

PREDICTION OF INSURANCE CHARGES BASED ON GIVEN INPUT PARAMETERS

GIVEN DATA SET: (SAMPLE)

age,sex,bmi,children,smoker,charges
19,female,27.9,0,yes,16884.924
18,male,33.77,1,no,1725.5523
28,male,33.0,3,no,4449.462
33,male,22.705,0,no,21984.47061
32,male,28.88,0,no,3866.8552

PHASE-1(MODEL CREATION)

Process 1:

Data Collection:

- =>Collected the data from the given data set
- =>The dataset contains 3800 data

Process 2:

Data Preprocessing:

- =>The given data set contains categorical data columns
- =>The data are pre-processed to convert all the column data into numerical values
- =>The columns are dropped to avoid repetition of same information.

Process 3:

Splitting the independent and dependent data:

Independent data:[INPUT]

Age	BMI	CHILDREN	SEX_MALE	SMOKER_YES
19	27.900	0	0	1
18	33.770	1	1	0
28	33.000	3	1	0
33	22.705	0	1	0
32	28.880	0	1	0

Dependent data: [OUTPUT LABEL]

Charges
16884.92400
1725.55230
4449.46200
21984.47061
3866.85520

Process 4:

Test and Train data Split:

- =>The input and output data is splited as training dataset and testing dataset in 70 and 30 percentage respectively.

Process 5:*Model Creation:*

=>The model is created by training with the train data set

Process 6:*Model Evaluation:*

=>The model is evaluated using the test dataset

Process 7:*Saving the best model*

=>After analysing the best model is saved

PHASE-2(DEPLOYMENT):**Process 1:**

=>The saved model is loaded

Process 2:

=>The input data is given

Process 3:

=>The output is predicted by the trained model