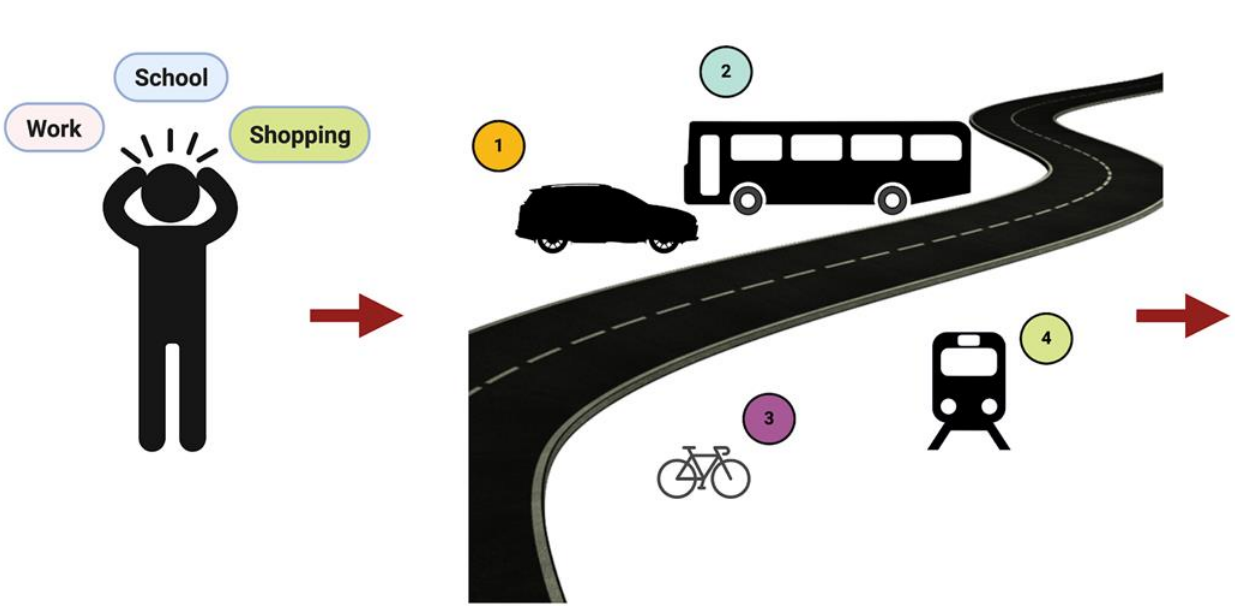
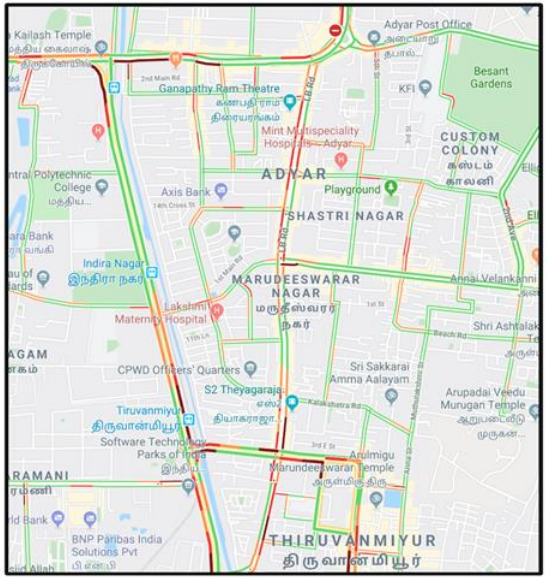
Machine Learning in Demand Forecasting

Vetri Venthan Elango
Workshop on Smart Mobility
Dec 17, 2019

Why Demand Forecasting





How was it done

- Traditionally data were collected through travel diaries.
  - Samples in the order of 1000s to 10,000s.
- Four Step Travel Demand Model
- Model typically used in planning

What has changed

- Tremendous improvements to computing power, storage and machine learning algorithms
- Massive amount of data generated and archived
  - A Traffic Management Center could collect traffic volumes every 20 seconds at each station
  - Uber completes approximately 14 million trips per day





What is new

- Trends in Research
  - Trip Generation
  - Mode Choice
  - Short-term demand forecasting
- Trends in Practice
  - Short term demand forecasting
  - Asset Allocation

Trends in Research

Data Sources

- Open Source Trip Data
  - Bike Sharing
- Weather Data
- Travel Surveys
  - National Household Travel Surveys
- Cell-phone Data
- Land Use / Geographical Data

