CREATING A TRAFFIC DETECTION ARTIFICIAL INTELLIGENCE

Daniel Huynh
Department of Computer Science
University of Virginia
Charlottesville, VA USA
tap7ke@virginia.edu

Srikar Gouru
Department of Computer Science
University of Virginia
Charlottesville, VA USA
mcj2vb@virginia.edu

ABSTRACT

Traffic is an exponentially growing problem in the United States, directly or indirectly causing more than $305 billion dollars in annual loss in the US alone. A major issue for consumers is the ability to properly predict traffic at a stoplight. Our research aims to apply Recurrent Neural Network (RNN) techniques to solve this issue. We used the UTD 19 dataset containing data from various traffic detectors in Switzerland across the span of 3.8 years. We constructed an RNN-based model that is trained on, and meant to run on, a single detector. The model did not learn as well as expected, due to an unbalanced dataset leading to lack of observations in which traffic was encountered. Moving forward, we hope to create an easy-to-use user interface to allow researchers to setup detectors and predict traffic at any traffic light in the world. This research can be applied by consumers when choosing their route, and also by mapping engines like Google Maps and Apple Maps when algorithmically determining the best route for a consumer to take.

1 PROBLEM AND IMPORTANCE

Oftentimes when driving, there are many similar routes that lead to the target destination. The ability to know which one is best may change based on many factors, including time of the day, day of the week, whether or not it's a holiday, if it's a long weekend, and many more. The size of the road, number of lanes, the location of the road, or even the speed limit may play a role. Determining which of these factors plays the most pronounced role may aid everyone on their daily commutes and make the right decisions on how to get to their destinations on time.

However, traffic causes more than just inconvenience for an average commuter. In 2017, in the US alone, commuters lost more than $305 billion dollars due to traffic jams, simply by sitting in traffic (INRIX, 2023). Simply sitting in a stationary vehicle on the road leads to a great buildup over time in fuel consumption, and especially builds up over the year. The power of AI can be leveraged to avoid these costly buildups, rerouting users from the worst routes to a destination, and allowing users to take similar routes that avoid these conditions. These benefits can help not just drivers and car owners, but can also prove useful for buses and other forms of public transportation on the roads. In addition to the economic and convenience benefits, there can be great environmental impacts– transportation has been found to be the leading source of greenhouse gas emissions, solely responsible for a net increase in 47% of total United States emissions since 1990 (Hagerman, 2023). Especially in the context of vehicles simply sitting stationary in traffic, the amount of pollution released into the atmosphere is a major problem to be addressed. In this case, reducing the amount of traffic can greatly reduce the amount of time spent releasing carbon into the atmosphere from fuel waste as a result of traffic. Even still, not only is it a major contributor to climate change, but traffic also contributes to noise pollution, air quality damage, and destroys ecosystems, especially in urban areas. The greater the number of cars that are on the street unfortunately increases the volume and disruptions to surrounding areas.

This paper outlines the creation of a model utilizing machine learning integrations into smart city applications such that users can accurately predict whether traffic will occur on routes of travel given a variety of different factors.

2 DATASET

The dataset used for this paper is the UTD19 dataset, compiled by ETH Zurich, a public university in Switzerland. The UTD19 dataset was compiled by the Institute for Transport Planning and Systems at ETH Zurich, and was done through the deployment of sensors on urban roads. It boasts itself as the largest multi-city traffic dataset publicly available, with over 23500 detectors across 40 cities worldwide, over the course of 3.8 years.

The dataset was initially created as three separate CSV files, each with separate features. We were able to merge these CSV files into a single dataset that contained the following information:

- Day
- Interval (time of day)
- Detector ID
- Flow
- Current average speed of traffic
- Speed Limit of Road
- Number of Lanes
- Length of Section of the Road
- Position of Sensor on Road (from an intersection)

The Detector ID was unique to each detector, so all of the data with a single Detector ID represented a single traffic intersection. Much of our preprocessing and data organization was based around this Detector ID value. The Flow at each measurement represents the number of cars passing the detector region per unit of time. Through preprocessing of the data, we were able to extract **Day of the Week** from the Day value as well.

Additionally, we ended up only using a few of the features mentioned here because of the fact that many features were constant throghout a single detector. Since we were only interested in developing a **By-Detector Model** for the purposes of this research paper, many of these values were not con.

3 APPROACH

To solve the problem, we researched and experimented with recurrent neural networks to work with the time series traffic-based data, of which many factors can occur that lead up to a traffic jam, of which we can identify.

This problem has been addressed before, in which an RNN was trained to predict if possible traffic jams will occur on a route, trained upon a similar data set from Waze. This particular project created a generalized model that simply predicted if whether or not traffic would occur. In order to expand on this project, we decided that simply a generalized model would not take into account the complexities that each road may have. A road in another part of the world, with different driving habits (such as even as simple as side of the road driven), different cultures, different speed limits or road layouts, or niche localized conditions cannot be completely generalized to a single model representative of all communities across the world. We decided on the effectiveness of the creation of a platform instead, relying upon citizen science and the development of smart city applications, rather than a single model.

