| **Criteria** | **Ratings** | **Points** |
| --- | --- | --- |
| **Feature Elimination** | 1) Feature selection is performed properly. Identified and eliminated the features that are not relevant from domain and functionality point of view. The columns ‘Row number’, ‘Customer ID’ and ‘Sur name’ are dropped from the dataset as part of feature selection, good. | 5/5 |
| **Bivariate** | 2) This question is attempted. EDA is not performed on the dataset. Univariate analysis and bi-variate analysis is missing. In multi variate analysis pairplots are not generated and correlation matrix and its heat map is also missing. (5 marks deducted) | 0/5 |
| **Data Split** | 3) The dependent/Target feature and the independents features are identified and separated properly. Then the given dataset is split into separate train and test dataset using sklearns train\_test\_split function, Good. | 5/5 |
| **Normalization** | 4) In data transformation step, the categorical/discrete features which has string data type (Geography, Gender) has been converted into numeric data type using one-hot encoding/ Label encoding, good. Further to make the dataset scale invariant the dataset is normalized using Minmax scaler, good. | 10/10 |
| **Modelling** | 5) A neural network model is built with an architecture of one input layer, 1 hidden layer and 1 output layer with an activation function of ‘sigmoid’. The optimizer is set as ‘adam’ and loss function as ‘binary cross entropy’ and epoch is set as 60, Good. But further optimization of the model is not performed. The performance of the model depends on various factors like Model’s architecture type, Number of layers, Number of Neurons in a layer, Regularization parameters, learning rate, back propagation technique, dropout rate, weight sharing technique etcs. Further iterating the model by varying these parameters might result in better performance of the existing model. Sbut this step is not performed (10 marks deducted) | 10/20 |
| **Prediction at 0.5 Threshold** | 6) Binary Predictions are obtained from the optimized neural network model by setting the prediction threshold at 0.5, Good. | 10/10 |
| **Model Performance Evaluation** | 7) The model evaluation metrics is generated by printing models accuracy score and confusion matrix along with classification report which involves precision, Recall and F1 score for both the classes of target variable, good. | 5/5 |
|  | Points | 45/60 |

**Comments:**

Overall, your approach to the problem is ok and you have attempted few parts of the assignment with right answer. Few improvements are required in few sections. Areas of Focus: The 2nd sub question is not attempted. EDA is not performed on the dataset. Univariate analysis and bi-variate analysis are missing. In multi variate analysis pair plots are not generated and correlation matrix and its heat map are also missing. Further optimization of the model is not performed. The performance of the model depends on various factors like Model’s architecture type, Number of layers, Number of Neurons in a layer, Regularization parameters, learning rate, back propagation technique, dropout rate, weight sharing technique etc. Further iterating the model by varying these parameters might result in better performance of the existing model. but this step is not performed.

Hemanth K V, Mar 21, 2021, 10:51 PM