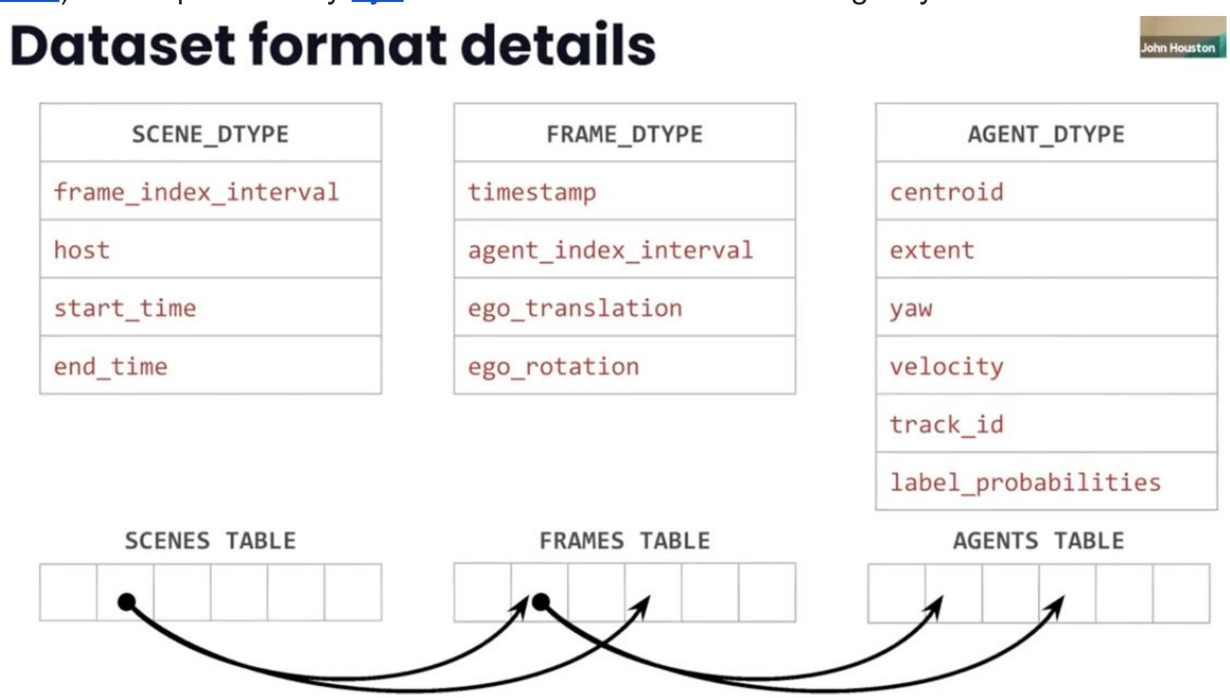
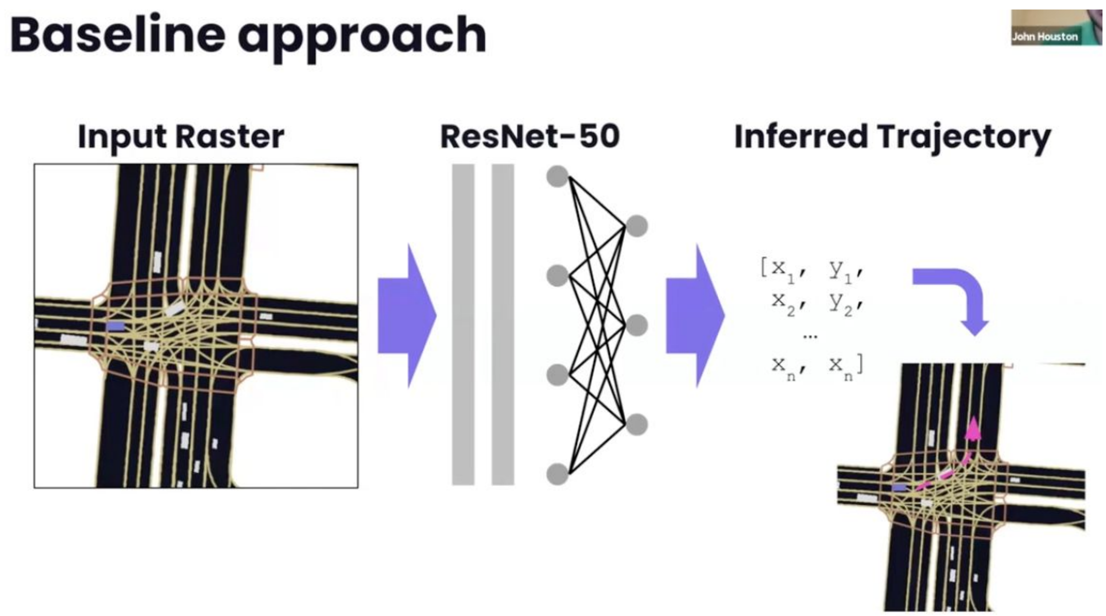
Proposal

