**PHASE-2**

**INNOVATION**

Here we use Support Vector Machine Algorithm to analyse the public transportation data to assess service efficiency. Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems.

**SVM Classification Algorithm:**

STEP-1: **Data Preparation:**

•Collect a labeled dataset with features (independent variables) and corresponding class labels (categories or outcomes). Ensure that the dataset is divided into a training set and a testing set.

STEP-2: **Feature Scaling:**

•Standardize or normalize the features to ensure they have similar scales. This step can help the SVM perform better and converge faster during training.

STEP-3: **Training the SVM:**

•Select a suitable SVM kernel function based on the nature of your data (e.g., linear, polynomial, radial basis function).

•Train the SVM on the training dataset by finding the hyperplane that best separates the data points of different classes while maximizing the margin between them. The training process involves solving the optimization problem associated with the SVM.

STEP-4: **Hyperparameter Tuning:**

•Perform hyperparameter tuning to optimize the SVM's performance. Key hyperparameters include the regularization parameter 'C' (controls the trade-off between maximizing margin and minimizing classification errors) and kernel-specific parameters (e.g., kernel degree, gamma for the RBF kernel).

•Use techniques like cross-validation to find the best hyperparameter values that yield the highest accuracy or other suitable evaluation metrics on the validation set.

STEP-5: **Testing and Evaluation:**

•Apply the trained SVM model to the testing dataset to make predictions on unseen data points.

•Evaluate the model's performance using appropriate classification metrics such as accuracy, precision, recall, F1-score, and the confusion matrix to assess how well it generalizes to new data.

**#USING SVM ALGORITHM:**

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

