1.**Project**

Name: Insurance Charges Prediction.
Developer: Aravind

2. **Dataset**

NamePath: “Dataset/insurance_pre.csv”.
Type: csv.
Size: 1337 rows, 6 columns.

3.**Preprocessing**

One hot encoding for nominal data columns, gender_sex and is_smoker was performed.

4.**Models tested**

SMR, SVM, DT, RT.

5.**Research Data**

Multiple Linear Regression (MLR)	**R2 score** = 0.78

Support Vector Machine (SVM)				
Hyper Parameter	Rbf (r2)	Linear (r2)	Poly (r2)	Sigmoid (r2)
C=10	-0.084	-0.001	-0.094	-0.09
C=100	-0.12	0.54	-0.099	-0.11
C=1000	-0.12	0.63	0.056	-1.71
C=10000	-0.01	0.74	0.35	-124.10
C=100000	0.55	0.74	0.77	-1167.41
C=1000000	0.83	0.74	0.86	-115
C=10000000	0.87	0.74	0.86	-115

Decision Tree (DT)		
Criterion	Splitter best (r2)	Splitter Rand (r2)
Squared Error	0.68	0.66
Friedman mse	0.73	0.68
Absolute error	0.74	0.72
Poisson	0.67	0.74

Random Forest (RT)				
N estimator	Squared error (r2)	Friedman mse (r2)	Absolute error (r2)	Poisson (r2)
C=10	0.82	0.82	0.84	0.82
C=100	0.82	0.82	0.84	0.82
C=1000	0.85	0.85	0.85	0.84
C=10000	0.85	0.85	0.85	0.85

Result
The ML model SVM with a 0.87 r2_score has better accuracy when compared with the other algorithms tested shown above.