First we import all the necessary modules and libraries.

Then we read dataframe. This data is of a telecommunication company predicting the reason of the leaving customers. We also convert the dependent variable column into integer value.

We set various columns as dependent variables and the last column as Y and convert them into arrays for computation using np.asarray().

Now we preprocess the data using StandardScalar() and standardize the data which uses the formulae x-mean/variance.

Then we split the data into test and train with a ratio of 80:20.

We then perform logistic regression using inbuilt function where c is the inverse of regularisation strength. (Regularization is the penalty applied to increase the weights and reduce error). Solver parameters tell us the algorithm we would like to use for applying logistic regression. We use liblinear algorithm to perform linear classification problems on a smaller dataset. Simultaneously we fit the X and Y data to this function. The number of iterations is 100 by default.

After that we calculate the probability value for test data. As we can 2 values in each row. These are probability values for one data for class 0 and class 1. This data point will belong to class 0 if the value of the first column is more than 0.5 else it will belong to class 1.

Now if we predict the value we can see the class 0 or 1.

We can calculate the accuracy of this model for this data and hyperparameters using the jaccard index. we can define jaccard as the size of the intersection divided by the size of the union of two label sets. That is the intersection between predicted and the actual values in our case. The accuracy here is 75 percent.