



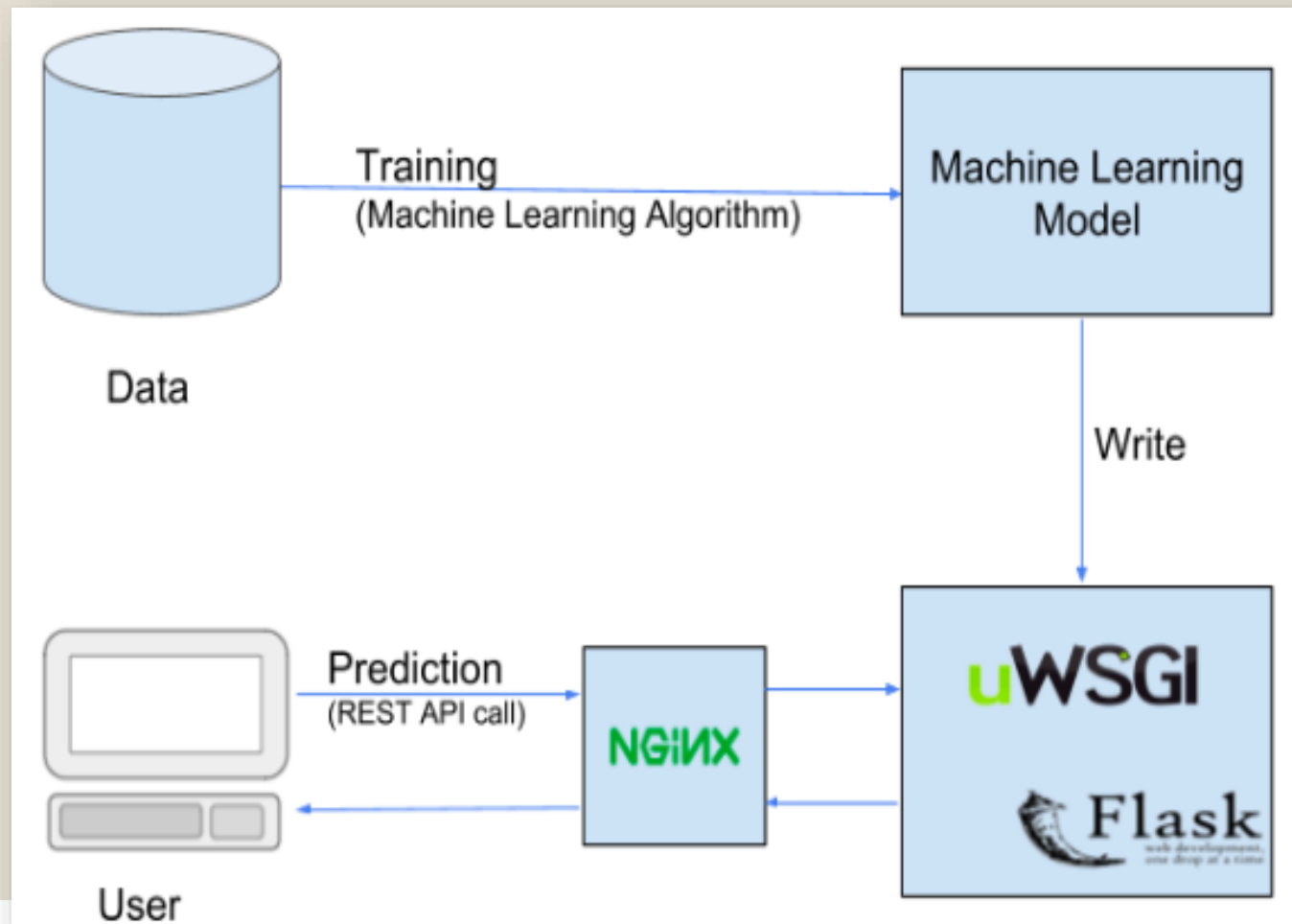
ADULT CENSUS INCOME PREDICTION

Umang Tank

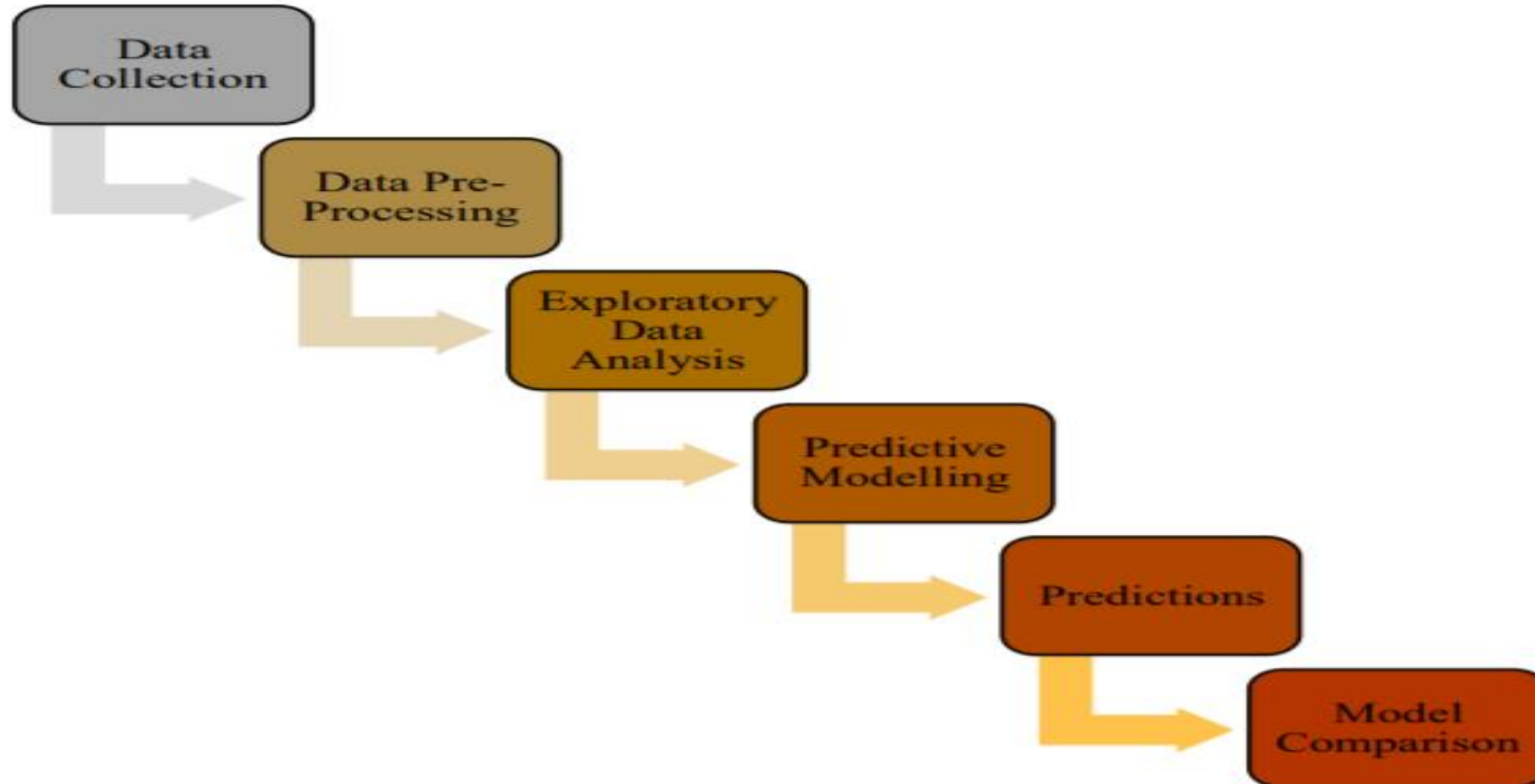
Objective

- The Goal is to predict whether a person has an income of more than 50K a year or not. This is basically a binary classification problem where a person is classified into the >50K group or ≤50K group.

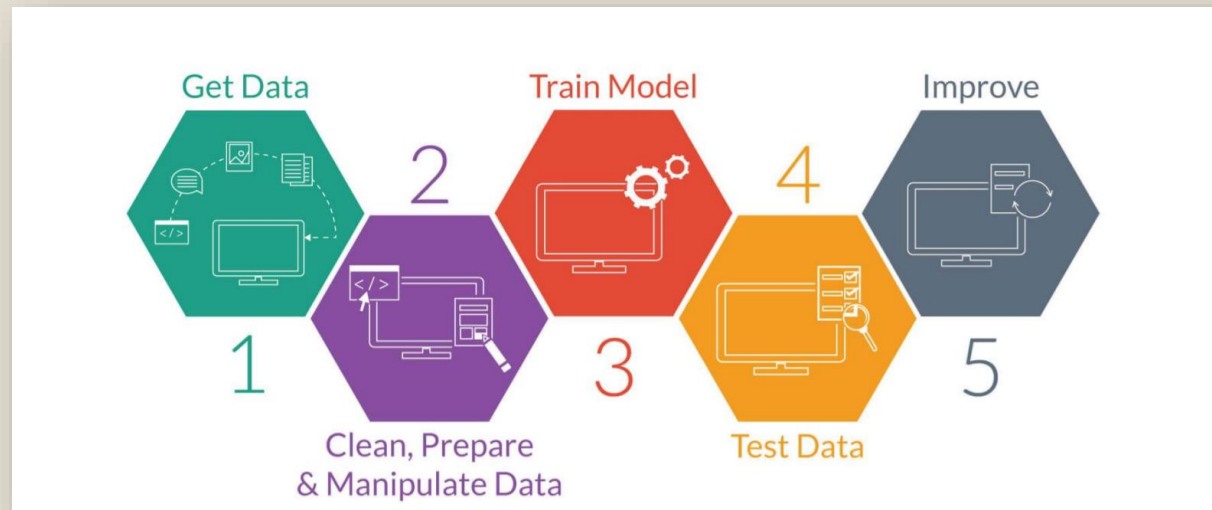
Architecture



Data Analysis step



Steps to Predictive Modelling



Training

- Once training data is in a suitable form to feed to the model, the training and testing phase of the model can proceed. During the training phase, the primary concern is the model selection. This involves choosing the best modelling approach for the task, or the best parameter settings for a given model. In fact, the term model selection often refers to both of these processes, as, in many cases, various models were tried first and best performing model (with the best performing parameter settings for each model) was selected.

Predictions

- The model was used to predict the income.
- The model used the relation between the features and the label to predict the amount.
- Accuracy defines the accuracy score and f1-score of the predicted income with binary classification.
- The model predicted the accuracy of model by using different algorithms, different features and different train test split size.
- The size of the data used for training of data has a huge impact on the accuracy of data. The larger the train size, the better is the accuracy.

Thank you

Umang Tank