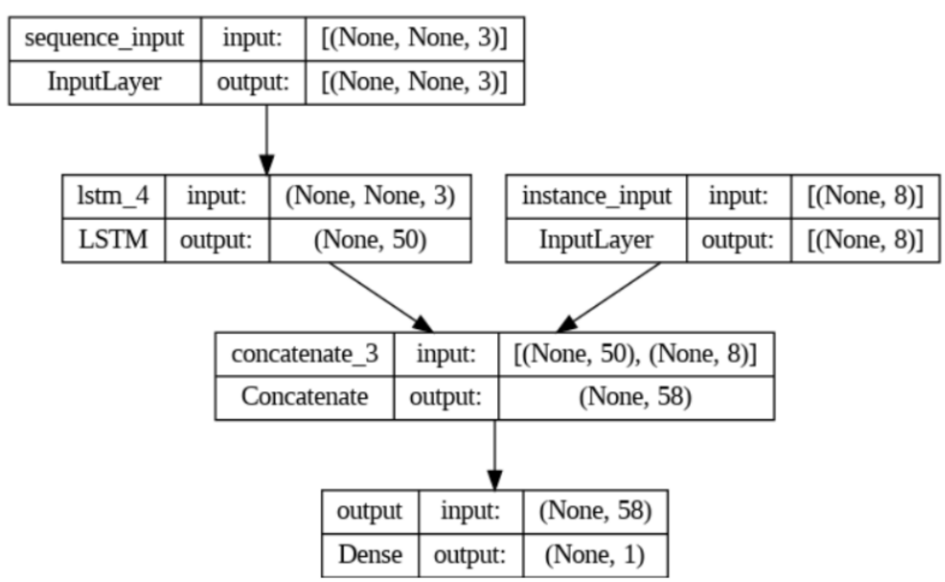
In our platform, we would allow local governments or communities to set up a detector much like ETH Zurich's detectors on a particular specific road to be studied. The platform we create will output specific predictions for that specific street or community, rather than a generalized prediction. This will allow for more accurate predictions that cater and account for specific regional or community cultures or habits. Collecting this specific data that we outlined earlier in our model, we will create a **By-Detector Model**, utilizing data such as:

- RNN that takes in previous data from the past hour:
  - Previous speeds
  - Previous flows
  - Previous times
- Prediction time
- Day of the week

Because this model was meant to be trained on a single detector, we were able to ignore values that were dependent on the detector, since those would be constant at every data point and would thus not affect the model training in any manner. These values included the speed limit of the road, the number of lanes in the road, the length of section of the road, and the position of the sensor on the road.

In order to achieve this model, we first had to preprocess the data such that we would only have dataframes including the data from each specific detector. Next, we created a function that implements a sliding window to achieve the compilation of ”previous” data to be used in the recurrent neural network. This sliding window selects all the observation data in a 60 minute period. This window of data will be used as the ”previous data” that will be used to predict whether or not traffic will occur in the time period 60 minutes into the future from the end of the window. In simpler terms, at a certain time, we will use the previous hour of observation data to predict an hour ahead from our current time into the future.

We will use the prediction time, the time that is an hour ahead into the future, as part of the prediction from the **By-Detector Model**. Additionally, our binary classification for what is traffic is determined by checking the speed of the traffic at the prediction time, and comparing it with the speed limit of that road. After careful consideration, we set the threshold to 0.8 times the speed limit on that road. If cars were predicted to go faster than this speed, then that would be classified as no traffic, and if cars were predicted to go slower than this speed, then that would be classified as having traffic. The general architecture of the model can be observed in the diagram below:



In order to build this model, we used TensorFlow's Keras model. Utilizing long short term memory (LSTM) to implement the recurrent neural network, dependent upon the previous 60 minutes, we feed the previous times, speeds, and flows into the LSTM layer. The LSTM layer had a Dropout of 0.2, with a ReLU activation. We also handle inputs that are dependent and related only to the specific instance, such as prediction time or day of the week, using the Input Layer called "instance_input" to handle these inputs, which fed into a Dense layer to encode the instance inputs. This layer also utilized a ReLU activation. We combine the two input layers, and feed it through one more hidden layer of 50 units, which then is connected to the final output layer. The model then outputs a prediction of whether or not there will be traffic one hour into the future (with respect to the time period we used for the "previous times" data), using softmax as the activation. Traffic in this case was preprocessed to be defined as having an average speed that is less than 80% of the speed limit of this specific road. In addition, this model used a learning rate of 0.0001.

