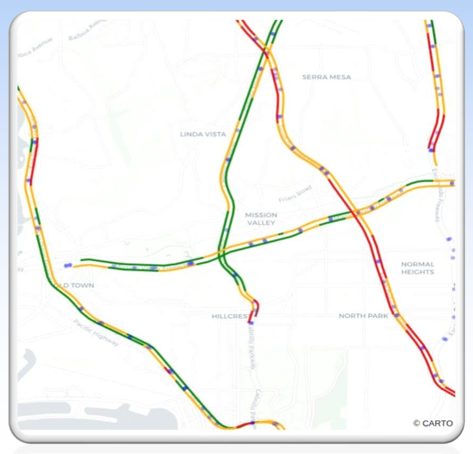
TORCHTS: Deep Learning based Traffic Forecasting in the Covid era

UC San Diego
Jacobs School of Engineering

Advisors: Dr. Rose Yu, Dr. Ilkay Altintas

Team: Akash Shah, Aparna Gupta, Daniel Roten, Kevin Lane, Raul Martinez



INTRODUCTION

From the onset of the COVID-19 pandemic, a dramatic change in traffic patterns has been observed across the country due to travel and other restrictions imposed by government agencies and health experts. We hypothesize that the performance of static time series models used for traffic forecasting will degrade beginning in early 2020. Dynamic models that do not rely solely on historical information will better forecast day-to-day traffic and be able to support long-term changes in traffic patterns. We present a graph convolutional recurrent neural network that captures both the inherent spatial and temporal complexities present in traffic forecasting.

DATA SOURCING AND STORAGE

Traffic dataset (30 second updates):

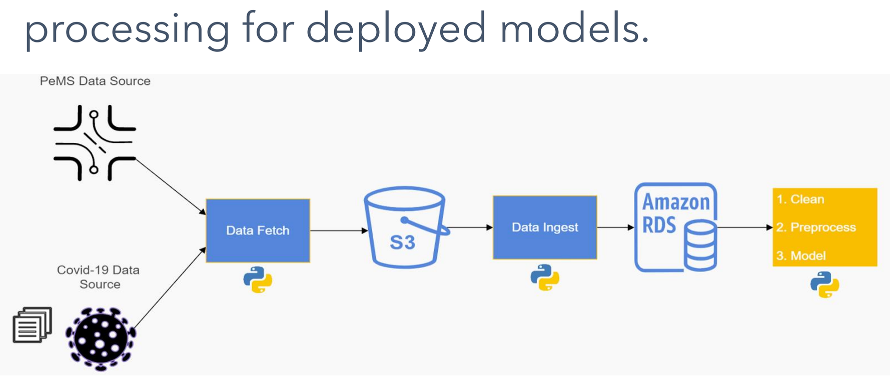
The Caltrans PeMs dataset contains near real-time traffic observations collected from more than 40,000 sensors installed on freeways in all California metropolitan areas. The raw data is provided in 30 second intervals and Caltrans also provides rollups on various intervals including 5 minutes, hourly, and daily.

Covid dataset (Daily updates) :

The John Hopkins dataset is provided in a Github repository which contains CSV files with the numbers of confirmed cases and deaths both in the U.S. and globally. The U.S. data is subdivided into counties, along with out-of-state and unassigned cases for each state.

Data Pipeline:

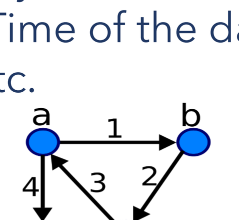
Raw PeMS data is provided by Caltrans in CSV format stored in GNU zip (gzip) files for a given district and day on AWS S3. Daily measurements, are stored in a database hosted on Amazon RDS to permit easy access for all team members and enable online processing for deployed models.



