JAI SHRIRAM ENGINEERING COLLEGE
TIRUPPUR – 638 660
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai
Recognized by UGC & Accredited by NAAC and NBA (CSE and ECE)

DEPARTMENT OF

ELECTRONICS AND COMMUNICATION ENGINEERING

IBM - Naan Mudhalvan

Internet of Things – Group 3

Phase 3 – Development Part 1

PUBLIC TRANSPORT OPTIMIZATION

NAME : C.Keerthana

NM ID : AU711221106011

YEAR : III

PUBLIC TRANSPORT OPTIMIZATION

Development Part 1

Define the Scope of the Project:

Before you start, it's essential to clarify the goals and objectives of your public transport optimization project. Decide on what specific aspects you want to optimize, such as route planning, scheduling, or passenger demand prediction.

Acquire the Dataset:

You'll need data related to public transport, which could include information about routes, schedules, passenger counts, and other relevant factors. This data can be obtained from public transportation agencies, open data sources, or collected through surveys and sensors.

Data Preprocessing:

Once you have the dataset, you'll need to preprocess it to make it suitable for analysis and model development. This typically involves several steps:

Data Cleaning:

Remove duplicates: Check for and remove any duplicate records from your dataset.

Handle missing data:

Identify and decide how to handle missing values. This might involve imputation or removal of records with missing values.

Outlier detection:

Identify and deal with outliers that could skew your analysis.

Data Integration:

If your data comes from multiple sources, you need to integrate it into a single dataset. Make sure that the data is consistent and that you have a common key to link data from different sources.

Convert data types: Ensure that data types are appropriate for analysis. For example, dates should be converted to datetime objects.

Normalize numerical data: If you have numerical data, scaling it to a common range (e.g., 0 to 1) can help in training certain machine learning models.

Encode categorical data: Convert categorical variables into numerical format using techniques like one-hot encoding or label encoding.

Feature Engineering:

- Create relevant features that might be useful for your optimization task. For public transport, this could include:
- Creating new time-related features (e.g., time of day, day of the week).
- Calculating distances between stops or stations.
- Aggregating data at different spatial or temporal levels.

Data Splitting:

Split your dataset into different subsets for training, validation, and testing. Common ratios are 70% for training, 15% for validation, and 15% for testing. Adjust these ratios according to your project's needs.

Normalization and Standardization:

Depending on your modeling techniques, you might need to normalize or standardize your data. This is important for algorithms that are sensitive to the scale of the input features.

