Assignment-Regression Algorithm

Problem Statement :

➔ Three stage of problem

Machine learning
Supervised
Regressor

➔ Info dataset

The number of is 1338 and columns is 6
Age , children are integer variables
Sex, smoker is object
bmi, charges are float variable

➔ Pre-processing

In this project pre-processing the sex and smoker they are nominal data so, converting the one hot encoding.

➔ Model

I developing the model in **Random Forest** algorithm

➔ Research values

Linear Regression

	params	mean_test_score
0	{'n_jobs': -1}	0.790884
1	{'n_jobs': -2}	0.785639
2	{'n_jobs': -3}	0.788081
3	{'n_jobs': -4}	0.789506
4	{'n_jobs': -5}	0.781319

Decision Tree

	param_criterion	param_splitter	param_max_depth	param_min_samples_split	param_max_features	param_min_impurity_decrease	mean_test_score
0	mse	random	8	2	sqrt	0.7	0.570878
1	mse	random	5	15	sqrt	0.2	0.434473
2	friedman_mse	best	5	15	auto	0.2	0.829773
3	mae	random	8	15	sqrt	0.8	0.297230
4	friedman_mse	random	15	8	auto	0.5	0.779399
5	mae	random	5	8	sqrt	0.2	0.547767
6	mse	best	8	15	sqrt	0.5	0.582035
7	mse	best	None	15	sqrt	0.3	0.638332
8	friedman_mse	random	None	15	sqrt	0.1	0.557748
9	mse	random	None	8	auto	0.7	0.750375

Random Forest

	param_n_jobs	param_n_estimators	param_min_impurity_decrease	param_max_features	param_max_depth	param_criterion	mean_test_score
0	-1	50	0.5	sqrt	15	mse	0.807952
1	-2	500	0.8	auto	None	mae	0.837993
2	-5	1000	0.7	auto	10	mae	0.663803
3	-3	200	0.2	auto	10	mae	0.810769
4	-3	500	0.8	sqrt	None	mse	0.794067
5	-1	500	0.3	sqrt	10	friedman_mse	0.053960
6	-2	100	0.3	sqrt	5	mae	0.428480
7	-5	50	0.8	log2	5	mae	0.773558
8	-5	50	0.1	log2	10	mse	0.782158
9	-1	1000	0.4	sqrt	5	friedman_mse	0.735926

Separate Vector Machine Regressor

s.no	Hyper	Linear	RBF	Sigmoid	poly
1	C=10	-0.0016	-0.0875	-0.0897	-0.0323
2	C=100	0.5433	-0.0746	-0.0897	0.5622
3	C=500	0.6270	-0.0495	-0.0897	0.7690
4	C=1000	0.6340	-0.0408	-0.0897	0.8190

➔ Why choose this algorithm

I choose the random forest algorithm why means that only gives more accuracy

And I do this r^2 score they give 0.8922 percentage