**TITANIC ASSIGNMENT**

**Process:**

* Import the Train.csv dataset.
* Clean the data by removing unnecessary columns by transforming categorical variables into numerical variables(Numerical to binomial), selecting attributes, handling missing values(Replace missing values), and set label using set role operator.
* Train a machine learning model on the training set using the "Random forest" operator and evaluate its performance using the "Performance(Classification)" operator through apply model operator, Here I have selected Random Forest because it gives a high accuracy comparing any other classification operator.
* Test the model on the train set and generate predictions for the "Survived" column and do the same for test set by importing test.csv and merging it with gender submission data file using “Join” operator.
* Merge the predictions with the "PassengerId" column from the test data and the "PassengerId " column from the gender submission file.

Diagram

Description automatically generated with medium confidence

Table

Description automatically generated

