



# Yard Departure Prediction: Case Study from Malmö Yard



Niloofar Minbashi  
Division of Transport Planning  
KTH  
FR8RAIL III, PRATA  
2021-11-23



# Agenda

- Motivation
- Problem
- Model Framework
- Case Study
- Results
- Conclusion

# Motivation



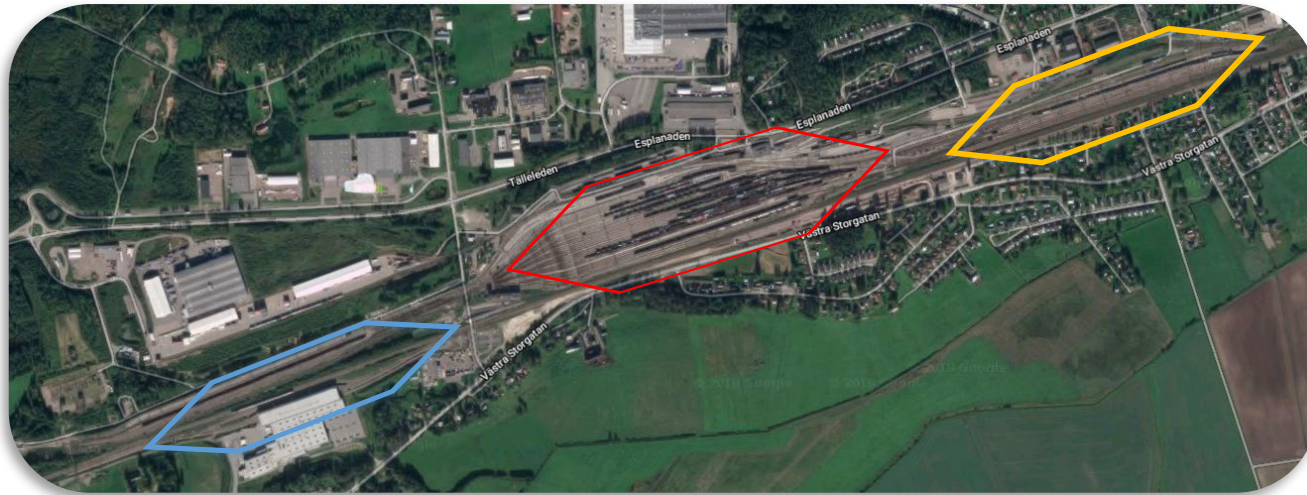




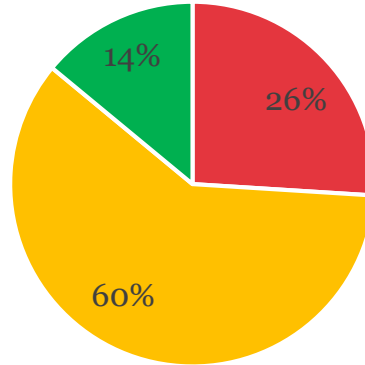
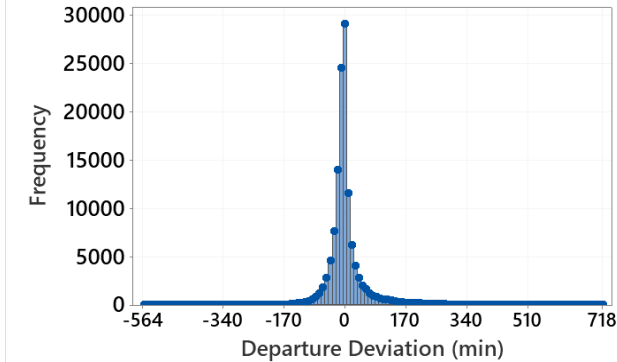
# Yard-Network in EU