4 Results

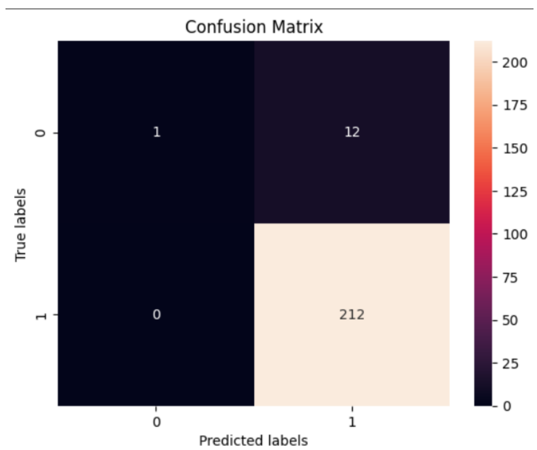


Figure 1: By-Detector Model Confusion Matrix

We trained our **By-Detector Model** on 50 epochs. This resulted in a final training accuracy of 92%, and a final validation accuracy of 94%. Additionally, we achieved a final training loss of 0.19 and a final validation loss of 0.16 when using the categorical_crossentropy loss function.

While this model appears to be performing well, this is not quite the case. As shown in the confusion matrix (**Figure 1**), the model did not learn well. In fact, it almost exclusively predicted a label of **1** (indicating no traffic), and it merely happened to get it right due to the disparity in our dataset. Due to a lack of data for any given detector, most of our data (90%) had a label of no traffic, and only a small proportion of our data (10%) had a label of traffic. We were unfortunately unable to balance our data because deleting that much data would result in very little left to train.

Our lack of data also showed up during training. After the first eight epochs, our model seemed to stay stagnant and not learn, with the training accuracy and loss both plateauing and not improving or worsening over the rest of the 42 epochs of training.

5 Conclusion and Next Steps

As seen in these results, there is still a lot of room for improvement. The development of such a platform requires both a more dependable set of models, as well as an incentive for local governments and citizens to participate in the study, even if for the greater good of the community. If developed further, given a longer period of time, these models can be fined tuned such that we can find the most optimal set of hyperparameters and neural network architecture such that we can improve our validation accuracy to an even greater accuracy. In addition, since our dataset was highly dependent on when these detectors were active on their respective streets, the time period could either be in a period of high volume or low volume of traffic. That means that in our By-Detector Model, the data that we would have provided to train our model for that specific detector may have been unbalanced. When we then test our model for validation accuracy, the confusion matrix would reveal that all observations that were used to validate the model were either mostly all labeled as having traffic or mostly all labeled as having no traffic. This made it less clear as to whether or not the model would simply be led to predict the classification that was most prominent of that set of time used, or if it had truly identified the features or patterns that would lead to such a prediction. As such, given more time, our next steps would have to try a new approach to the preprocessing of our data, and will take a deeper look into the nature of each of these detectors' data in case there were any patterns that could have caused such an unbalanced set. This may also be the result of the fact that while there was a wide variety of detector sensors that were included in this dataset, the amount of data per detector was only over the span of a few days. As such, a By-Detector Model would be trained upon and tested upon a small set of data, likely causing the effect of an unbalanced set of data that we encountered, as well as a less accurate model overall due to the less amount of data it was trained upon.

Additionally, we wish to create a more general, Generalized Model that is not dependent on individual detectors. This would be much more similar to that of previous developments from the Waze dataset. The goal of this model would be for it to work automatically on any detector, given some extra features related to that detector. This way, researches could set up a detector and immediately start predicting, instead of having to train a separate model for that detector. This architecture would also allow the model to scale easier in terms of storage, since a single model can predict traffic at any detector set up anywhere in the country.

In addition, another difficulty of our project was that there were not as many features available in our dataset. We believed initially that the By-Detector Model was an approach that would allow us to cater to each specific city and region of the world, realizing that there were many more factors that could play a role in traffic on a certain day or time. We believe that the amount of features that were available to us were not enough for a generalized model, leading us to move in the direction of the By-Detector Model over the General Model as our main model, but that simply more features about the road such as the width of the road, or about the city such as the city's population, population density or distribution, or even the weather patterns or weather data of the day and time of detection would be of major benefit to the creation of a much more accurate General Model.

ACKNOWLEDGMENTS

We would like to thank our professor, Prof. Ferdinando Fioretto for teaching us this content and encouraging us to apply it to the real world. We would also like to thank the ETH Zurich students for compiling the UTD19 dataset. Finally, we would like to acknowledge Google Colab as we depended heavily on their resources to compute our results and collaborate with one another.

REFERENCES

Hagerman, K. (2023, May 25). How commuting impacts the environment. RideAmigos. https://rideamigos.com/commuting-impacts-environment:~text=The%20impact%20of%20driving %20to%20work%20Here's,car/SUV%20(14%20MPG%20fuel%20economy):%205.7%20tons.

Inrix. (2023, October 4). Congestion costs each American 97 hours, $1,348 a year. https://inrix.com/press-releases/scorecard-2018-us/:~text=Boston%20drivers%20lost%20up%20to, studied%20at%20$304%20per%20driver.