Description

This project requires you to understand what mode of transport employees prefers to commute to their office. The attached dataset Cars.csv, includes employee information about their mode of transport as well as their personal and professional details like age, salary, work exp. We need to predict whether or not an employee will use Car as a mode of transport. Also, which variables are a significant predictor behind this decision?

Following is expected out of the candidate in this assessment.

EDA

- Perform Exploratory Data Analysis on the dataset
- Illustrate the insights based on EDA
- Multicollinearity check and summarization of problem statement for business stakeholders

Data Preparation

- Prepare the data for analysis

Modelling

- Create multiple models and explore how each model perform using appropriate model performance metrics
    - KNN
    - Naive Bayes (is it applicable here? comment and if it is not applicable, how can you build an NB model in this case?)
    - Logistic Regression
- Apply both bagging and boosting modelling procedures to create 2 models and compare its accuracy with the best model of the above step. **Actionable Insights & Recommendations**

- Summarize your findings from the exercise in a concise yet actionable note