INSURANCE CHARGE PREDICTION

1.PROBLEM STATEMENT

- ❖ STAGES

A. MACHINE LEARNING
B. SUPERVISED LEARNING
C. REGRESSION

2.DATASET

A. ROWS-1338
B. COLUMNS-6
C. SEX AND COLUMS ARE CATEGORICAL VALUE.

3.PREPROCESSING METHOD

THE CATEGORICAL VALUE IS CONVERTED BY NUMERICAL VALUE BY USING NOMINAL DATA.THEN USING get_dummies IN PANDAS LIBRARY.

INPUT/OUTPUT SPLITATION PROCESS

TEST/TRAINING SET CREATED.

4.MODEL:

- ➢ MULTIPLE LINEAR REGRESSION:

R^2 VALUE=0.798

- ➢ *SUPPORT VECTOR MACHINE*

kernel	**R2 value**
Poly	-0.0961
sigmoid	-0.0910

Linear	-0.029
Precomputed	-
RBF,C=3000	0.995
Poly, C=3000	0.6370
Sigmoid,C=3000	0..5913

- *DECISION TREE*

Criterion	**Splitter**	**R2value**
Mae	Best,SQRT	0.7652
Friedman_mse	best	0.6091
Friedman_mse	random	0.417
Mae	random	0.7650
Mae	Random,sqrt	-0.068

- RANDOM FOREST

N_Estimators	**Max_Features**	**R2 value**
100	sqrt	0.816
10	Log2	0.647
100	Log2	0.216
50	sqrt	0.5667
100	auto	0.7656

100	Auto,random_state=0	0.8660

5.COMPARING ALL THE MODELS SVM IS A BEST MODEL AND HIGHEST VALUE 0.99%.