To find the R2 value using following the Machine learning Regression algorithm for same dataset to predict the profit:

Dataset : https://github.com/Kuppusamy104/Machine-Learning/blob/main/2.Multiple%20Linear%20Regression/50_Startups.csv

Find out the 3 -Stage of Problem Identification

Stage 1 – Machine Learning

Stage 2 – Supervised Learning

Stage 3 – Regression

1.Multiple Linear Regression r2 Score = 0.9358

2.Support Vector Machine:

S.No	Hyper Parameter	Linear R2 score	RBF(Non linear) R2Score	POLY R2 Score	SIGMOID R2 Score
1	C 1.0	-0.05569	-0.05741	-0.05710	-0.05720
2	C 10	-0.03964	-0.05680	-0.0536	-0.05471
3	C 100	0.10646	-0.05072	-0.01980	-0.03045
4	**C 1000**	**0.78028**	0.00676	0.26616	0.18506

The **R^2 value is 0.7865** for the **SVM algorithm** using the hyper parameter settings: **C= 100** and **kernel=linear**.

This indicates that approximately **78.65%** of the variance in the target variable is explained by the model with these settings.

3.Decision Tree

S.No	Criterion	Max Features	Splitter	R2 Score
1	squared_error	None	best	r2_score [33]: 0.9246310243504303 [ ]: default

2	squared_error	sqrt	best	[53]: r2_score [53]: 0.9025277081984453
3	squared_error	log2	best	[58]: r2_score [58]: -0.9237400598856129
4	**squared_error**	**None**	**random**	[70]: r2_score [70]: 0.9445085531896364
5	squared_error	sqrt	random	[82]: r2_score [82]: -0.5134892769861088
6	squared_error	log2	random	[100]: r2_score [100]: -0.8779930315386868
7	friedman_mse	None	best	[112]: r2_score [112]: 0.8930106646150924
8	friedman_mse	sqrt	best	[124]: r2_score [124]: 0.9147606703357929
9	friedman_mse	log2	best	[136]: r2_score [136]: 0.2663948195021004
10	friedman_mse	None	random	[148]: r2_score [148]: 0.908355349840323
11	friedman_mse	sqrt	random	[160]: r2_score [160]: 0.7131452576799653
12	friedman_mse	log2	random	[172]: r2_score [172]: -0.020012438623537854
13	absolute_error	None	best	[186]: r2_score [186]: 0.9266557112595282
14	absolute_error	sqrt	best	[198]: r2_score [198]: 0.9318597326088528
15	absolute_error	log2	best	[212]: r2_score [212]: 0.7347476074594279
16	absolute_error	None	random	[226]: r2_score [226]: 0.7867381562052261

17	absolute_error	sqrt	random	[242]: r2_score [242]: 0.21795801898313838
18	absolute_error	log2	random	[260]: r2_score [260]: 0.8801692305313137
19	poisson	None	best	[274]: r2_score [274]: 0.9278753336364314 [45]: r2_score [45]: 0.9194670101901314
20	poisson	sqrt	best	[288]: r2_score [288]: 0.4157704680673042 [312]: r2_score [312]: 0.47299414406543616 [325]: r2_score [325]: 0.7864413524347799 [33]: r2_score [33]: 0.5300829327048621
21	poisson	log2	best	57]: r2_score 57]: 0.847311358679689
22	poisson	None	random	[77]: r2_score [77]: 0.6096480036257195
23	poisson	sqrt	random	[91]: r2_score [91]: 0.7908010377802845
24	poisson	log2	random	105]: r2_score 105]: 0.6606748234063025 [123]: r2_score [123]: 0.12869969460543706

The **R^2 value is 0.949508** for the **Decision tree regression algorithm** using the parameter settings: **criterion="squared_error" ,** default **max_features =”None”, splitter="random" .**

This indicates that approximately **94.95%** of the variance in the target variable is explained by the model with these settings.

4. Random Forest:

The **R2 value is 0.9446336** for the Random forest regression using the parameter setting : **n_estimators =50,random_state=0**

```
r2_score
```

```
0.944633639431341
```

```
import pickle
filename="Finalized_model_RandomForest.sav"
pickle.dump(regressor,open(filename,'wb'))
```

Conclusion : Created the models using different algorithm (Multiple Linear Regression , Support Vect or Machine, Decision Tree, Random Forest:)to predict the profit using the same data set to find the b est model.

This indicates that approximately **94.46%** of the variance in the target variable is explained by the model with these settings in Random forest and we have saved this as best model to predict the profit