# Yard Studies



# Specific Problem (Swedish Context): Yard Departure Prediction



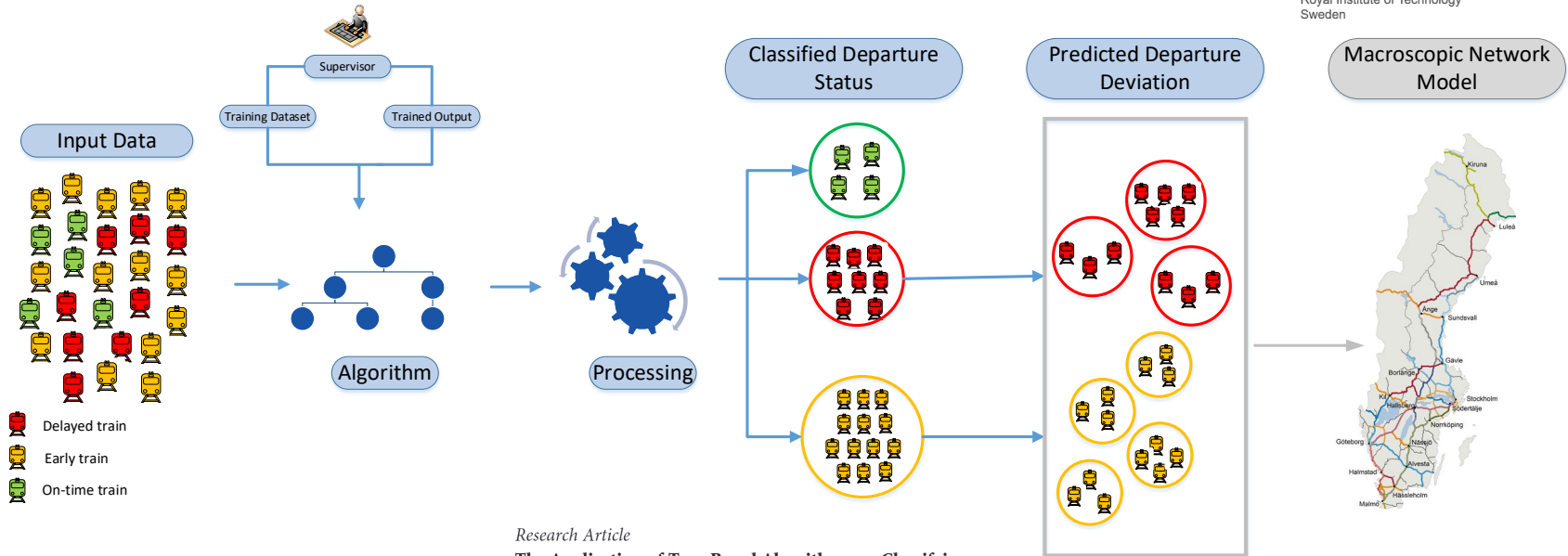
■ Delayed ■ Early ■ On Time



# Model Framework



Niloofer Minbashi, Markus Bohlin  
KTH  
Royal Institute of Technology  
Sweden



Research Article

### The Application of Tree-Based Algorithms on Classifying Shunting Yard Departure Status

Niloofer Minbashi <sup>1</sup>, Markus Bohlin <sup>1</sup>, Carl-William Palmqvist <sup>2</sup> and Behzad Kordejad <sup>1</sup>

<sup>1</sup>Division of Transport Planning, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden

