

Methodology-

My first step was to clean and prepare the data for analysis. I went through different steps of data cleaning. First, I checked the data types focusing on numerical and categorical to simplify the correlation's computation and visualization. Second, I tried to identify any missing values existing in our data set. Last, I researched each column/feature's statistical summary to detect any problem like outliers and abnormal distributions.

I cleaned the data then I did feature selection then I applied different machine learning algorithms and random forest classifier gave the best results. Before that we hypertuned our model and atlast we did hypertuning with a random forest classifier.

