

Short-Answer Portion

1. According to the definition of F1- score, the higher the F1-score, the less (FN + FP). So higher score means better precision and recall.

For the logistic model with 0.6 F1-core and the neural network model with 0.63 F1-score, I will recommend the bank the neural network model.

$$F1 = \frac{2TP}{2TP + FN + FP}$$

2. In general way, I would suggest to use the function of *f_egression* from *sklearn.feature_selection* to evaluate the importance of features. This function calculates the correlation between each feature and the target.

Also, there are some specific approaches for different model (linear regression, random forest, neural network).

For linear regression, I would suggest compare coefficient if the model is regularized and dataset is normalized. Large values of w_j signify higher importance of the j th feature in the prediction.

For Random forest model, if the model is implemented using sklearn, I would suggest to check the *feature_importances_* attribute of the model. The higher the value, the more important the feature.

For Neural network regressor, I would suggest to check the weight matrix of first layer. For neural network, if one feature is important, then it will have a highly weighted connection to the first layer of neural net model.

$$z^1 = w^1 * X + b$$

According the definition of weight matrix, I will sum the absolute value in i th column which represents the weight connect from i th feature to first layer.