Hope AI

1.Problem statement :

By using (ML) and the dataset the client has provided which includes age , number of children, bmi index etc and using the supervised learning(numerical data), to predict the person's insurance charges based on the given parameters in the dataset .

STEPS:

*Machine Learning

*Supervised learning

*Regression(numerical data)

2. Number of columns:1339

Number of rows: 6

3.Conversion from string to nominal data has been done

4,5. Finding the best Model **:**

1.Simple linear regression(r_score)=0.7894

2.Multiple linear regression(r_score)=0.7894

3.Support vector machine:

SL.NO	Hyper Parameter	Linear (r_score)	Rbf(non-linear) (r_score)	Poly (r_score)	Sigmoid (r_score)
1.	C10	0.4624	-0.0322	0.0387	0.0393
2.	C100	0.6288	0.3200	0.6179	0.5276
3.	C500	0.7631	0.6642	0.8263	0.4446
4.	C1000	0.7649	0.8102	0.8566	0.2874
5.	C2000	0.7440	0.8547	0.8605	-0.5939

6.	C3000	0.7414	0.8663	0.8598	-2.1244

4.Decision Tree :

SI.NO	CRITERION	SPLITTER	r_score
1	squared_error	best	0.8975
2	squared_error	random	0.6327
3	friedman_mse	best	0.8999
4	friedman_mse	random	0.9373
5	absolute_error	best	0.9514
6	absolute_error	random	0.9250
7	poisson	best	0.9328
8	poisson	random	0.9109

5.Random forest (n_estimators=100, random_state=0)
r_score = -0.8538

6)Final model - Decision tree(absolute error,best) r_score=0.9514
The above model has the best r_score out of all the other models , hence it was chosen to be the final model .