**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**](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 **R² 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**

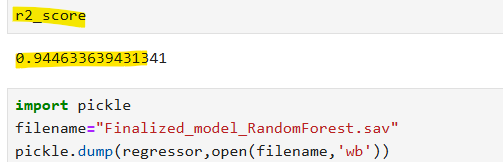
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S.No** | **Criterion** | **Max Features** | **Splitter** | **R2 Score** |
| 1 | squared\_error | None | best | default |
| 2 | squared\_error | sqrt | best |  |
| 3 | squared\_error | log2 | best |  |
| 4 | **squared\_error** | **None** | **random** |  |
| 5 | squared\_error | sqrt | random |  |
| 6 | squared\_error | log2 | random |  |
| 7 | friedman\_mse | None | best |  |
| 8 | friedman\_mse | sqrt | best |  |
| 9 | friedman\_mse | log2 | best |  |
| 10 | friedman\_mse | None | random |  |
| 11 | friedman\_mse | sqrt | random |  |
| 12 | friedman\_mse | log2 | random |  |
| 13 | absolute\_error | None | best |  |
| 14 | absolute\_error | sqrt | best |  |
| 15 | absolute\_error | log2 | best |  |
| 16 | absolute\_error | None | random |  |
| 17 | absolute\_error | sqrt | random |  |
| 18 | absolute\_error | log2 | random |  |
| 19 | poisson | None | best |  |
| 20 | poisson | sqrt | best |  |
| 21 | poisson | log2 | best |  |
| 22 | poisson | None | random |  |
| 23 | poisson | sqrt | random |  |
| 24 | poisson | log2 | random |  |

The **R² 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**

****

**Conclusion** : Created the models using different algorithm (Multiple Linear Regression , Support Vector Machine, Decision Tree, Random Forest:)to predict the profit using the same data set to find the best 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