Classification Model to Identify Employee Attrition

To Identify Employee Attrition by using Classification Model (Random Forest Classifier)

**STEPS:**

Step 1: Import Libraries

*# Import the numerical algebra libs*

import pandas as pd

import numpy as np

*# Import visualization libs*

import seaborn as sns

import matplotlib.pyplot as plt

Step 2: Import and Read data

Data Source:

'https://github.com/ybifoundation/Dataset/raw/main/EmployeeAttrition.csv’

Step 3: Study Data: After importing the file we have to run 4 functions to understand data

1. X.head() : The head() method returns a specified number of rows, string from the top.
2. X.info (): The info () method prints information about the DataFrame
3. X.describe(): The describe() method returns description of the data in the DataFrame.
4. X.isnull(). sum (): It checks missing values. If there are certain numbers of null values present in each feature, then we have to impute any other value to fill null values

Step 4: Divide data into Y{Target} & X{Features}

Step 5: Split the data into Test and Train

Eg: from sklearn.model\_selection import train\_test\_split

Step6: Select, train(fit) & predict the Model

Step 7: Error Calculation (Model Accuracy)

Step 8: Future Prediction

**CONCLUSION:**

Analyzed the Employee’s Attrition and used Machine Learning to identify the Employee Attrition. We have used Random Forest Classifier and its variations, to make predictions and compared their performance.