Models

- Trip Generation / Short Term Demand Forecasting
  - Gradient Boosted Regression Tree
  - Neural Network
  - Recurrent Neural Network
  - Long Short Term Memory
  - Support Vector Machines
- Activity participation
  - Input-Output Hidden Markov Model
  - Convolutional Neural Network
- Mode Choice
  - Decision Tree Classifier
  - Artificial Neural Networks

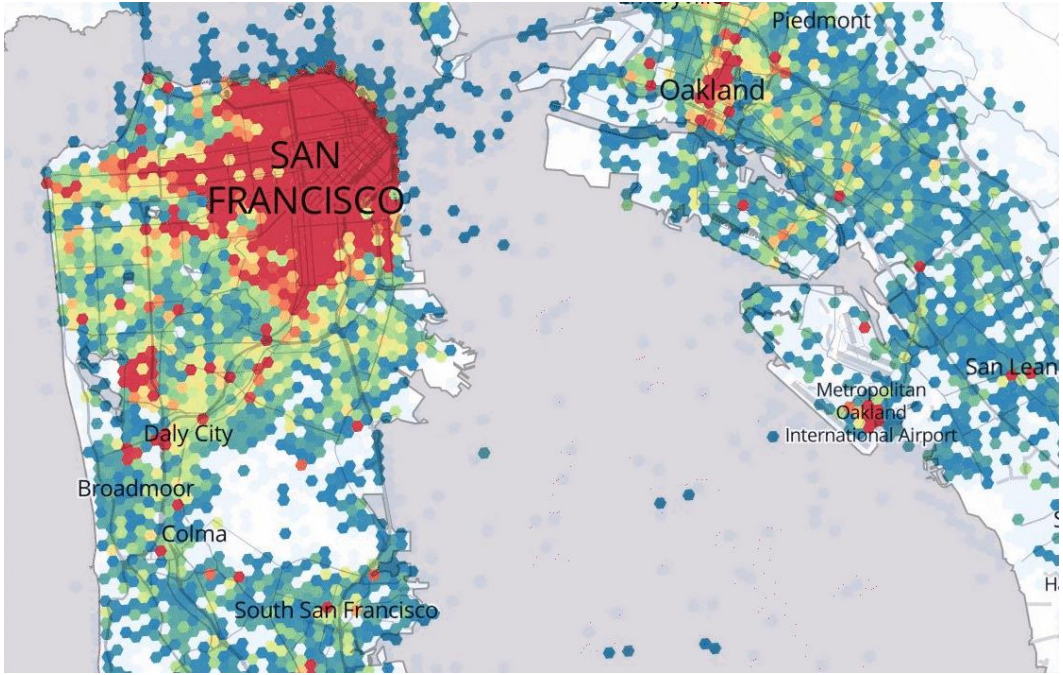
Learnings

- ML approaches outperform traditional statistical methods - BUT not by too much
- Use of new signals - e.g. weather
- Potential to increase accuracy with additional signals