Problem Statement:
Currently, many self-driving cars rely on rule based models to make decisions about about when, where, and how to maneuver; however, the real world often doesn't follow these rules and comes with many uncertainties. In order to make a future where self-driving cars make transportation safer, environment-friendly and more accessible for everyone, they must be able to reliably predict the movement of traffic agents around the autonomous vehicle (AV), such as cars, cyclists, and pedestrians. Given an HD map, the location of an AV on the map, and the current and past positions for cars, bicyclists, and pedestrians around the AV, we aim to create a deep learning model that predicts how these agents will move in the next 50 frames. The dataset (white paper here) will be provided by Lyft and is structured in the following way:



The result could then theoretically be used in the decision making process for the AV to take the correct maneuver.

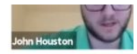


Approach: We intend on using a multimodal model to generate multiple hypotheses (up to 3) - further described by a confidence vector - to predict the trajectories of other traffic participants in a given scene. Our approach will largely build on [Multi-Modal Trajectory Prediction of Surrounding Vehicles with Maneuver based LSTMs]() by Deo and Trivedi by learning additional maneuver classes based on roadway traffic with traffic lights. While our dataset doesn't explicitly include the lane structures and intersections of the roadway in the present scene, perhaps we can somehow determine this based on the information we do have; however, we must first determine whether or not this information will be useful for our maneuver based model.

Experiments and Results:
All the data we intend to use has been provided by Lyft thus no data collection is necessary. In addition there are several unique solutions to this problem which have been posted on Kaggle and we intend to reference those separate solutions to garner new ideas and approaches if we become stuck. Regarding the code, we plan to use the [documentation]() that already exists from Lyft about the problem to come up with a baseline solution. From there, we plan to modify the solution with both the knowledge gained in CS 4476 and our combined personal knowledge in order to achieve our goals. To obtain our results, we plan to analyze both the accuracy and efficiency of both our initial solution and our modified solution. We will use the baseline metrics as a control group of sorts, and analyze how well the modifications we made in our modified solution improved upon our baseline solution. We also will compare our solution to the various others submitted to Kaggle. Given this we hope to achieve a model that can correctly predict the trajectory of any agent with an accuracy around or above 70%. This is an ambitious goal due to the fact that we collectively have never done something like this

before and on such a large scale as well. We hope to achieve this goal but more importantly we hope to contribute a new and unique perspective to this challenging problem. Thus our true goal is merely a contribution to the ever-growing field of autonomous vehicles to ensure the safety of everyone.

Large Datasets Enable Deep Prediction



Statistic	**Value**
# self driving vehicles used	20
Total data set size	1,118 hours / 26,344 km / 162k scenes
Training set size	928 hours / 21,849 km / 134k scenes
Validation set size	78 hours / 1,840 km / 11k scenes
Test set size	112 hours / 2,656 km / 16k scenes
Scene length	25 seconds
Total # of traffic participant observations	3,187,838,149
Average # of detections per frame	79
Labels	Car: 92.47% / Pedestrian: 5.91% / Cyclist: 1.62%
Semantic map	15,242 annotations / 8,505 lane segments
Aerial map	74 km^2 at 6 cm per pixel