Modeling preparation:

1. **Train test split**: Traffic data was ingested for the first half of the year 2020 (01-01-2020 to 06-30-2020) , when the number of COVID-19 cases started rising. Rolling average methods over the prediction horizon was used to impute missing values. Further, data was subdivided into train (70%), val (10%) and test (20%). Various methods of infusing COVID-19 and traffic datasets were also explored.

2. **Building the adjacency matrix:** Traffic data is spatiotemporal and is represented as a graph of sensors with time series at every node. DCRNN takes input signals a Directed Weighted Acyclic graph as an adjacency matrix with N features. e.g. Traffic speed, Time of the day, No. of COVID cases, etc.



	1	2	3	4
a	1	0	-1	1
b	-1	1	0	0
c	0	-1	1	0
d	0	0	0	-1

MODELING APPROACH: DIFFUSION CONVOLUTIONAL RECURRENT NEURAL NETWORK (DCRNN)

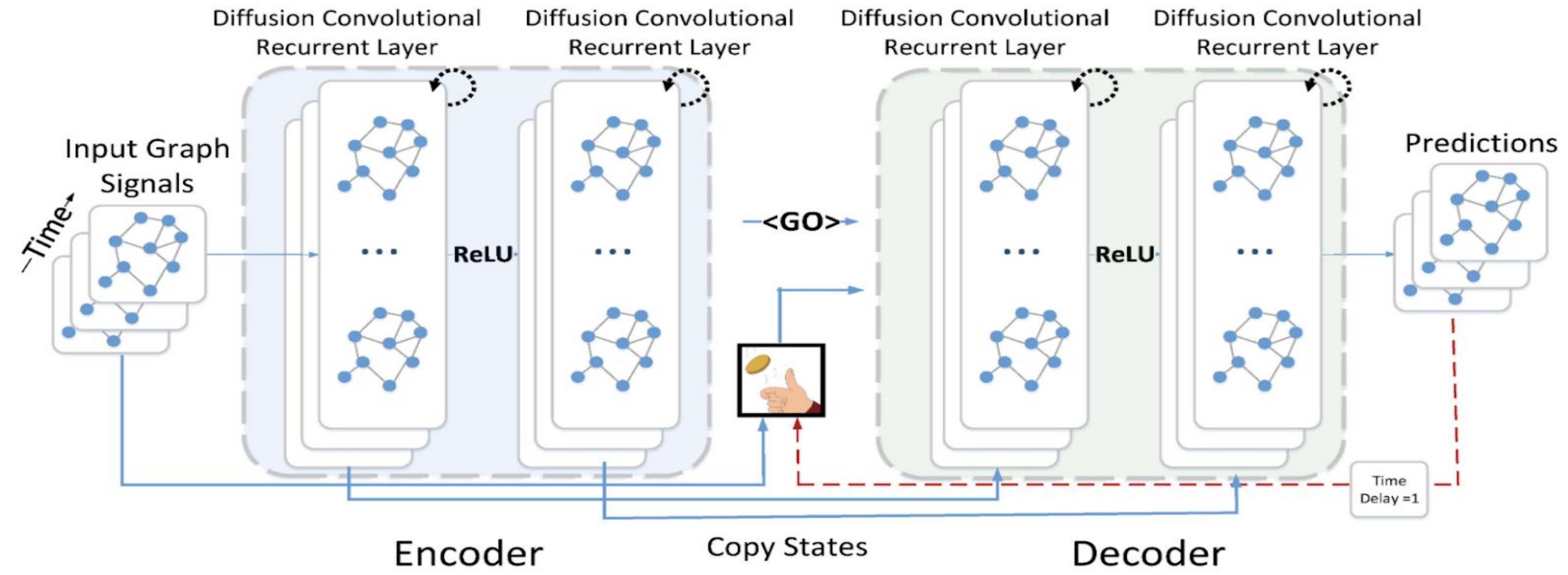
Diffusion spectral graph convolution:

