

Predicting Sick Patient Volume in a Pediatric Outpatient Setting using Time Series Analysis

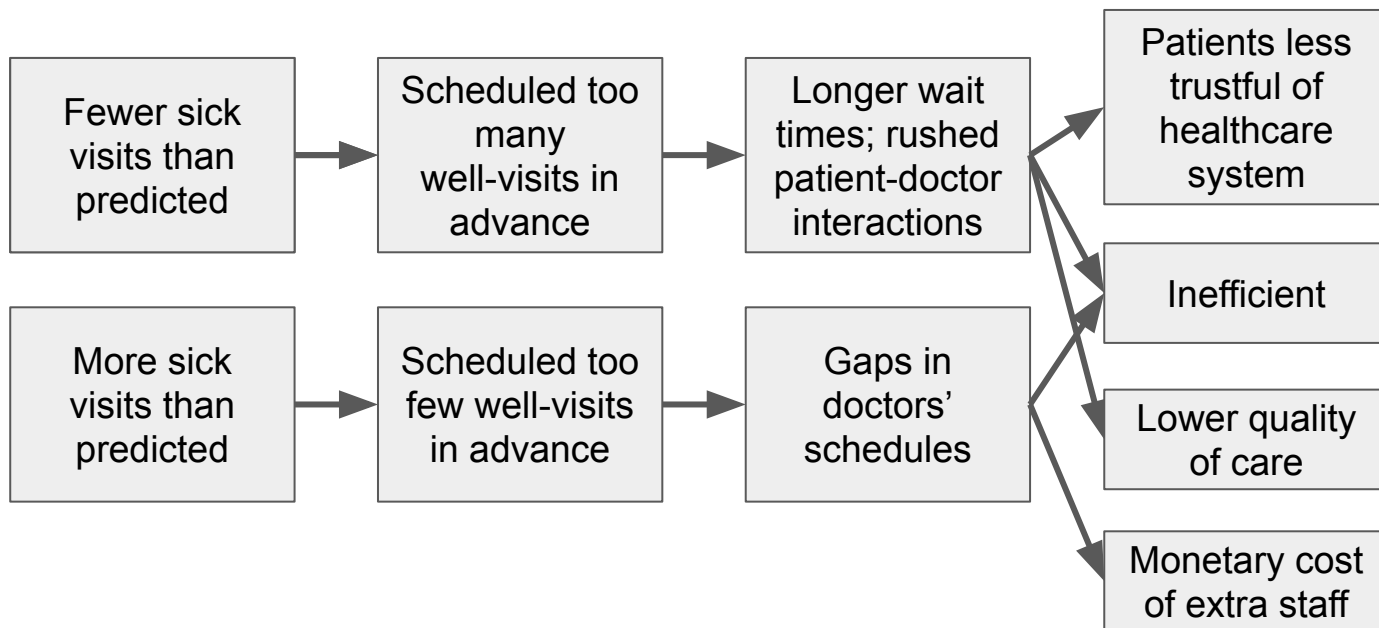
Grace Guan (speaker), Barbara E. Engelhardt
Princeton University

Problem Background

- Pediatric group w/ 30,000 yearly patients; 9 years' data (2010-2018)
- 2 types of visits:
 - Well visits = booked around 3 months in advance, check-ups
 - Sick visits = booked no more than 24 hours in advance, walk-in
- There are a constant amount of well visits scheduled for each day
- The remainder of a doctor's time is intended to be used for sick visits

Problem Background

- Need to accurately predict the number of sick visits...

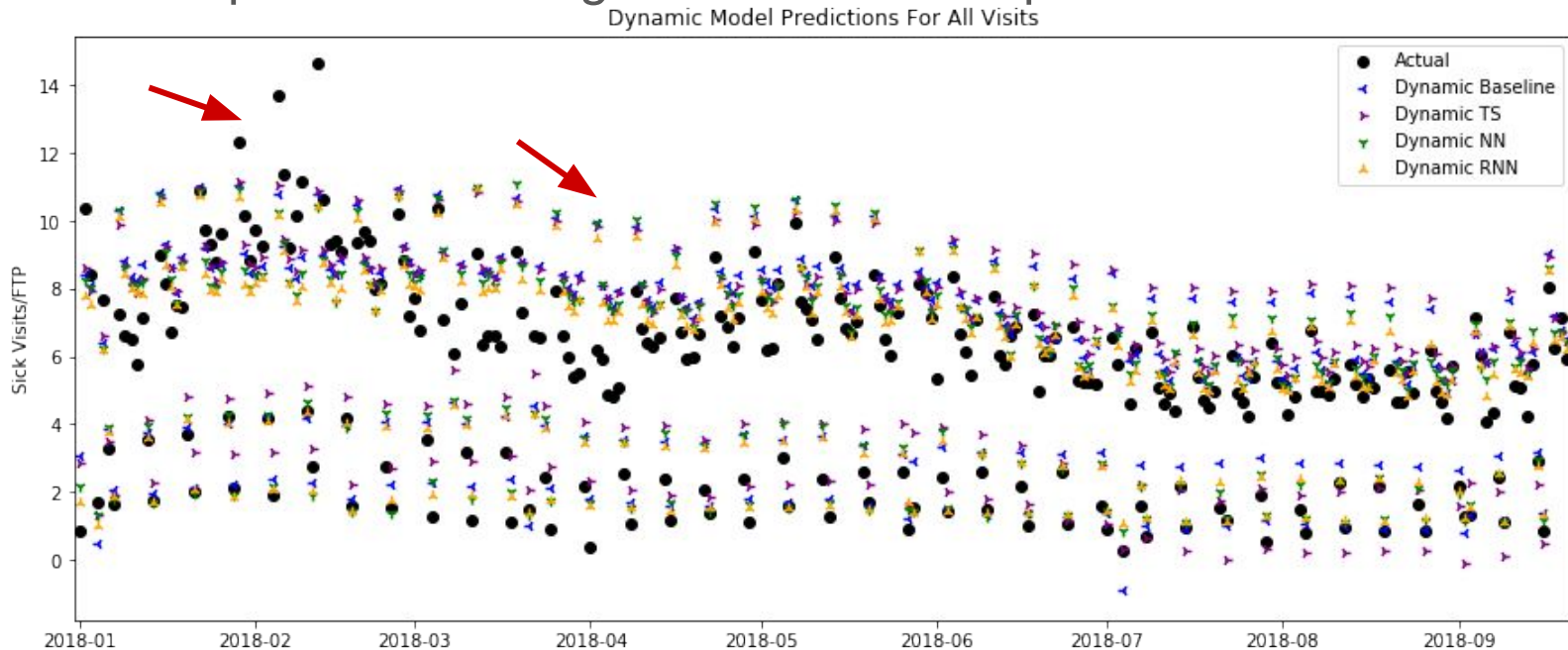


Models

- Linear regression (baseline), time series SARIMA from literature
 - Non-linear neural network and RNN/LSTM models
-
- Many different prediction tasks - Can RNN/LSTM models learn the time patterns and other features in the data?

Key Results

- Train on 2010-2016 data, Validate/Withhold 2017, Test on 2018
- RNN performs best; however, all models underpredicted in February and overpredicted throughout March and April.



Key Results

- **Different time scales of staff planning; yearly disease dynamics**
- We improve model performance (especially in Feb) by:
 - One day ahead planning using walk forward validation
 - Using lag features
 - Decomposing visits by ICD diagnosis code and using separate models for prediction
- Despite all models improving, RNN still performs best!