Absenteeism Prediction and Extra-Board Driver Scheduling for Metro Transit

Xiaochen Zhang

University of Minnesota
Joint work with Qie He (UMN) and Soniya Somani

November 1, 2018

Outline

- Motivation
- Existing Results
- Prediction Model
- Experiment Results

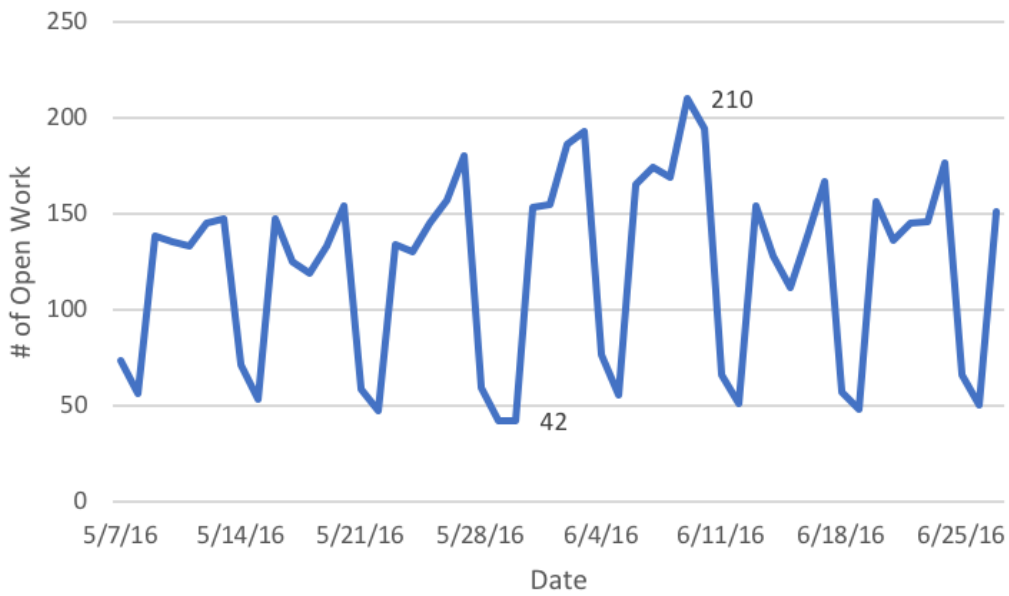
Background

The Minneapolis Metro Transit operates 127 bus routes with over 250,000 annual riders. It's extensive network of bus service employs more than 1500 operators at five garages.

Motivation

1. Assigned runs of regular drivers can be open work because of absences, personal day-off, and injuries.
   - According to the U.S. Bureau of Labor Statistics (2018), the absence rate was 3%, and the lost worktime rate was 1.7% in transportation occupations in US for 2017.
2. The amount of open work can vary a lot daily and seasonally.

Figure: Number of Open Work



Motivation

Table: Goal and Performance of Call Driver Utilization YTD October 2017

Metric	Nicollet	Heywood	Ruter	South	East	System
Goal	65%	65%	65%	65%	65%	65%
Utilization	58.5%	81.4%	56.1%	53.8%	69.5%	62.2%

Table: Goal and Performance of Late Pullouts YTD October 2017

Metric	Nicollet	Heywood	Ruter	South	East	System
Goal	3	3	3	3	3	15
Lost Pullouts	4	25.2	14.6	1.8	2.8	48.7

Existing Results

General Absenteeism Prediction:

- Job satisfaction [Schaumberg and Flynn, 2017] [Cohen and Golan, 2007]
- Prior absent history [Ivancevich, 1985]
- Driver information such as age [Markussen et al., 2011], gender [Markussen et al., 2011] [Laaksonen et al., 2008].