The Convolution operation over a graph signal $X \in R^{N \times P}$ and P features, a filter f_θ is defined as:

$$\sum_{k=0}^{k-1} (\theta_{k1}(D_O^{-1}W)^k + \theta_{k2}(D_I^{-1}W)^k)X_{:,p}$$

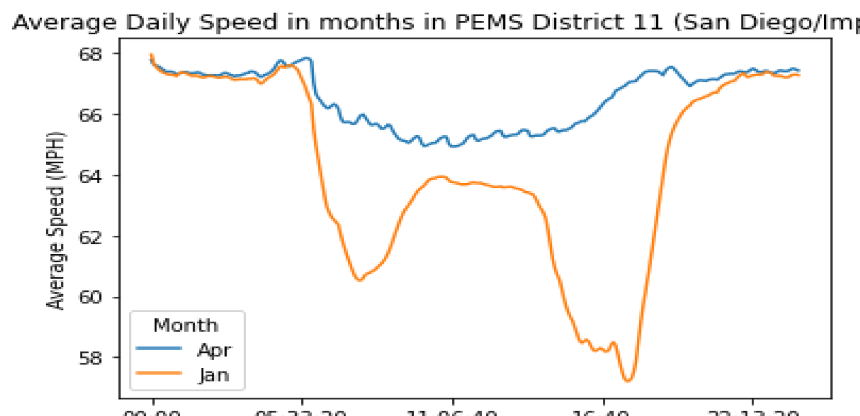
where $(D_O^{-1}W)$ *and* $(D_I^{-1}W)$ are the transition matrices for the diffusion and the reverse diffusion process and k is the maximum number of diffusion steps.

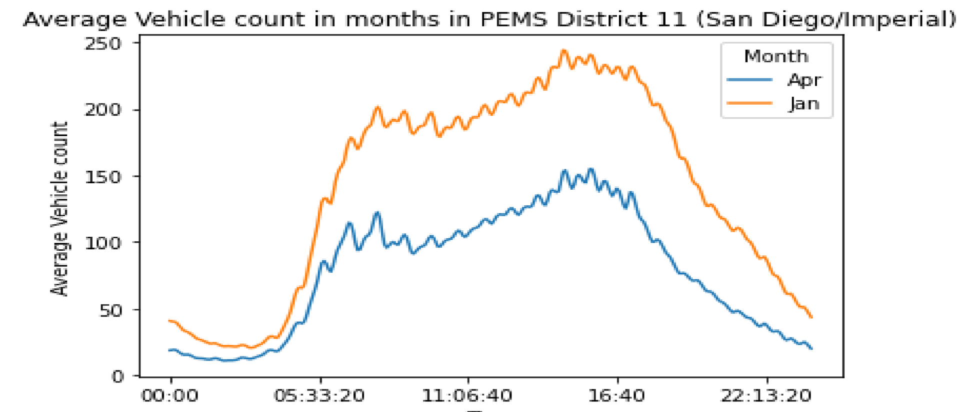
Network Architecture: An Encoder Decoder network is used to learn the temporal dynamics of the system which is trained using curriculum learning. The seq2seq is set up to 1 hour, with each decoder timestep predicting a 5-minute interval. The maximum diffusion steps was set to 2 and the diffusion process is captured using a dual random walk process, that ensures bidirectional behavior. The spatial dynamics are modeled using a Diffusion Convolutional GRU Cell, which is a specialized GRU with graph convolutions at the reset and update gates of each timestep.



