Finding Best model based on accuracy and hyper tuning the parameters

PROBLEM STATEMENT: **HR - domain: Salary Prediction**

1. SIMPLE LINEAR REGRESSION:

Model accuracy = 0.9740993407213511 (r2_score)

PROBLEM STATEMENT: **Profit prediction for startups based on the investment rate in different departments**

1. MULTIPLE LINEAR REGRESSION

Model accuracy = 0.9358680970046243 (r2_score)

2. SUPPORT VECTOR MACHINE – REGRESSION

BEST ACCURACY MODEL REPORT FOR SVM - REGRESSION
Problem Statement: Profit prediction for startups based on the investment rate in different departments

SL. NO	KERNEL TYPE	MODEL ACCURACY							Remark
		Before Standardization	After Standardization	Regularization Parameter (C parameter in SVM) - (after standardization)					
				C=0.10	C=100	C=1000	C=2000	C=3000	
1	linear	0.87425742	-0.157242386	-0.159614949	0.054876147	0.784912165	0.847721563	0.89592876	comparatively better model
2	rbf	-0.15947378	-0.159758443	-0.159866731	-0.147910466	-0.06276812	0.003968747	0.07670393	poor model
3	poly	-0.147933345	-0.159338555	-0.159824735	-0.107408756	0.268053478	0.526492254	0.68029043	poor model
4	sigmoid	-0.159905544	-0.159469157	-0.159837799	-0.119525566	0.149688096	0.415822569	0.59617543	poor model
5	precomputed	Precomputed kernel type is not suitable for this dataset Reason: Precomputed matrix must be a square matrix. Our Input is a 40x5 matrix.							

Best model:

Linear kernel type with accuracy (r2_score) : 0.895928760783282

3. DECISION TREE – REGRESSION

MODEL ACCURACY BEFORE HYPER TUNING PARAMETERS = 0.9103968927441213 (r2_score)

BEST ACCURACY MODEL REPORT FOR DECISION TREE - REGRESSION Problem Statement: Profit prediction for startups based on the investment rate in different departments					
SL.NO	**criterion**	**splitter**	**max_features**	**model accuracy**	**Remark**
1	squared_error	best	None	0.961681355	**BEST Model**
2	friedman_mse	best	None	0.958174567	good model
3	absolute_error	best	None	0.960309408	good model
4	poisson	best	None	0.961681355	good model
5	squared_error	best	sqrt	0.737630797	poor model
6	friedman_mse	best	sqrt	0.550952689	poor model
7	absolute_error	best	sqrt	0.749180879	poor model
8	poisson	best	sqrt	0.735414346	poor model
9	squared_error	best	log2	-0.274055276	poor model
10	friedman_mse	best	log2	0.393389135	poor model
11	absolute_error	best	log2	0.111477846	poor model
12	poisson	best	log2	0.064075335	poor model
13	squared_error	random	None	0.773375319	poor model
14	friedman_mse	random	None	0.941171288	good model
15	absolute_error	random	None	0.951210159	good model
16	poisson	random	None	0.957416075	good model
17	squared_error	random	sqrt	0.624568348	poor model
18	friedman_mse	random	sqrt	0.236486068	poor model
19	absolute_error	random	sqrt	0.845083542	poor model
20	poisson	random	sqrt	0.888024182	poor model
21	squared_error	random	log2	0.67268384	poor model
22	friedman_mse	random	log2	0.681635312	poor model
23	absolute_error	random	log2	0.019468788	poor model
24	poisson	random	log2	-0.070106056	poor model

Best model:

squared_error criterion with accuracy (r2_score) : 0.961681355225867