Trends in Practice

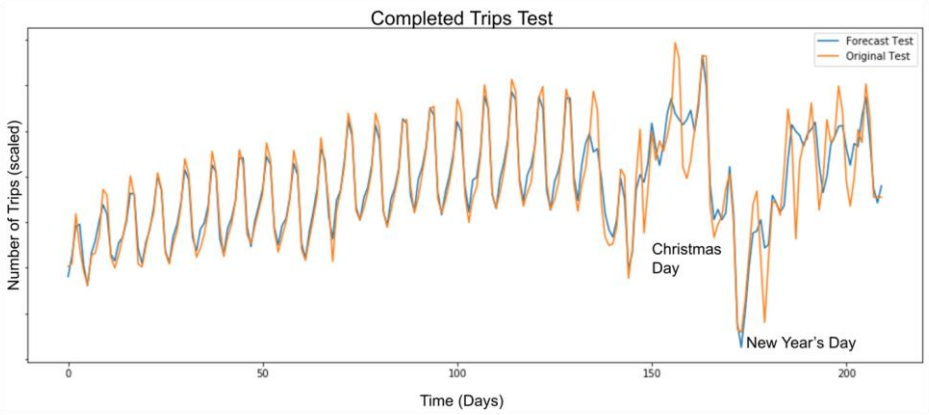
Uber Demand Forecasting

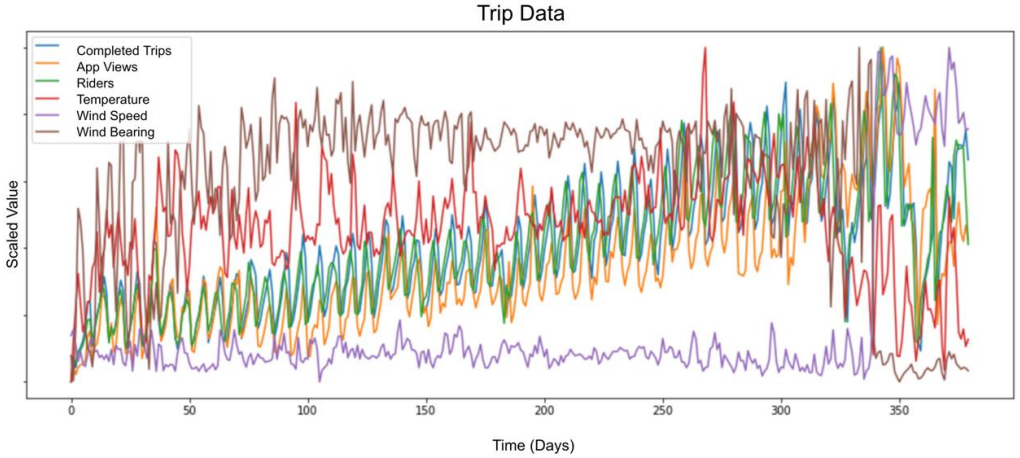
- Critical Element of Uber's Platform
- Classical Statistical models are simple and interpretable
- ML models are complex and flexible
- Compare model performance to naive approach
- Very important to measure the uncertainty of the prediction.



Extreme Event Forecasting – Uber

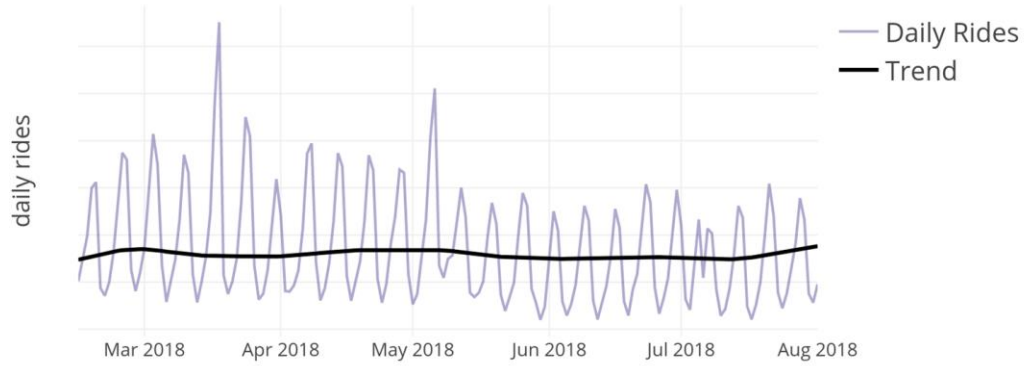
- Extreme Events prediction is critical
- Challenging with Limited Data
- Solution Built on LSTM architecture
- Training with sliding windows
- Results suggest 2-18% improvement over their existing model

