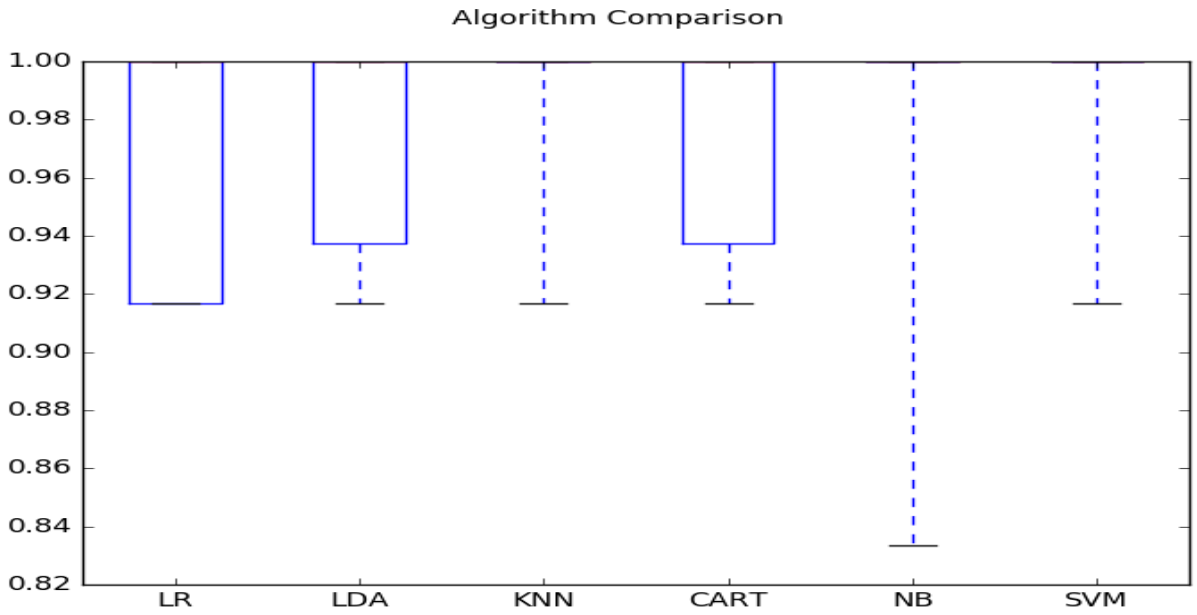
MACHINE LEARNING CASE STUDY

Submission Date: 30th December 2023

- **Problem Definition**
  - **Give a concrete description of your machine learning problem in no more than 50 words**
- **Datasets**
  - **Brief on the datasets used in the project.**
  - **The dataset should consist of at least 1000 samples and at least 10 features. The features should be a mix of categorical and numerical features.**
- **Prepare Data**
  - **Explain the pre-processing done on your Dataset to make it suitable for applying machine learning algorithms.**
  - **Summarization:**
    - **Use statistical methods to understand the data and apply the required methods such as**
      - Statistical summary of all attributes.
      - Breakdown of the data by the class variable.
  - **Data Visualization:**
    - **Visualize the data using various plots like scatterplot, histograms, box plot etc and record your interpretations with varying values**
- **Python packages**
  - **Brief on the python packages used for implementation of Machine learning algorithms pertaining to your project.**

- **Supervised Learning Algorithms**
  - **At least 3 supervised machine learning algorithms are to be used**
    - Brief on the 3 supervised ML algorithms chosen for creating learning model from your dataset.
    - Split your data into training, validation, and testing
    - Use at least 1 cross validation technique to evaluate your ML algorithm
    - Create models of the data and estimate their accuracy on unseen data using the specified ML algorithms.
      - Example: If Logistic regression, SVM, and Bayesian learning are used for classification, create models for different algorithms. Select the best model.
      - Plot various graphs for each of the ML algorithm separately.
      - Plot a comparison graph showing the accuracy comparison of various algorithms on each of your datasets. Comparison can be based on accuracy, f1 measure, precision, recall, etc.
        Example : Following figure shows an instance of average accuracy comparison of various algorithms on a particular dataset.



Prepare a video presentation and upload by 30th December, 2023.