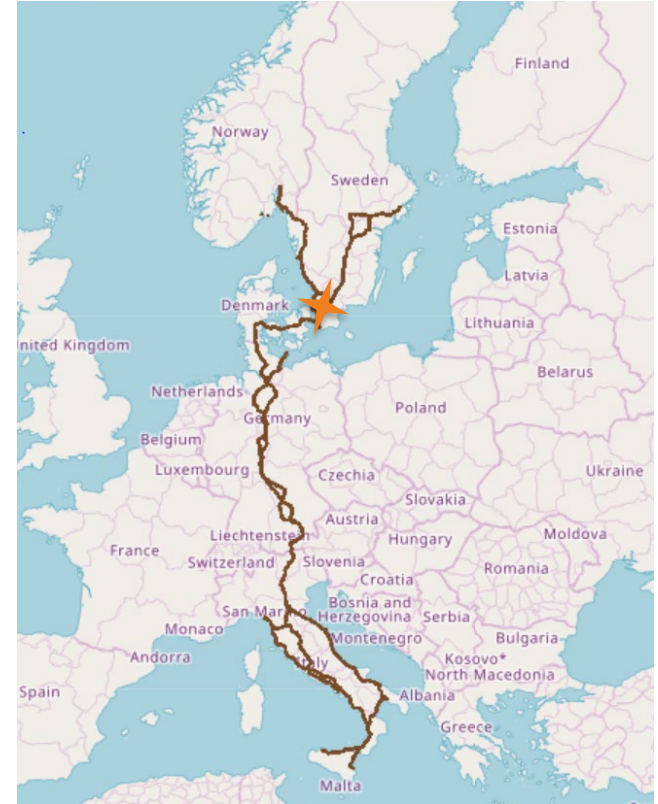
<sup>2</sup>Division of Transport and Roads, Lund University, P.O. Box 118, 221 00 Lund, Sweden



# Case Study



Malmö Yard Layout



Scan-Med Corridor





# Data

- One year data (2019)
- Three datasets:
  - Wagon connection
  - Train Punctuality
  - Train Features
- 161 162 wagon connections
- 10% duplicates in the data: 144 362 wagon connections
- **30 548 train connections**
- 8 156 train departures

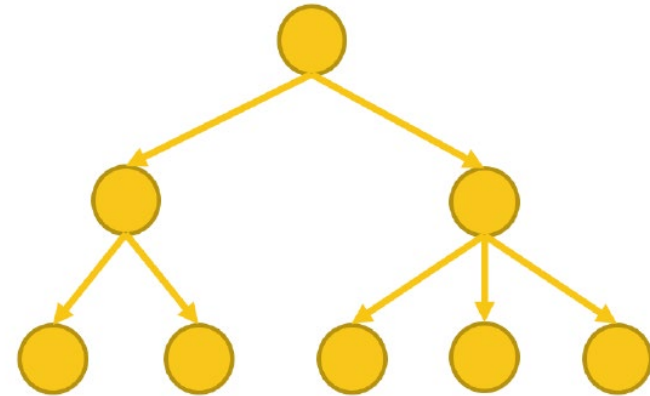


# Predictors

Wagon Connection	Min Wagon Dwell Time
	Max Wagon Dwell Time
	Number of Wagons per Departing Train
	Number of Arriving Trains to a Departing Train
Train Punctuality	Departing Train Number
	Scheduled Departure Hour
	Departure Week-Day
	Departure Month
Train Features	Maximum Operated Length
	Maximum Operated Weight

# Model

- Decision Tree Algorithm
- Gradient Boosted Trees Algorithm
- Random Forest
- Baseline model (Median)
- 10-fold cross-validation
- Evaluation Criteria:
  - R-Squared
  - Mean absolute error
  - Mean squared error
  - Root mean squared error
  - Mean signed difference



A Single Decision Tree



# Results

Model	R-Squared	Mean Absolute Error (min)	Mean Squared Error	Root Mean Squared Error (min)	Mean Signed Difference
Base Line (Median)	-	19	1584	40	-
Simple Decision Tree	0.86	2.6	200.38	14.15	0.79
Gradient Boosted Trees	0.89	2.8	158.5	12.6	0.08
Random Forest	0.92	2.9	117.13	10.83	0.12



# Conclusion

- The amount of data matters in this problem.
- Simple approach to be used by both the IM and the YM.
- Not much difference between a single tree and an ensemble version for departure deviation.
  
- **Future Direction:**
- Running the model on data from other yards
- Connecting the model to a macroscopic network model



# Thanks for your attention!



**Niloofar Minbashi**

PhD Candidate

Division of Transport Planning

KTH Royal Institute of Technology

[minbashi@kth.se](mailto:minbashi@kth.se)

