

Matej Jusup

matej.jusup@gmail.com | mjusup@ethz.ch | +41764505807

DOCTORAL RESEARCH

- **ETH Zurich, Institute for Transport Planning and Systems** Zurich, Switzerland
PhD *Sep 2020 – present*

THESIS TITLE: Real-time Railway Rescheduling under Uncertainty using Learning Algorithms
SUPERVISOR: Prof. Dr. Francesco Corman
KEYWORDS: Reinforcement learning | Approximate dynamic programming | Stochastic optimization
Railway optimization

EDUCATION

- **University of Zagreb, Department of Mathematics** Zagreb, Croatia
MSc in Mathematical Statistics; graduated with honours *Oct 2013 – Feb 2017*
- **University of Bielefeld** Bielefeld, Germany
Erasmus exchange student *Sep 2015 – Jul 2016*
- **University of Zagreb, Department of Mathematics** Zagreb, Croatia
BSc in Mathematics *Oct 2010 – Jul 2013*

RELEVANT COURSES: Advanced probability | Mathematical statistics | Applied statistics | Stochastic processes | Markov chains | Time-series analysis | Numerical analysis | Operations research | (Non) Linear and convex optimization | Data structures and algorithms
Database systems | Computer architecture | Combinatorics

ONLINE COURSES: (Graph) Neural networks | Computer vision | Support vector machines | Random forests | Natural language processing | Gradient boosting | Clustering algorithms

PUBLICATIONS

1. M. Jusup, A. Trivella, F. Corman (2021), A review of real-time railway and metro rescheduling models using learning algorithms, In 30th International Joint Conference on Artificial Intelligence (IJCAI-21).
<https://rl4its-ijcai21.github.io/workshop/>

TALKS AT CONFERENCES AND WORKSHOPS

- **STRC 2021 – 21st Swiss Transport Research Conference** Monte Verità, Switzerland
Sep 2021
- **IJCAI 2021 – RL for Intelligent Transportation Systems Workshop** Montreal, Canada
A Review of Real-time Railway and Metro Rescheduling Models using Learning Algorithms *Aug 2021*
- **DevArena – software development conference (invited)** Zagreb, Croatia
Machine Learning - From Idea to Production *Oct 2019*

MASTER THESIS

- **Network Optimization in Railway Transport Planning**
Thesis link

SUPERVISORS: Prof. Dr. Sc. Marko Vrdoljak | Prof. Dr. Sc. Andreas Dress
SUMMARY: In-depth theoretical discussion of the approaches for representing and solving railway timetabling problem for a single-line network
KEYWORDS: Graph theory | Minimum-cost flow problem | Linear programming | (Revised) Simplex method | Multi-commodity flow problem | Lagrangian relaxation | Column generation

COLLABORATIONS

- **Collaboration with Norbert Fogarasi**

C++ implementation of the algorithm proposed in the "On Partial Sorting in Restricted Rounds" - link

GOALS: Improve Norbert's Matlab implementation and generate publication results

OUTCOME: *Space complexity* - reduced $\mathcal{O}(n^2 n!)$ to $\mathcal{O}(n^2)$

of test cases - $n!$

Time complexity - $\mathcal{O}(n^4)$ per test case

Execution time - use of lookup tables made 4x improvement

EXPERIENCE

- **Cantab Predictive Intelligence**

Zagreb, Croatia

Senior Data Scientist - team leader

Mar 2019 - Jul 2020

E-Commerce Recommender System: Recommending products to potential as well as active shoppers

GOALS: Develop a recommender system which can attract new shoppers and give personalized recommendations to active shoppers

APPROACH: In progress

Delivery Delay Estimation: Estimating delivery delays for shopping malls during COVID-19 pandemic

GOALS: Develop a model which can predict delays given purchase related information

APPROACH: Given courier service capacity and the list of orders, estimate potential delays within given time period

TECHNOLOGY: Python | ARIMA models

OUTCOME: Used by operations department of a client providing ordering application
Upon placing orders, customers are notified about potential delays

Behavioural Credit Scoring: Developing the model and system for assessing consumer's default risk