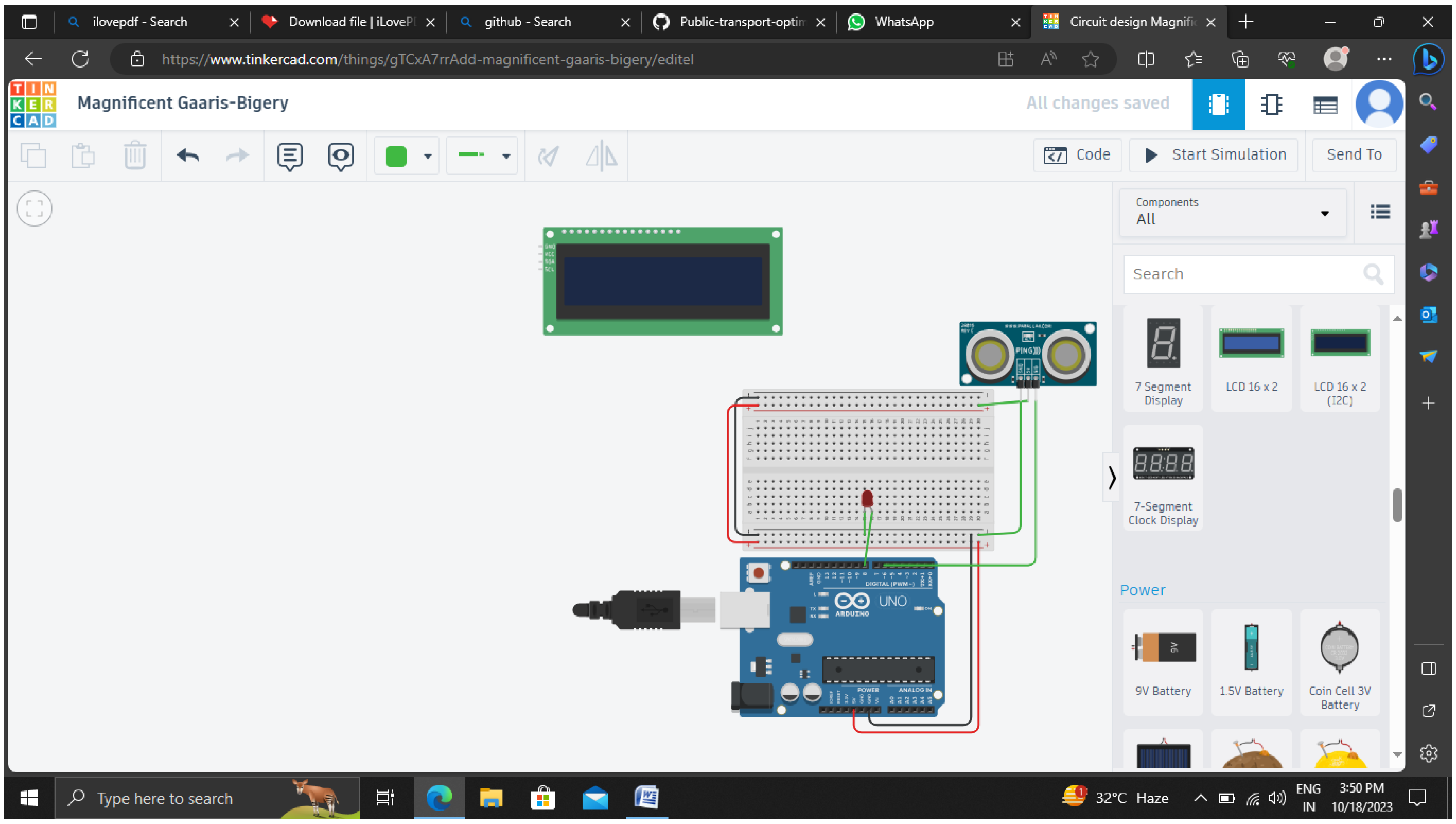
Data Visualization (Optional):

Create visualizations to explore the data further and gain insights. Visualizations can help in identifying patterns and trends in your public transport data.

Data Resampling (Optional):

If your data is imbalanced (e.g., there are significantly more records for one type of transport route than others), you may need to resample it to achieve a balance. This could involve oversampling the minority class or undersampling the majority class.

Circuit Diagram:



Save Preprocessed Data:

Once you've completed the preprocessing steps, save the preprocessed dataset to a file for easy access in your modeling and analysis phase. Common formats are CSV or database storage.

Documentation:

It's essential to keep a record of all the preprocessing steps, including the decisions made for handling missing data, encoding categorical variables, and creating new features. This documentation ensures transparency and reproducibility.

The preprocessing steps will depend on the specific dataset and the goals of your public transport optimization project. Keep in mind that data preprocessing is often an iterative process, and you may need to revisit these steps as you delve deeper into your analysis and modeling.

ExploratoryDataAnalysis(EDA):

Conduct EDA to gain insights into your dataset. This involves visualizations and statistical analysis to understand the distribution of data, correlations between variables, and potential patterns.

Model Development:

Depending on your project's objectives, you might need to develop various models or algorithms. Some common models for public transport optimization include:

a. Route Planning: Graph-based algorithms like Dijkstra's or A* for finding optimal routes.

b. Scheduling: Linear programming or optimization algorithms to create efficient schedules.

c. Demand Prediction: Time series forecasting or machine learning models to predict passenger demand.

Training and Evaluation:

Train your models using the training data and evaluate their performance using the validation dataset. Adjust model parameters and features as needed to improve performance.

Model Deployment:

Once you have a satisfactory model, deploy it in a real or simulated public transport system. Make sure the model is integrated with the existing infrastructure for real-world application.

Continuous Monitoring and Maintenance:

Monitor the performance of your optimized public transport system over time. Update the model as necessary to adapt to changing conditions and demands.

Documenation:

Document your project thoroughly, including data sources, preprocessing steps, model details, and results. This documentation is essential for reproducibility and knowledge transfer.

Communication:

Share your findings and solutions with relevant stakeholders, such as public transportation authorities, to gain support and feedback.

Remember that public transport optimization is a complex and dynamic field, and the success of your project will depend on the quality of your data, the appropriateness of your models, and the collaboration with relevant stakeholders.