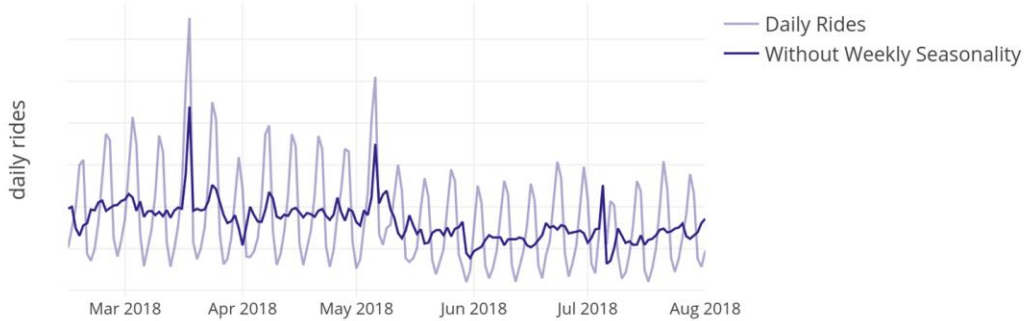




Seasonality of Demand - Lyft

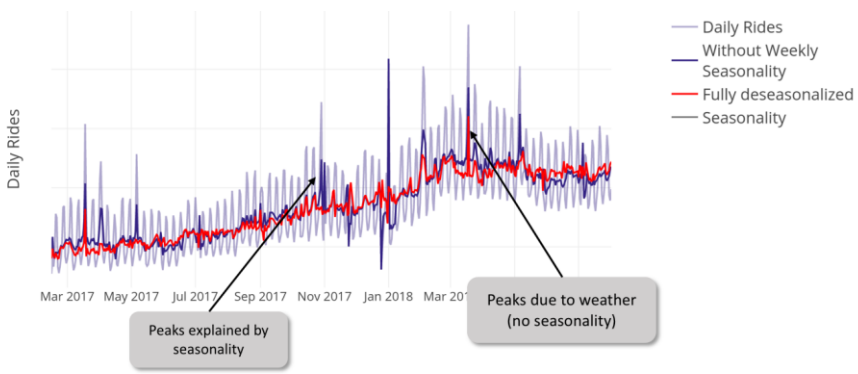
- Ensuring Supply and Demand balance is critical
- Weekday vs Weekend seasonality needs to be accounted
- Kalman Filters are used evaluate the hidden trend in the time series

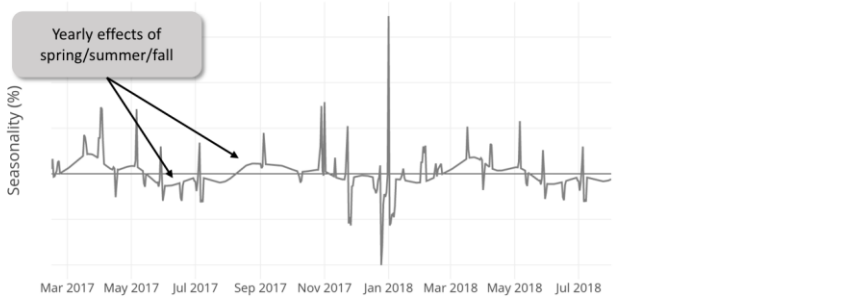


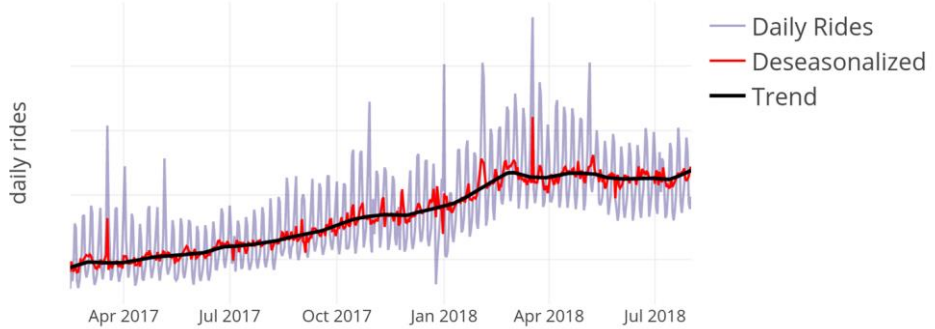


Holiday and Events - Lyft

- Recurring events (holidays) impact the time series the same way.
- Seasonality such as Summer vs Winter modeled using Fourier Decomposition
- After eliminating the effects of all seasonality the real underlying trend can be forecasted.

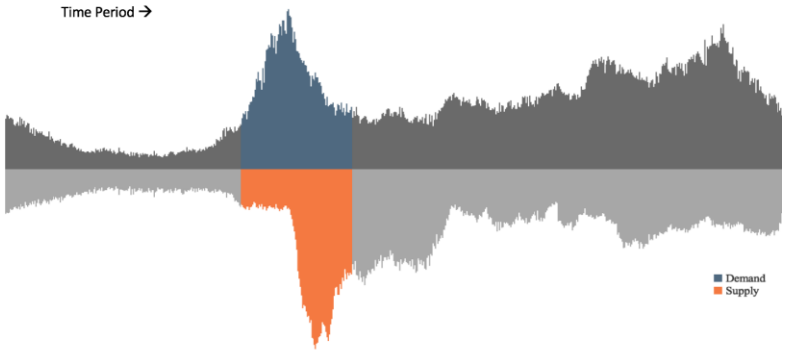






Supply and Demand - Grab

- Determine Effective Supply to each Demand
- Aggregate Metric is Supply Demand Ratio or Supply Demand Difference
- Spatial Opportunities to move Supply
- Temporal Opportunities to move Demand



Learnings

Use Statistical Models for inferring and ML models for predicting.

Prediction confidence intervals are critical to decision-making

Rare and extreme events are critical to balance demand and supply

Historical statistical models can provide insights to improve existing operational inefficiencies.

Questions for the future

Can we use Predicted Short-term Demand to predict transportation network conditions.

Transportation Operations - Reactive to Proactive

Data analytics to Actionable Insights