**(M.HARISH – AU732721106016)**

**PUBLIC TRANSPORTATION OPTIMIZATION**



* Machine learning algorithms can be incorporated to improve arrival time prediction accuracy based on historical data and traffic conditions in a number of ways.
* One approach is to use a supervised learning algorithm, such as random forest or support vector regression. These algorithms learn from historical data to predict the arrival time of a vehicle or person based on a set of features, such as the vehicle's current location, speed, and direction of travel, as well as the current traffic conditions.
* Another approach is to use an unsupervised learning algorithm, such as k-means clustering. This type of algorithm can be used to identify patterns in historical arrival time data, such as different traffic patterns during different times of day or on different days of the week. This information can then be used to develop more accurate arrival time predictions.

**Machine learning algorithms can be used to improve arrival time prediction accuracy:**

* Use a random forest algorithm to predict the arrival time of buses at a particular stop. The algorithm could be trained on historical data, such as the bus's current location, speed, and direction of travel, as well as the current traffic conditions. The algorithm would then be able to predict the bus's arrival time with greater accuracy than a traditional method, such as using a fixed schedule.
* Use a support vector regression algorithm to predict the arrival time of ships at a port. The algorithm could be trained on historical data, such as the ship's current location, speed, and direction of travel, as well as the weather conditions and other factors that may affect the ship's speed. The algorithm would then be able to predict the ship's arrival time with greater accuracy than a traditional method, such as using a fixed schedule.
* In general, machine learning algorithms can be used to improve arrival time prediction accuracy by taking into account a wide range of factors that can affect travel time. This is in contrast to traditional methods, which often rely on fixed schedules or simple formulas.

**Considerations for incorporating machine learning algorithms to improve arrival time prediction accuracy:**

* **Data quality and quantity:** The quality and quantity of the available data are important factors for the success of any machine learning project. For arrival time prediction, this means having access to a large and accurate dataset of historical arrival time data, as well as real-time data on traffic conditions.
* **Feature selection:** It is important to select the right features to train the machine learning algorithm on. The features should be relevant to the prediction task and should be able to capture the factors that affect arrival time.
* **Model evaluation:** Once the machine learning algorithm is trained, it is important to evaluate its performance on a held-out test set. This will help to ensure that the model is not overfitting the training data and is able to generalize to new data.
* **Model monitoring:** Once the machine learning model is deployed, it is important to monitor its performance over time. This is because the factors that affect arrival time can change over time, such as due to new road construction or changes in traffic patterns.

By following these considerations, machine learning algorithms can be used to improve arrival time prediction accuracy and provide users with more reliable information about their travel time. 

**Simplified Block Diagram:-**

**PUBLIC TRANSPORTATION OPTIMIZATION**

**11. Reporting & Visualization**

- Dashboards

- Reports

**10. Arrival Time Alerts**

- Notify Users

- Delay Alerts

**9. Feedback Loop**

- User Feedback

- Model Refinement

**8. Continuous Model Updating**

- Regular Model Retraining

**7. Arrival Time Prediction**

- Real-time Predictions

**5. Model Evaluation**

- Metrics (e.g., MAE, RMSE)

- Validation Data

**6. Real-time Data Integration**

- Traffic Conditions

- Weather Updates

**4. Model Training & Hyper parameter Tuning**

-Training on historical Data

- Cross Validation

**3. Machine Learning Models**

- Regression Models

- Decision Trees

- Random Forest

- Gradient Boosting (e.g., XGBoost)

- Neural Networks

2 **2. Feature Selection**

- Relevant Features

- Feature Scaling/Normalization

**1 1. Data Collection & Preprocessing**

- Historical Arrival Data

- Real-time Traffic Data

- Weather Data

- Feature Engineering

- Data Cleaning

**Conclusion:-**

Optimizing public transport is an ongoing process that requires a multidisciplinary approach, incorporating technology, data analysis, infrastructure development, and public engagement. The goal is to create efficient, sustainable, and convenient public transport systems that serve the needs of urban populations while reducing the negative impacts of private car use on cities and the environment.

**THANK YOU…**

**Team Members:-**

1. **K.S. HARI PRASATH**
2. **M. HARISH**
3. **K. KISHORE SARATHI**
4. **R. SANTHOSH**