McCurry Consultancy – Attrition Analysis

Background:McCurry Consultancy is an MNC that has thousands of employees spread across the globe. The  
company believes in hiring the best talent available and retaining them for as long as possible.  
A huge number of resources is spent on retaining existing employees through various  
initiatives. The Head of People Operations wants to bring down the cost of retaining  
employees. For this, he proposes limiting the incentives to only those employees who are at  
risk of attrition. As a recently hired Data Scientist in the People Operations Department, you  
have been asked to identify patterns in characteristics of employees who leave the organization.  
Also, you have to use this information to predict if an employee is at risk of attrition. This  
information will be used to target them with incentives and in turn curb attrition.

Objective:To predict if an employee is going to attrite or not.

Data:The data contains demographic details, work-related metrics and attrition flag.

Deliverables:

1. Import the datasets and libraries and Check: datatype, statistical summary, shape, null  
values or incorrect imputations.

2. EDA:  
• Univariate analysis – description of the independent attributes. Strategies to address the  
different data challenges such as data noise, outlier’s treatment, and missing values  
treatment.  
• Bi-variate analysis between the predictor variables and target column. Comment on  
your findings in terms of their relationship. Visualize the analysis using boxplots and  
pair plots, histograms, or density curves. Get the data model ready.

3. Split the data into training and test set in the ratio of 70:30 respectively.

4. Use Logistic Regression model to predict whether the employee would leave or not.  
Print all the metrics related for evaluating the model performance.  
5. Give your reasoning on how the model can perform better.  
6. Give Business understanding of your model.