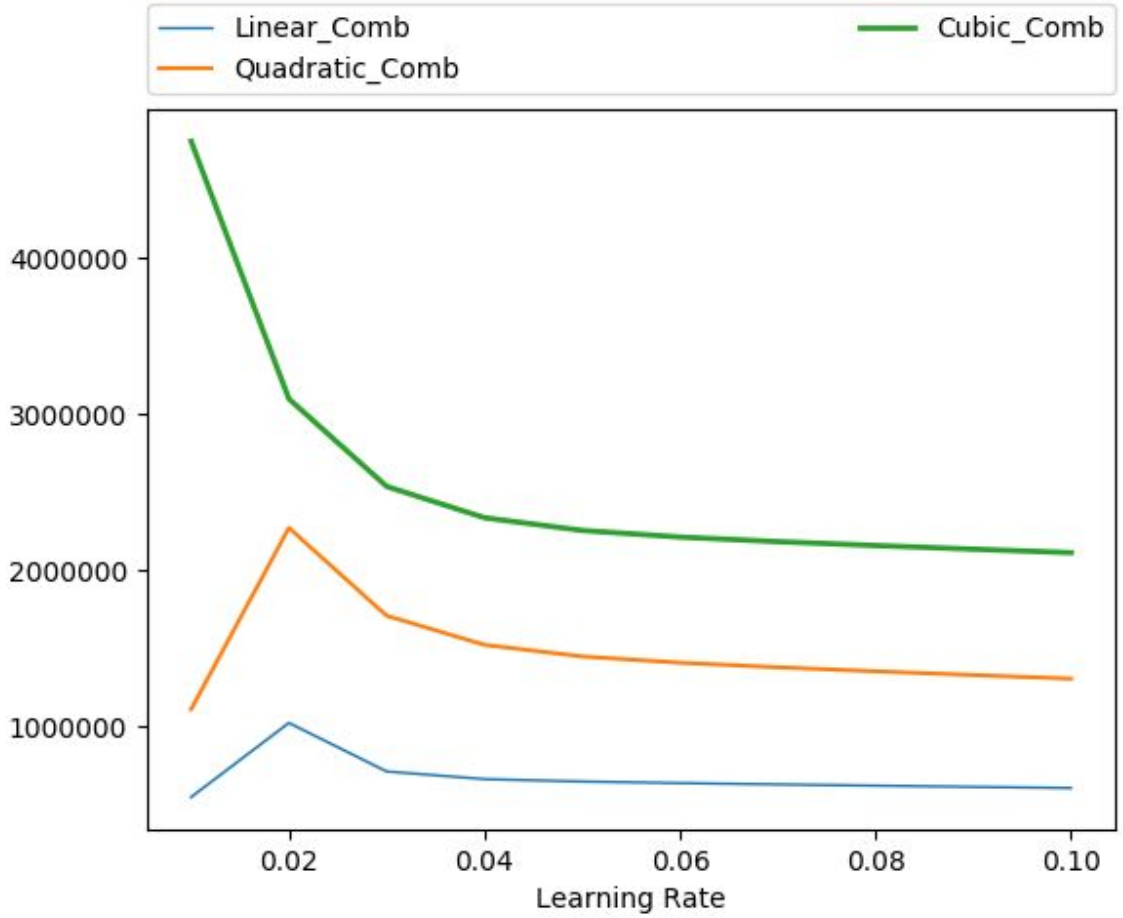
Final Learned Values (Theta Values)

Learning Rate	Linear Combination	Quadratic Combination	Cubic Combination
0.01	[-0.10543436 0.08786005 0.18436629 0.96560801 0.4582587]	[-0.1034223 0.08625711 0.31018084 0.96365585 0.61573651]	[0.24570956 0.09335038 0.6386838 0.9767122 0.77345081]
0.02	[-0.08032039 0.09274871 0.29872628 0.975974 0.5199437]	[-0.03273072 0.09265528 0.50792871 0.97581969 0.70279975]	[0.02963346 0.09273884 0.60687929 0.97583202 0.76285811]
0.03	[-0.13119729 0.09240358 0.26999613 0.97533727 0.50147136]	[-0.10638316 0.09219639 0.48069087 0.97505901 0.69050685]	[-0.04508841 0.09227843 0.58996231 0.97504633 0.75735701]
0.04	[-0.14194175 0.09210503 0.25853204 0.974731 0.4947055]	[-0.12863299 0.09179553 0.46403546 0.97433607 0.68320598]	[-0.06990571 0.0918725 0.57837988 0.97429527 0.75367127]
0.05	[-0.14229371 0.09181967 0.2516477 0.97413377 0.49103284]	[-0.13388513 0.09141558 0.45105242 0.9736275 0.67761615]	[-0.07750236 0.0914865 0.56871677 0.97355746 0.75062976]
0.06	[-0.14009194 0.09153865 0.24599701	[-0.13368306 0.09104403 0.43941974	[-0.0793005 0.09110835 0.55977844

	0.97353992 0.48818179]	0.97292545 0.67263536]	0.97282541 0.74782297]
0.07	[-0.13732519 0.09125974 0.24072876 0.972948 0.48557469]	[-0.13180269 0.09067679 0.42837437 0.9722273 0.66790197]	[-0.0791994 0.09073391 0.55115609 0.97209654 0.74511092]
0.08	[-0.13447746 0.09098236 0.23563351 0.97235763 0.48306965]	[-0.12944631 0.09031252 0.41766983 0.97153219 0.66329968]	[-0.07849795 0.0903618 0.54271253 0.97136998 0.74244674]
0.09	[-0.13166087 0.09070636 0.23066053 0.97176871 0.4806325]	[-0.12698579 0.08995078 0.40722644 0.97083983 0.65879137]	[-0.0776178 0.08999155 0.53440221 0.97064546 0.73981502]
0.10	[-0.12889943 0.09043168 0.22579667 0.9711812 0.4782546]	[-0.12453228 0.0895914 0.39701691 0.9701501 0.65436479]	[-0.07669213 0.08962301 0.52620973 0.96992284 0.73721067]

Plot :

Test RMSE vs Different Combination of Features



From the above Graph we can infer that :

I. **Cubic Combination** results in tremendously larger values of RMSE and decreases with increase in Learning Rate,but never attained the Minimum

Ii. **Quadratic Combination** resulted in RMSE increase initially may be due to random initialization,and then started decrease with increase in Learning Rate,also never attained the minimum.

Iii. **Linear Combination** resulted though initially increased,then gradually decreased with increase in Learning Rate, but attained the **Minimum** of the three.

From the above observations, and also considering less number of features for the given problem,**Linear Combination** is most preferable.

PS : Because of Random initialization,every execution of the trained model gives different Learned Values

The above graph is corresponding with the values submitted in part_c result file.