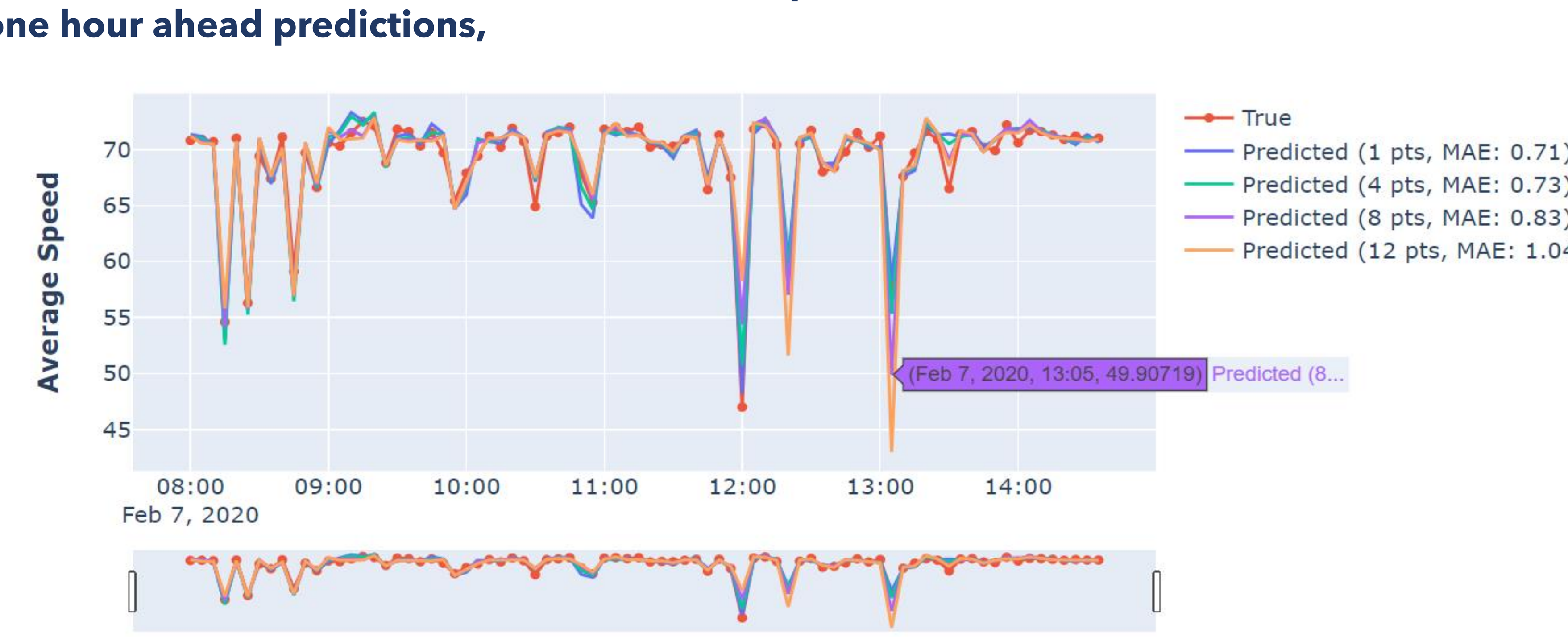
RESULTS AND INSIGHTS

Traffic patterns have changed dramatically in 2020, with a dramatic reduction in rush hour traffic and large decrease in the number of vehicles.