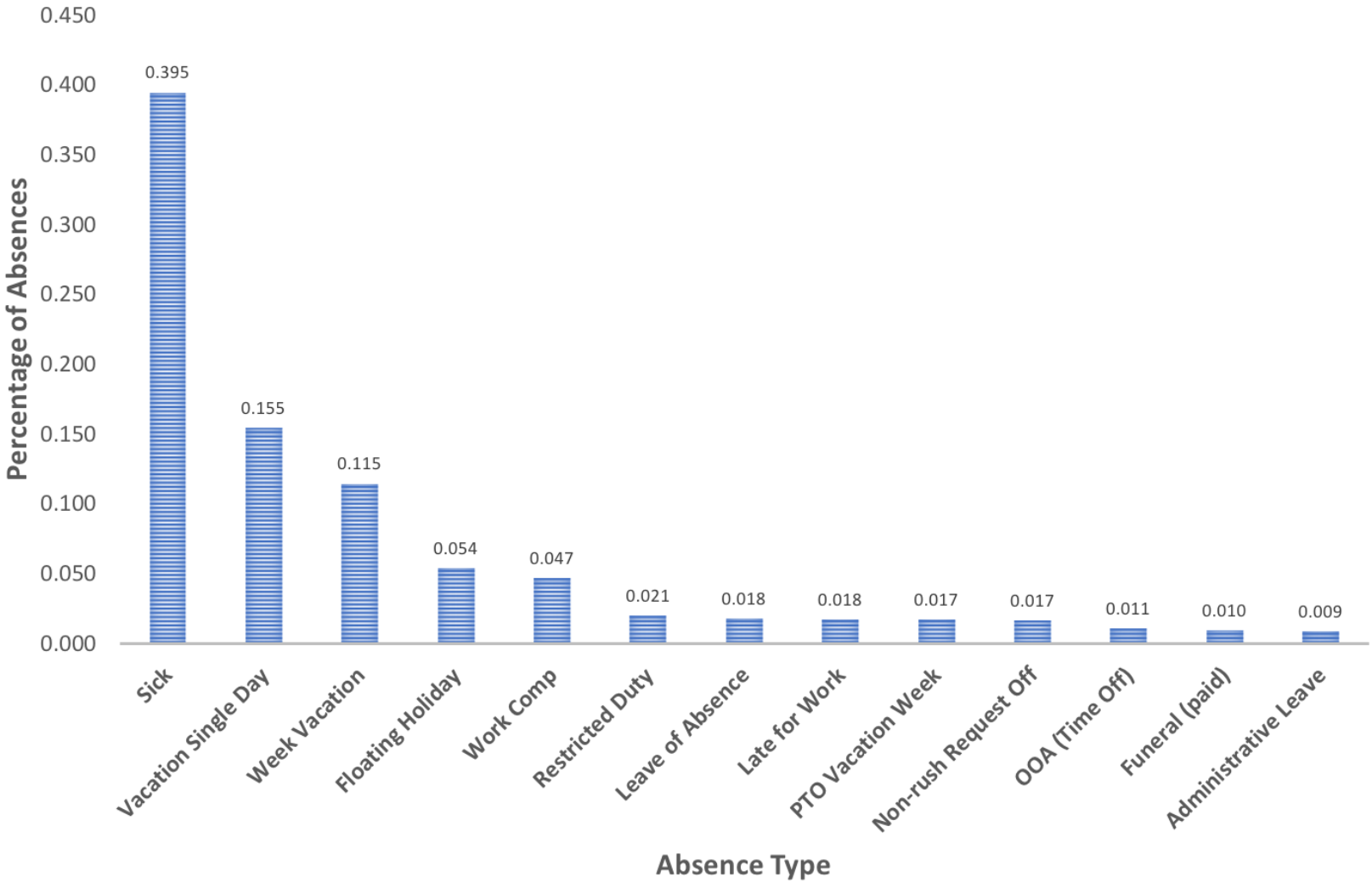
Existing Results

Absenteeism Prediction in Transportation:

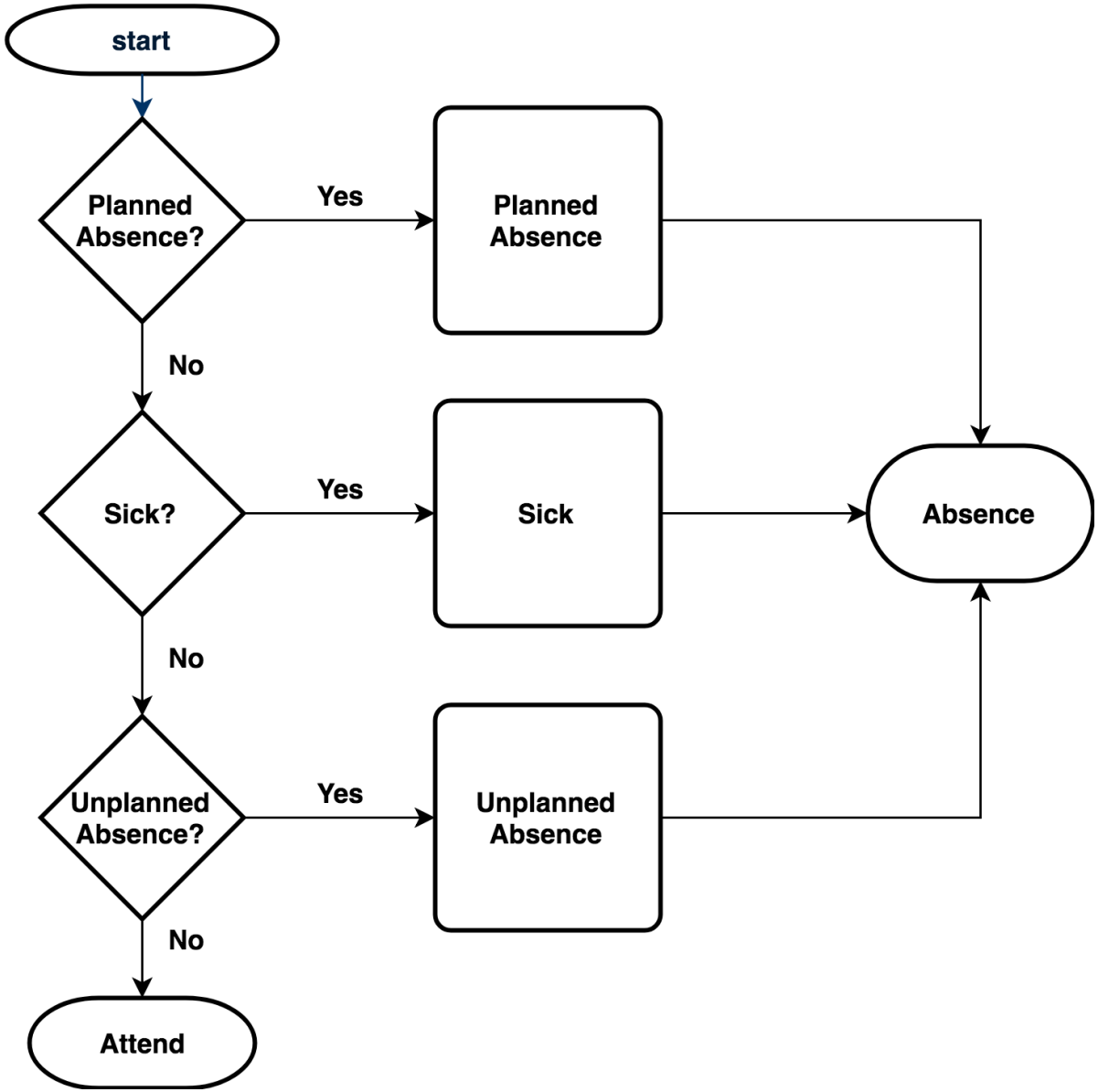
- Individual level absence likelihood prediction [Strathman et al., 2009] [Strathman et al., 2012]
- Aggregated level absence prediction [Diab et al., 2014]
- ...

