MSc Projects for Smart Mobility

Machine Learning for Smart Mobility group

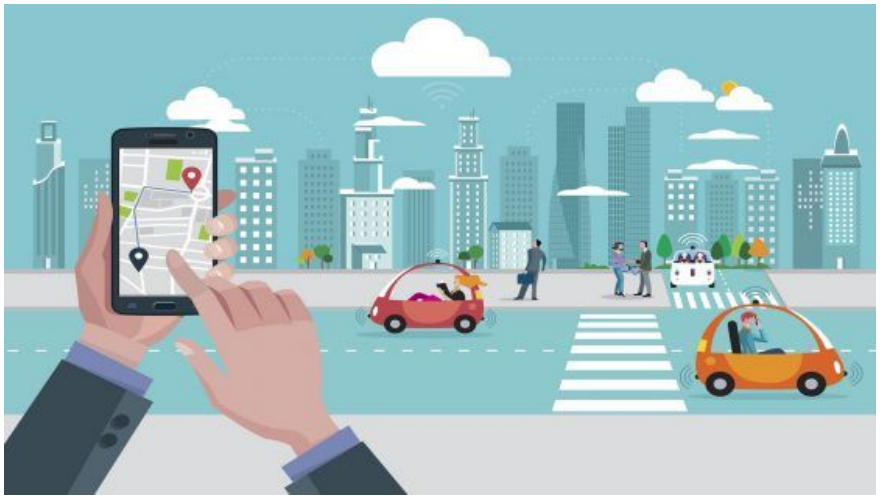
DTU Technical University of Denmark

Transport Modelling Division, Department of Management Engineering

Table of Contents

IMPORTANT: This document contains general project ideas. If you have your **own ideas** about projects or **variations** of the projects described in this document, feel free to **get in touch with us**!

1. Mobility-on-Demand Meets Machine Learning



https://bestmile.com/bestmile-part-of-mobility-nation-working-group-planning-the-mobility-evolution/

PROJECT ADVISORS: Francisco Pereira (camara@dtu.dk), Filipe Rodrigues (rodr@dtu.dk), Stanislav Borysov (stabo@dtu.dk), Jeppe Rich (rich@dtu.dk)

KEYWORDS: Machine Learning; Data Science; Demand Prediction

BACKGROUND: Mobility-on-Demand (MoD) is an innovative concept in transportation field where the mobility service is provided on a request from a customer. Nowadays, we witness a growing number of companies operating in Denmark and worldwide which are based on this concept, such as DriveNow, Green Mobility or Donkey Republic. The shared mobility services provided by these companies allow for efficient utilization of transport network and vehicles hence contributing to sustainable development of modern cities. However, being highly responsive to customer needs, this business model more than ever relies on accurate prediction of short- and long-term travel demand, defining business areas for further expansion, fleet rebalancing, etc.

PROJECT DESCRIPTION: The project is based on the data provided by two large car-sharing companies operating in Denmark: DriveNow and Green Mobility. The project is aimed at increasing efficiency of the current fleet and understanding new business opportunities. The project might involve one or several following directions:

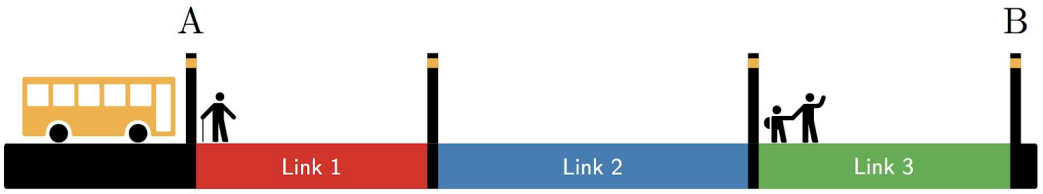
1. Demand prediction using modern deep learning techniques, probabilistic graphical models or other spatio-temporal models.
2. Survival analysis to predict idle time of cars.
3. Identifying business areas for further expansion.
4. Your own ideas are also welcome :)

PREREQUISITES: Good programming skills in Python; Basic knowledge of machine learning and data mining.

LANGUAGE: English.

NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

2. Uncertainty Estimation in Neural Networks for Bus Arrival Time Prediction



PROJECT ADVISORS: Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Assistant Professor Stanislav Borysov (stabo@dtu.dk), Professor Francisco Pereira (camara@dtu.dk)

EXTERNAL ADVISOR: Niklas Petersen (ncp@moviatrafik.dk)

KEYWORDS: Prediction Uncertainty; Bayesian Neural Networks; Deep Quantile Regression

BACKGROUND: Bus arrival times are well known to vary dramatically due to various contextual factors. Even for repeating conditions (e.g. same day of the week, same type of day), bus arrival times can change significantly due to traffic lights, unexpected dwell times, etc. Hence, bus arrival time prediction are inherently uncertain. Therefore, providing single point-estimates (e.g. “the bus will arrive at 13:20”) for bus arrival times is insufficient. Instead, this project proposes the use of prediction intervals (e.g. “the bus will arrive between 13:16 and 13:23 with 95% certainty”), thus transmitting uncertainty information by providing a time interval within each the bus is very likely to arrive. From a user's perspective, knowing how uncertain an arrival time estimate is constitutes a very important piece of information for travel decision making. For example, an underestimated arrival time can cause the user to miss the bus, forcing him/her to wait for potentially long periods for another one or even to search for a different alternative, while if he/she knew in advance that the arrival time prediction was likely to be wrong (high uncertainty), he/she might have opted for a more reliable alternative. Similarly, an overestimated arrival time can cause users to wait for long periods at the stop, unnecessarily.

PROJECT DESCRIPTION: This project proposes the use of state-of-the-art approaches for estimating uncertainty in bus arrival times predictions. Possible approaches includes the use of Bayesian deep learning methods, normalizing flows or more direct approaches based on quantile regression techniques to jointly estimating multiple quantiles for the bus arrival times. For this purpose, we will use data from the bus provider for Copenhagen (Movia). The prediction intervals by the proposed model can then be transmitted directly to the user, in order to enable more informed decision making, or used internally by travel information services/apps in order to deprioritise unreliable travel alternatives, guide users into making better choices and ultimately improve the user’s perception of the bus public transport and overall satisfaction.

PREREQUISITES: Good programming skills in Python; Statistics; Basic knowledge of machine learning and data mining.

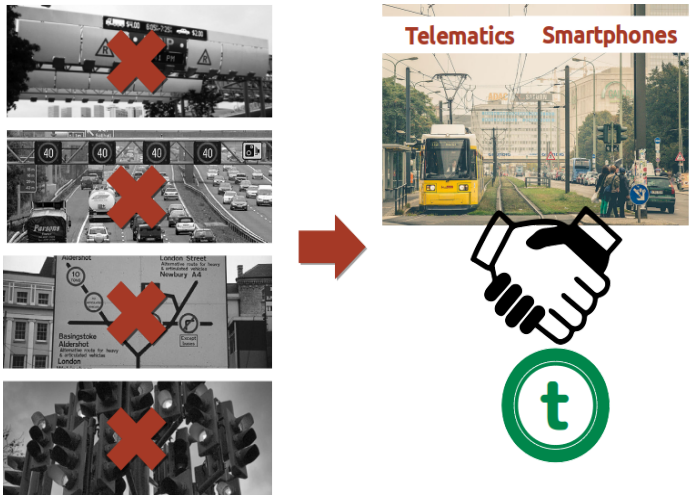
LANGUAGE: English.

NUMBER OF STUDENTS: 1 master student (1 projects of 30-35 ECTS Credits).

3. Coordination with Mobility Tokens: Tradable Electronic Permits for Mobility Management

PROJECT ADVISORS: Carlos Lima Azevedo (climaz@dtu.dk)

KEYWORDS: mobility management; tradable permits; dynamic control.



BACKGROUND: Congestion in any transportation system is ultimately due to a problem of limited space which, in the end is a scarcity problem that can be dealt with a pricing, quantity control or a combination of both. Tradable permit schemes are such a combined system where a regulator provides an initial endowment of mobility permits to all potential travelers. In order to use a transportation system, the road, the public transportation or a shared system (such as car or bike sharing) users need to spend a certain number of permits (i.e.: tariff) that could vary with the travel conditions on the specific mobility alternative used (congestion, pollution, etc.). If needed, the permits can be bought and sold in a market at a price that is determined by demand and supply interactions. While many countries, including Denmark, are considering the deployment of road congestion pricing mechanisms, academic studies that (1) propose and conceptualize the policy and economics of mobility tradable permit schemes; (2) formulate a mathematical approach for understanding the user and market equilibrium under different theoretical assumptions; and (3) empirically investigate the individual behavior under such schemes have increasingly being put forward in the research field. The theoretical gains in terms of wealth distribution, dynamic adaptation and efficiency have been shown, but what does it take to actually make it happen?

