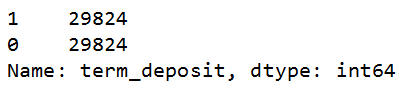
**Model Building**

# **Introduction**:

Model building in data analytics is aimed at achieving not only high accuracy on the training data but also the ability to generalize and perform well on new, unseen data. Therefore, the focus is on creating a model that can capture the underlying patterns and relationships in the data, rather than simply memorizing the training data.

**Steps involved in Model building:**

* Identifying the output and target variable.
* Here identified target variable is booking\_status and the rest are the input variables.
* Splitting the data into training and testing.
* By splitting the input and output variables into train and test we got 20800 train points and 5200 test points.
* Dividing categorical and numerical features.
* Pre-processing the train\_data.
* Pre-processing the test\_data.
* Scaling the numerical features of test\_data using Standardization.
* Scaling the categorical features of test\_data using OneHotEncoding.
* Concatenating the numerical and categorical features.
* Balancing the data using SMOTE (Synthetic Minority Oversampling Technique) to rectify the imbalance in the given data. This technique is used after data transformation technique.
* After implementing SMOTE our data was successfully balanced resulting like this



* Building the logical structure using Keras tuner a powerful Hyperparameter optimization library helps in shaping the architecture of deep learning model.
* **Creating a function.**
* **Calling the function using tuner() for training and testing of the data.**

Keras tuner library is used for searching the optimal hyper-parameters for Deep learning models. Keras tune is a great way to check for different numbers of combinations of kernel size, filters, and neurons in each layer.

* The hyperparameters that yielded the best-performing model were discovered,

including model architecture with six hidden layers, specific neuron counts in each layer.

* 'he\_uniform' weight initialization.
* 'rmsprop' optimizer.

These hyperparameters define the final model's configuration.

The final model architecture is composed of six dense hidden layers, each followed by batch normalization and dropout layers to enhance model generalization and prevent overfitting.

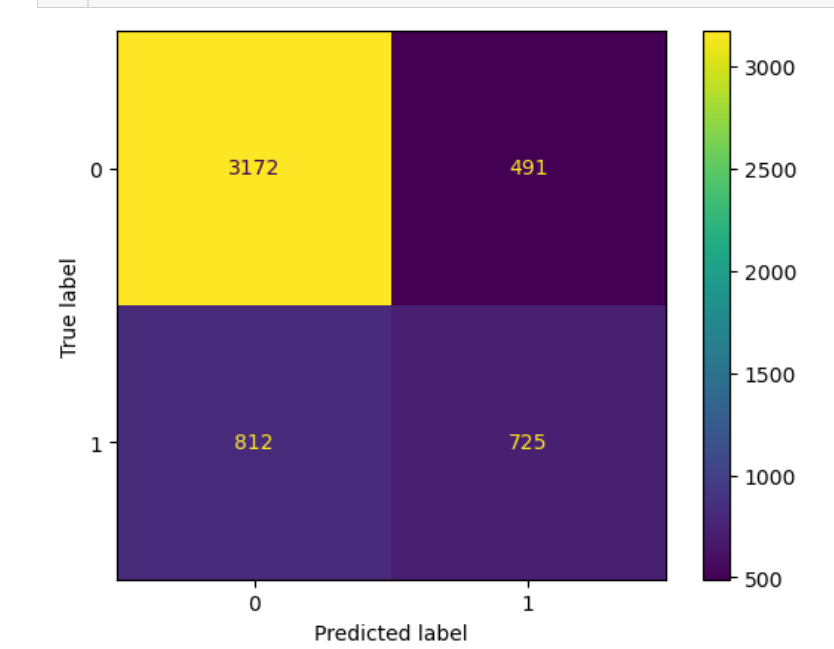
The final model was trained on the preprocessed dataset with 20 as the batch size

and 20 training epochs. During training, a validation split of 15% was used for

performance monitoring.

The model achieved an accuracy of approximately 74% on the validation dataset.

The model's performance was further evaluated using a confusion matrix. The goal was to maximize the diagonal elements of the matrix, indicating correct predictions.

****

* **A high diagonal element count signifies the model's capability to make accurate classifications.**