Prediction Model

Figure: Frequency of Different Absence Types



Hierarchical Prediction Model



Hierarchical Prediction Model

1. Logistic regression on each layer
2. Features
   - Work information (month, day of the week, start hour)
   - Driver information (gender, age, sick carry)
   - Weather information (precipitation)

Hierarchical Prediction Model

Test Procedure:

1. Individual absence likelihood:

$$p_{absence} = p_{planned} + (1 - p_{planned}) p_{sick} + (1 - p_{planned})(1 - p_{sick}) p_{unplanned}$$

2. Aggregated number of absences: Add up all the absence probabilities

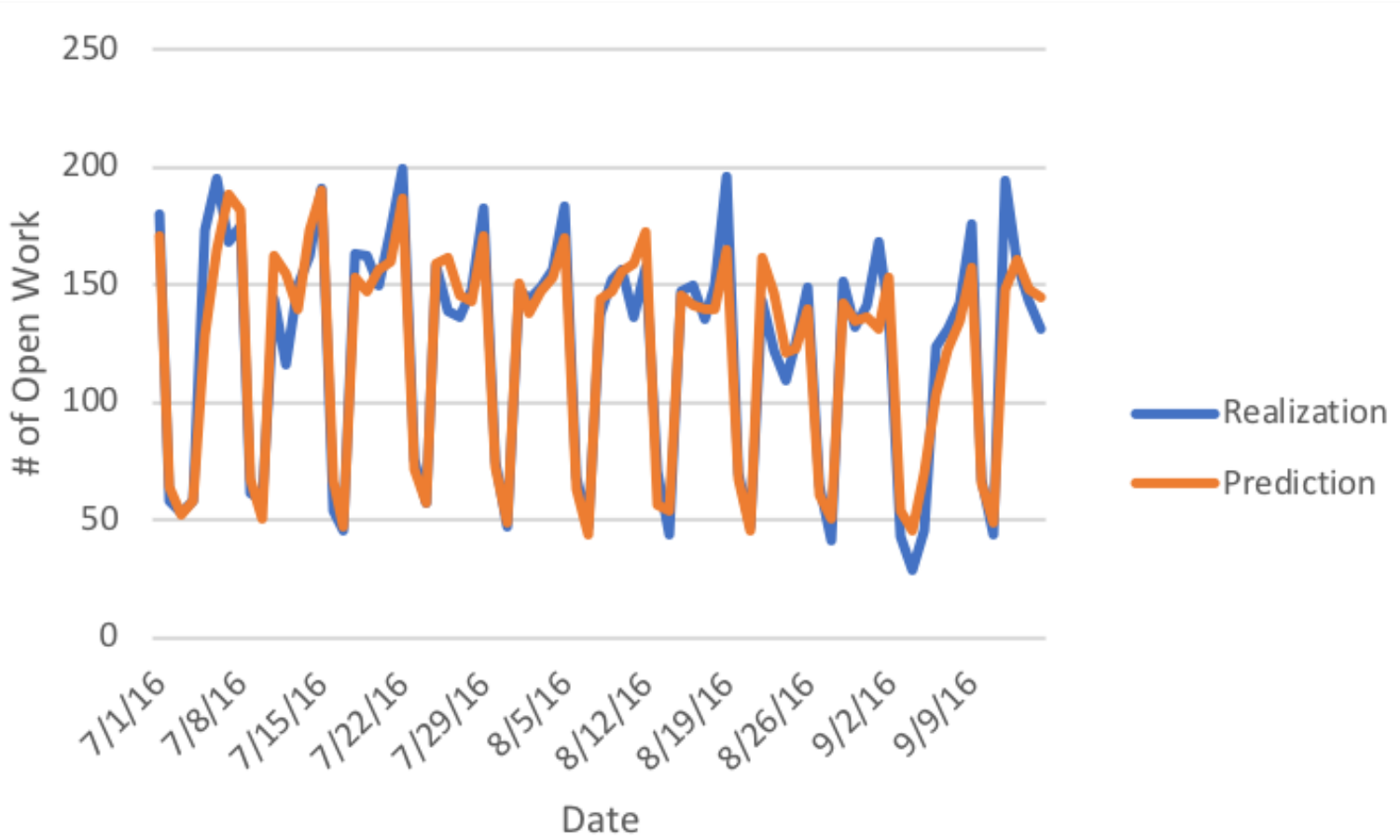
Experiment Results

Table: Basic Statistic and Errors in 2016

Avg Number of Pieces of Work	Avg Number of Open Work
1705	173
Mean Squared Error (Individual)	Mean Aggregated Error
0.05143759	9.2388