GOALS: Develop a model which relies on transactional information only

Develop a service which can handle a huge data-load

APPROACH: Extract useful features from consumer's transactions

Find the best model according to Gini metric

Develop a service which can be offered to various banks

TECHNOLOGY: Python | PySpark

OUTCOME: We reported one of the best scores on the market, Gini up to 75%

We presented it on a couple of conferences, most notably Credit Scoring and Credit

Control Conference XVI in Edinburgh

Already in the test environment of two retail banks

Service is still under development

Personalized Newsletter Recommender System: Developed for a private equity fund

APPROACH: Generate features based on historical user behavior

Find the best model according to ROC AUC curve

TECHNOLOGY: PySpark | Azure | Databricks

ML-Driven Marketing Plan: Increased sales and patient share of a newly launched hearth disease drug for one of the leading pharmaceutical companies

GOALS: Improve the intensity and reach of calls and meetings with doctors

Find the optimal number of congresses, one-to-one meetings and other relevant factors

APPROACH: Find market segments/clusters

Predict sales on the segments

Find optimal number of calls for each segment

TECHNOLOGY: Pandas | Numpy | Scikit-learn | Matplotlib | Seaborn | OR-Tools

OUTCOME: Sales in the test sample where recommendations were applied is around 10% higher than in the control group

- **Morgan Stanley**

Budapest, Hungary

Data Scientist

Oct 2017 - Mar 2019

Systemic Risk Model Execution Efficiency: Improved the execution time of a mixed-integer convex programming model used for better understanding of the risk structure

GOALS: Execution should be under 3 minutes

Solution should be as close as possible to the optimum in given time

APPROACH: Probabilistic methods like simulated annealing, random-restart hill climber (RRHC), and problem-specific heuristic were implemented and compared
TECHNOLOGY: Python | Cplex solver | Google OR-Tools
TESTS: Mathematical proof that the heuristic converges to the optimum
Benchmark against a set of optimal solutions generated by solvers
OUTCOME: RRHC outperformed other approaches | On average 5% from the optimum | Worst case 15%

E-Trading Execution Limits Calibration: Examined e-trading clients and calibrated a model which blocks real-time executions in case of a high risk

APPROACH: Statistical analysis of historical executions
Model parameters were calibrated to capture risky behaviors
TECHNOLOGY: Pandas | Numpy | Matplotlib | Sybase | DB2
TESTS: Historical backtests | Tracking real-time trading | Formal proof of parameters stability in a limit

Treasury Department Cash Traceability: Built uncollateralized debt tracking system

APPROACH: Combining various daily feeds to produce the firm-wide reporting system
TECHNOLOGY: *PyQ kernel* - Q/kdb+ | Python | SQL

Listed Derivatives Liquidity: Created proof of concept for data-driven liquidation model based on intraday futures trading

TECHNOLOGY: Q/kdb+ array programming language specialized for time-series analysis

Software Developer

TECHNOLOGY: Java | Spring | Google protobuf | JUnit | JBehave

Budapest, Hungary
Dec 2016 - Oct 2017

Technology Analyst Program

RELEVANT COURSES: C++ | Java | Scala | JavaScript | Operating systems

New York & London
Aug 2016 - Dec 2016

- **University of Zagreb, Department of Mathematics**
Junior Teaching Assistant - Euclidean Spaces

Zagreb, Croatia
Oct 2013 - Mar 2014

PROGRAMMING SKILLS

Advanced: Python | C++
Intermediate: PySpark | Q/kdb+

Basic & University experience: Java | JavaScript | C | R | Matlab
VCS & Other: AWS | MS Azure | Git | Perforce | SQL

LANGUAGES

English: Professional working proficiency
Croatian: Native proficiency

INTERESTS AND AWARDS

Chess: Silver medalist at individual Croatian junior (under 20 years) championship in 2011
Member of a Reti club participating in the 1st Swiss league
Current ELO rating 2267

REFEREES DETAILS

Prof. Dr. Francesco Corman | *publications* | francesco.corman@ivt.baug.ethz.ch | +41446333350
Prof. Dr. Sc. Siniša Slijepčević | *publications* | sinisa.slijepcevic@math.hr | +38514605895
Prof. Dr. Sc. Marko Vrdoljak | *publications* | marko.vrdoljak@math.hr | +38514605736