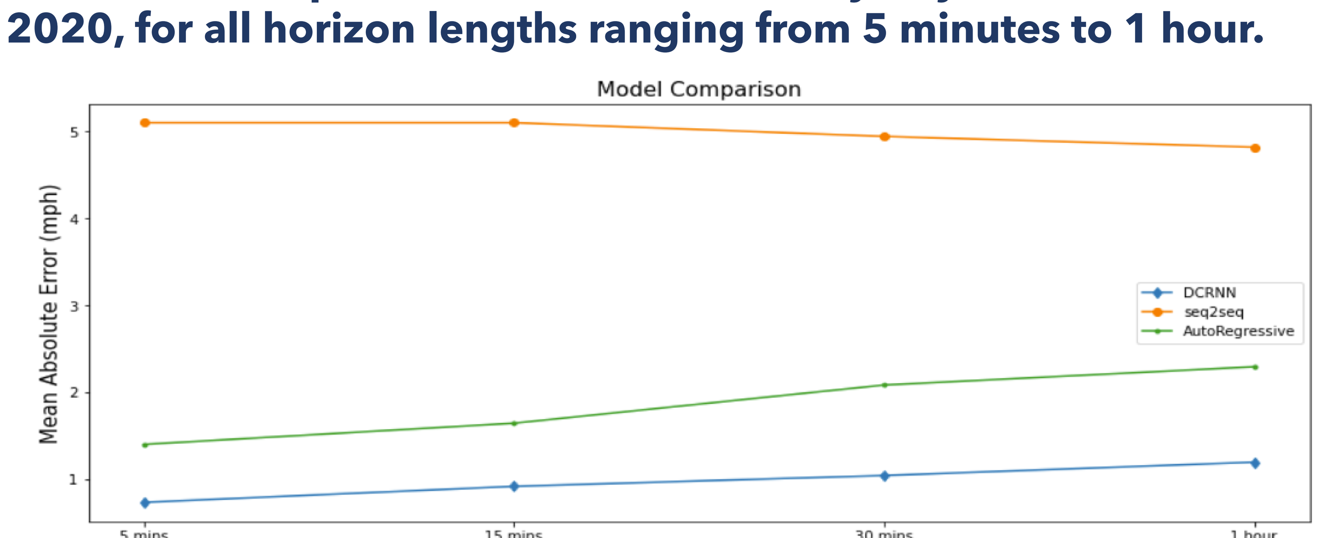




DCRNN's diffusion convolution helps model traffic effectively, predicting traffic speeds with an MAE (Mean Absolute Error) of close to 1mph across 320 traffic sensors in 2020. Even for one hour ahead predictions,

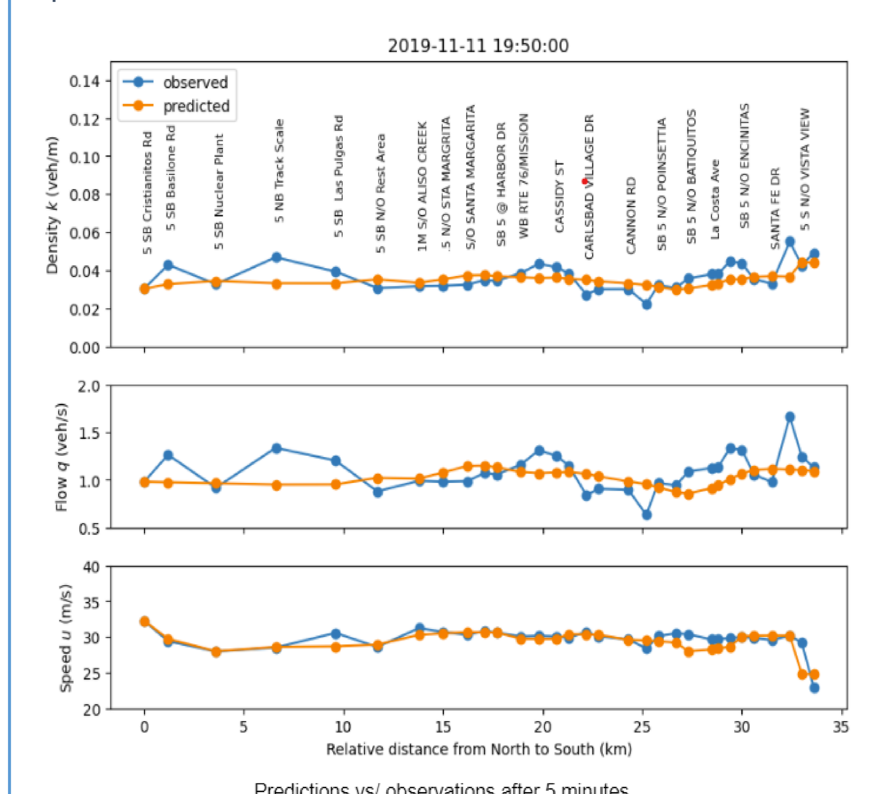


As expected, DCRNN out-performs models that solely rely on historical information on data from June 2020, for all horizon lengths ranging from 5 minutes to 1 hour.



