AI Traffic Jam Prevention

Business Understanding

- What problem are you trying to solve, or what question are you trying to answer?

  The problem that I am trying to solve are urban traffic jams through vehicle detection and traffic flow control

- What industry/realm/domain does this apply to?

  Automotive, Government Department of Motor Vehicles and Transit

- What is the motivation behind your project? (Saying you needed to do a capstone project for flatiron is not an appropriate motivation)

  Traffic jams are one of the biggest inconveniences of living in modern society, especially in big cities and as more people migrate into these cities, the problem just gets more complicated. Cities keep developing projects to mitigate traffic jams but not much has changed on the drivers end of things where lack of coordination between drivers trying to get through a traffic jam may end up complicating the event unnecessarily.

Data Understanding

- What data will you collect?

  I will collect data related to traffic jams,

- Is there a plan for how to get the data (API request, direct download, etc.)?

  Direct download

- What are the features you'll be using in your model?

  Vehicle ID, Initial Position, Final Position, Time, Date, Location

Data Preparation

- What kind of preprocessing steps do you foresee (encoding, matrix transformations, etc.)?

  Data cleaning and normalization, feature engineering for positions and data augmentation

- What are some of the cleaning/pre-processing challenges for this data?

Finding erroneous data, creating relations between positions and time and setting coordinates

Modeling

- What modeling techniques are most appropriate for your problem?

  Spatial related statistics, point pattern analyses, k-means, regression models and clustering

- What is your target variable? (Remember - we require that you answer/solve a supervised problem for the capstone, thus you will need a target)

  Time

- Is this a regression or classification problem?

  Regression problem

Evaluation

- What metrics will you use to determine success (MAE, RMSE, etc.)?

  MAE, RMSE, R-squared

Tools/Methodologies

- What modeling algorithms are you planning to use (i.e., decision trees, random forests, etc.)?

  Regression models, clustering, random forests and SVM