

Problem Statement

Predict the insurance charges based on the input 'age', 'bmi', 'children', 'sex_male', 'smoker_yes'

MultipleLinearRegression

<i>fit_intercept</i>	<i>copy_X</i>	<i>n_jobs</i>	<i>positive</i>	<i>r2_score</i>
True	<i>True</i>	None	False	0.710092369
False	True	None	False	0.6061988575
False	False	None	False	0.6061988
False	False	-1	False	0.6061
False	False	10000	False	0.60619
True	True	-10	True	0.71008
True	True	1	True	0.71008

The best *r2_score* comes for default parameter values

SupportVectorMachine Regression

kernel	rbf	linear	poly	sigmoid	rbf	linear
<i>degree</i>	3	3	3	3	3	3
<i>gamma</i>	scale	scale	scale	scale	auto	auto
<i>coef0</i>	0.0	0.0	0.0	0.0	0.0	0.0
<i>tol</i>	<i>1e-3</i>	<i>1e-3</i>	<i>1e-3</i>	<i>1e-3</i>	<i>1e-3</i>	<i>1e-3</i>
<i>C</i>	1.0	1.0	1.0	1.0	1.0	1.0
<i>epsilon</i>	0.1	0.1	0.1	0.1	0.1	0.1
<i>shrinki</i>	True	True	True	True	True	True

<i>ng</i>						
<i>cache_size</i>	200	200	200	200	200	200
<i>verbose</i>	False	False	False	False	False	False
<i>max_iter</i>	1	1	1	1	1	1
<i>r_score</i>	-34167.663	-11.39	-174.883	-3443455.3	-34167.663	-11.3966

DecisionTreeRegressor

Criterion	<i>squared_error</i>	<i>poisson</i>	<i>friedman_mse</i>	<i>absolute_error</i>
<i>splitter</i>	<i>best</i>	<i>best</i>	<i>random</i>	<i>random</i>
<i>max_depth</i>	<i>None</i>	<i>20</i>	<i>None</i>	<i>None</i>
<i>min_samples_split</i>	2	20	2	2
<i>min_samples_leaf</i>	1	10	1	1
<i>min_weight_fraction_leaf</i>	0.0	0.0	0.0	0.0
<i>max_features</i>	None	None	None	None
<i>random_state</i>	None	None	None	None
<i>max_leaf_nodes</i>	None	100	None	None
<i>min_impurity_decrease</i>	0.0	0.0	0.0	0.0
<i>ccp_alpha</i>	0.0	0.0	0.0	0.0

<i>monotonic_cst</i>	None	[1,1,1,1,1]	None	None
<i>r2_score</i>	0.71862	0.8753	0.7222	0.7499

Random Forest Regressor

<i>n_estimators</i>	100	100	1000	100	1000
<i>criterion</i>	<i>squared_error</i>	<i>absolute_error</i>	<i>poisson</i>	<i>friedman_mse</i>	<i>squared_error</i>
<i>max_depth</i>	None	None	1000	None	1000
<i>min_samples_split</i>	2	2	20	2	20
<i>min_samples_leaf</i>	1	1	3	1	10
<i>min_weight_fraction_leaf</i>	0.0	0.0	0.0	0.0	0.0
<i>max_features</i>	1.0	1.0	1.0	1.0	1.0
<i>max_leaf_nodes</i>	None	None	None	None	None
<i>min_impurity_decrease</i>	0.0	0.0	0.0	0.0	0.0
<i>bootstrap</i>	True	True	True	True	True
<i>oob_score</i>	False	False	False	False	False
<i>n_jobs</i>	None	None	None	None	None
<i>random_state</i>	None	None	None	None	None

<i>verbose</i>	0	0	0	0	0
<i>warm_start</i>	False	False	False	False	False
<i>ccp_alpha</i>	0.0	0.0	0.0	0.0	0.0
<i>max_samples</i>	None	None	None	None	None
<i>monotonic_cst</i>	None	None	None	None	None
<i>r2_score</i>	0.8461	0.8467	0.8767	0.8461	0.8794

The best score model is Random Forest Regressor with 0.8794