

13 cars to produce data

4 nodes to store, train the model, and giving predictions

Neural Network model and random forest

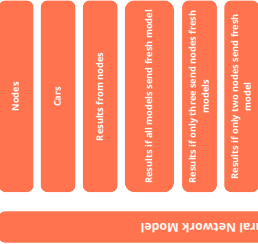
1 central server to receive the models from nodes, make refinement and aggregation, broadcast the general model to nodes

the script produces dataset and uses kafka as a producer

continuously receives data from a kafka topic, predicts whether a vehicle needs charging, updates a machine learning model with new data, and periodically exchanges the model with a server. It also monitors and plots the prediction accuracy over time. Moreover, it loads a pre-trained model and uses it in the beginning.

Load Data, Data Cleaning: Adds a 'needs_charger' column based on a condition. Converts categorical variables into dummy/indicator variables. Removes duplicate rows. Handles outliers by replacing extreme values with medians. **Feature Selection:** Selects specific features based on Recursive Feature Elimination (RFE) with a logistic regression model. Visualizes feature distributions and relationships. **Data Visualization, Random Forest Classifier:** Builds a Random Forest Classifier model. Performs hyperparameter tuning using GridSearchCV. Evaluates the model's accuracy and provides a classification report. Conducts cross-validation to assess model performance. Neural Network Model: Constructs a Neural Network model using TensorFlow/Keras. Compiles the model. Defines callbacks for early stopping and model checkpointing. Trains the model on the training data. Evaluates the model's performance on the test data. **Prediction function, Model Saving, Model Loading, Additional Predictions.** **This ML aims to cover various aspects of data preprocessing, feature selection, model training, evaluation, and prediction for a binary classification task related to determining whether a car needs charging based on its attributes and environmental conditions.**

A server that collects machine learning models and their accuracies from multiple nodes, aggregates these models using a simple average when a threshold is met, and sends the aggregated global model back to the nodes that request it. This setup can be used for federated learning or collaborative model training across multiple nodes.



Node - 1 -

11 cars

at starting 99%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate

Node - 2 -

12 cars

at starting 99%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate

Node - 3 -

not more than 12 (depends on battery charging level)

at starting 99%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate

Node - 4 -

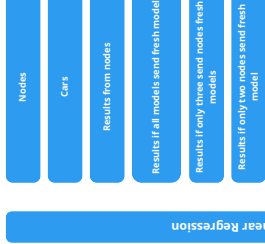
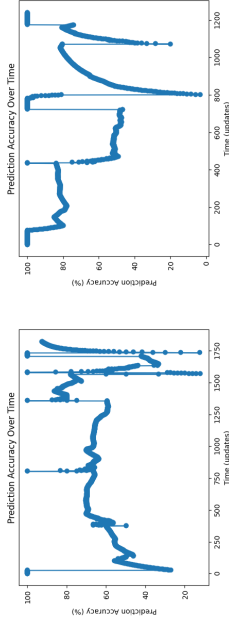
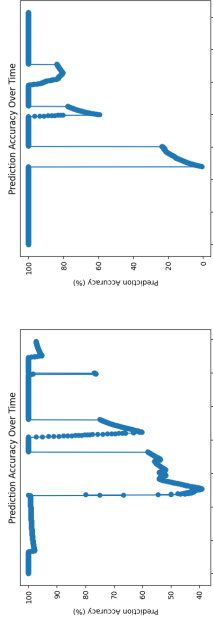
1 certainly connected and 12 not certainly connected (depends on battery charging levels)

at starting 99%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate



Node - 1 -

11 cars

at starting 100%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate

Node - 2 -

12 cars

at starting 100%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate

Node - 3 -

not more than 12 (depends on battery charging level)

at starting 100%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate

Node - 4 -

1 certainly connected and 12 not certainly connected (depends on battery charging levels)

at starting 100%

The general model becomes more accurate, specially, when it drops down by the time because of changing environment conditions

The general model becomes more accurate, but less than all models.

Not accurate

