

### 1) How cleaning/EDA was performed?

Ans:

- a) Checking null values and null value percentages in each column
- b) Ignoring columns having null values more than 30%
- c) Replacing null values of object column type by mode of the variable
- d) Replacing the numeric type variable by median/mean value. We have imputed median to avoid discrepancy in the dataset.
- e) We performed univariate analysis to check the variables and its value

### 2) Your independent and dependent feature

Ans:

- a) Dependent variable = 'Gender'
- b) Independent variable = Apart from gender all variables are considered as independent variable but since we have dropped many variables due to high null value percentages, we have listed down the remaining variables

### 3) Why and how selection/engineering/scaling were performed

Ans:

- a) We have import MinMaxScaler() scaler function from library sklearn.preprocessing.
- b) Min Max Scaler helps to scale data in the range of 0 to 1.
- c) It basically helps to normalise the data within a particular range.

### 4) Which activation function was chosen and why?

- a) The activation function used is "ReLU".
- b) ReLU is linear identity for all positive values, and zero for all negative values.
- c) It's cheap to compute as there is no complicated math. The model can therefore take less time to train or run.
- d) It doesn't have the vanishing gradient problem suffered by other activation functions.
- e) Since 'ReLU' is zero for all negative inputs, it's likely for any given unit to not activate at all.

### 5) Which optimizer was chosen and why?

- a) The optimiser that is used is adam.
- b) Adam combines the best properties of the AdaGrad algorithms to provide an optimization algorithm that can handle sparse gradients on noisy problems.
- c) Adam is relatively easy to configure where the default configuration parameters do well on most problems.

## 6) Which neural network and why? Describe how your neural structuring?

- a) Artificial neural network using sklearn is used in this mini project. Artificial neural networks (ANNs), simply called neural networks, are computing systems vaguely inspired by the biological neural.
- b) ANN uses the processing of the brain as a basis to develop algorithms that can be used to train complex patterns and prediction problems.
- c) Step to build Neural network and check accuracy
  - i. Define a sequential model.
  - ii. Add a dense layer.
  - iii. Compile the model with an optimizer and loss function.
  - iv. Fit the model to the train dataset.
  - v. Test the built model in test dataset.
  - vi. Check predictions
  - vii. Based on predicted value and actual value calculate the confusion matrix and check accuracy.