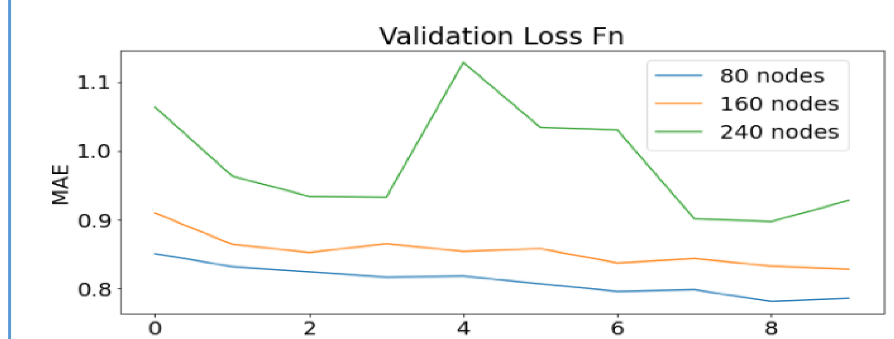
PHYSICS BASED PDE MODEL

The physics-based neural network for traffic prediction is based on the Lighthill-Whitham-Richards (LWR) model. It treats traffic as a fluid stream and couples a partial differential equation representing conservation of mass with a fundamental equation relating traffic flow to traffic density. PDE in the LWR model are solved numerically using an explicit finite difference scheme. We use PyTorch's auto differentiation feature and the Adam optimizer to find the parameters (e.g., the jam density and free velocity) which result in the lowest misfit between training data and predictions. With these optimized parameters, traffic conditions (speeds, density and flow) can be predicted into the future.

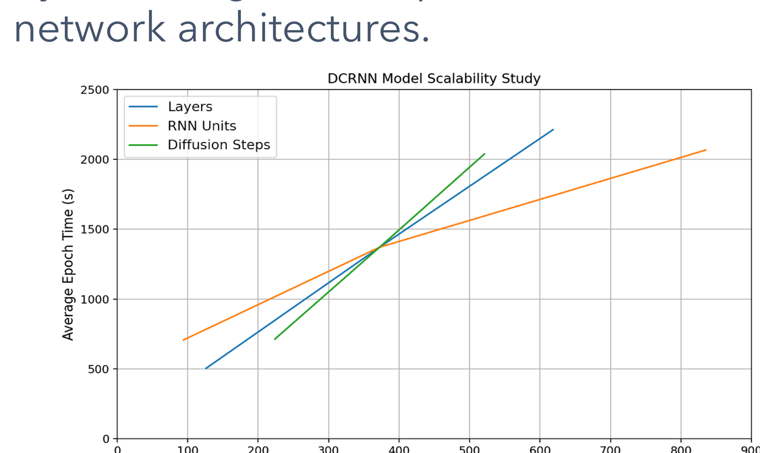


DATA AND MODEL SCALABILITY

Data Scalability: Model performance evaluated by incorporating different subsamples of the traffic graph, 25%, 50%, and 75% to analyze data scalability.



Model Scalability: Model scalability is tested by assessing the compute time for different network architectures.



Compute Scalability: Compute scalability was achieved by wrapping DCRNN around PyTorch Lightning. The model scales well in a data parallel environment with ~25x speed ups.

Number of GPUs	1	2	4	6	8
Training time/epoch (mins)	103.04	25.27	13.23	7.37	4.13

CONCLUSIONS

Although our experiments with the four different time-series forecasting models show that all methods have their strengths in the prediction of traffic and Covid-19 cases, the DCRNN emerged as our viable modeling product as it accurately predicts short-term variations in road traffic across the freeway network.