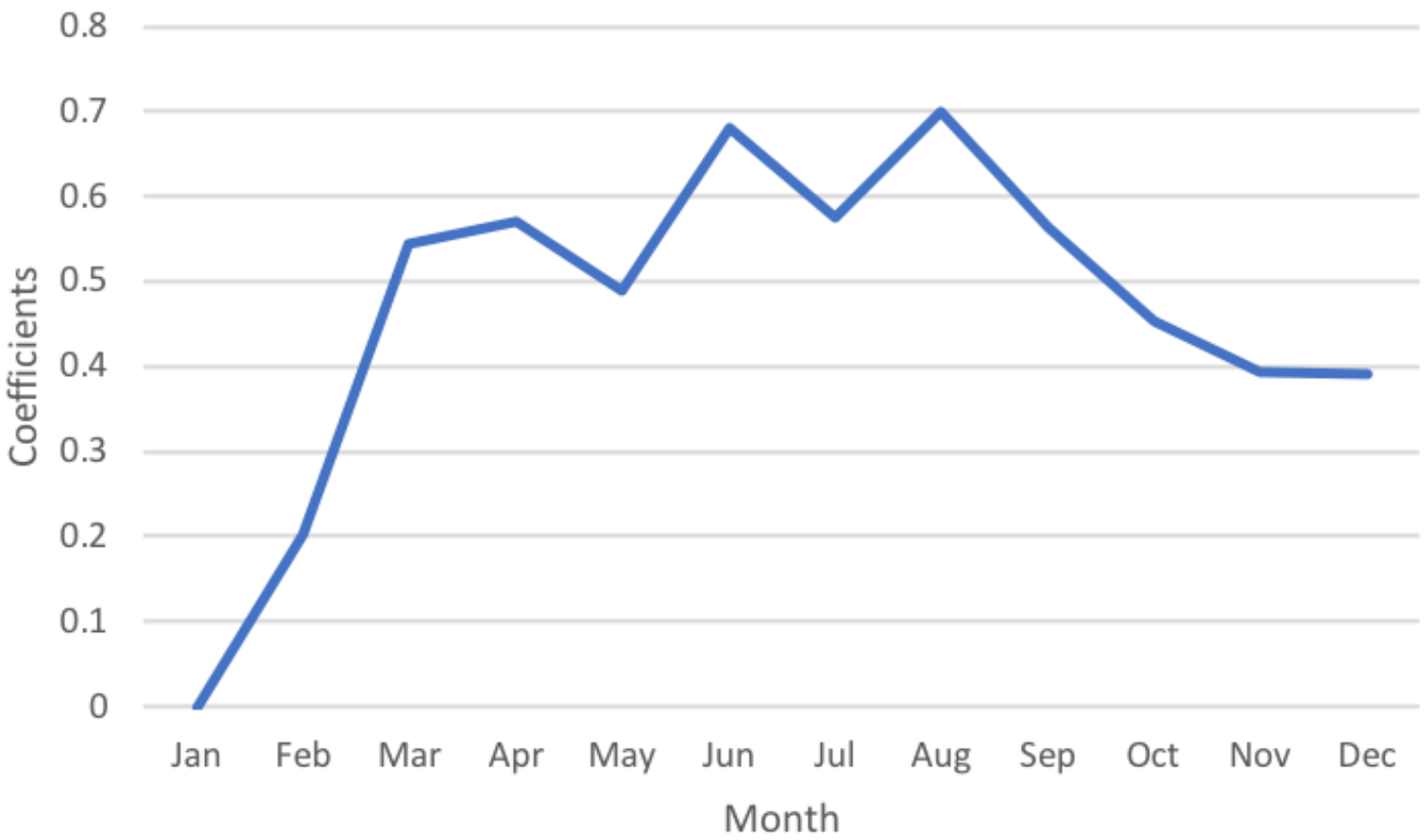
Experiment Results

Figure: Comparison Between Predicted and Realized Values of Summer 2016



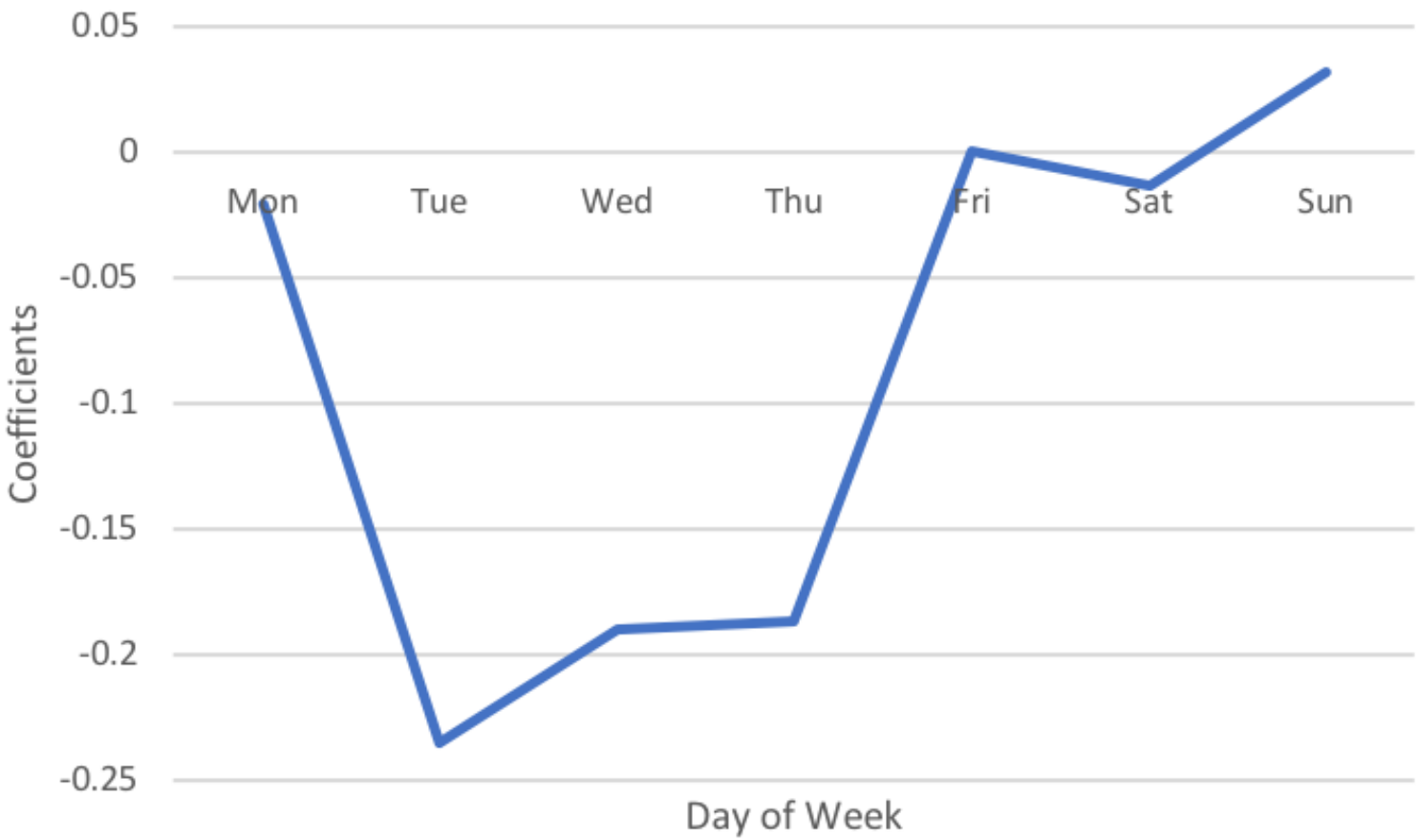
Experiment Results

Figure: Coefficients of Months in Planned Absence Model



Experiment Results

Figure: Coefficients of Day of the Week in Unplanned Absence Model



Summary

We

- Developed a hierarchical prediction model helping dispatcher predict daily absenteeism.
- Provided insights on important features for absenteeism.

Future Application:

- Help to decrease the absence rate
- Realtime open work assignment algorithm

Thank you!