PROJECT DESCRIPTION: Having in mind the understanding of the practical gains and the application feasibility of Mobility Tradable Tokens, two different MSc thesis fall under the larger umbrella of this project:

1. Develop and numerically assess a model for the demand side of Mobility Tradable Tokens from existing and new formulations of market equilibrium, purchase and sales of permits and travel behaviour;
2. Develop and numerically assess a model for the supply side of Mobility Tradable Tokens from existing and new formulations by constructing a dynamic tariff control system and a network simulation framework;
3. Your own ideas are also welcome :)

PREREQUISITES AND INTERESTS: For 1.: economic, game theory or discrete choice modelling; For 2.: network theory, control theory and simulation. Knowledge in transport modelling and Python coding is a plus.

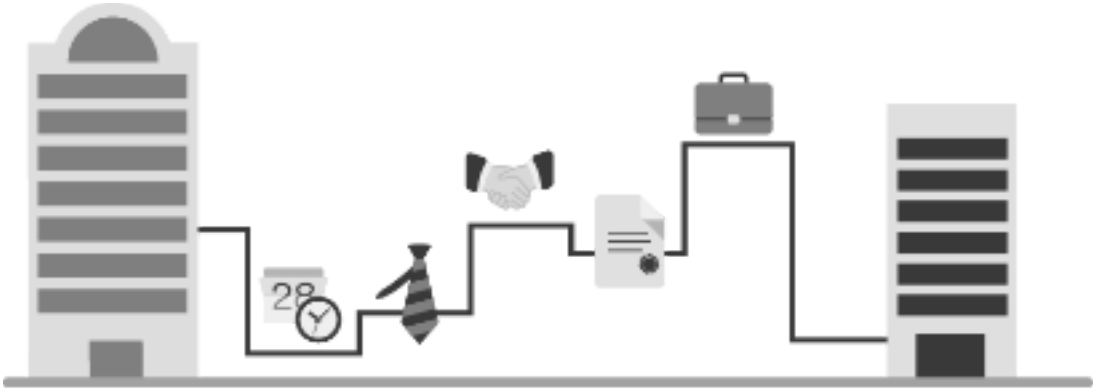
LEARNING OBJECTIVES:

- Understand the fundamental behavior and mechanisms behind a Tradable Electronic Permits schemes
- Demonstrate the capacity to conceptualize demand or supply modelling frameworks for dynamic control systems, deploy them and assess them with numerical experiments
- Leverage your algorithms and coding skills along with current international research, to develop a new solution
- Be able to work, communicate and report research results orally and in practice and academic oriented written form in Danish and English.

LANGUAGE: Communication: English. Thesis: English or Danish.

NUMBER OF STUDENTS: 1 or 2 master student (1 or 2 project of 30-35 ECTS Credits).

4. Human Activity Embeddings



PROJECT ADVISORS: Professor Francisco C. Pereira (camara@dtu.dk), Assistant Professor Filipe Rodrigues (rodr@dtu.dk)

KEYWORDS: Representation Learning; Deep Learning; Word Embeddings; Human Activity Data; Travel Diaries

BACKGROUND: For multiple reasons, including little data and computing power, and a traditionalist community, representations of models and people (or, "agents") in transport demand behavior have changed very little through the decades. Among this is the representation of categorical variables (such as "gender", "education level", or "city"), which usually end up as "dummy variables" in the models. However, new paradigms exist that have been under-explored, including natural-language-based. In natural language, virtually everything is "categorical" in nature (e.g. words), and a new and very successful representation approach has been using "text embeddings" (search for word2vec in Google!).

A text embedding is a deep neural network based highly abstract representation of data, and it has been behind many well known tools, such as Google Translate or even autocomplete functionalities.

PROJECT DESCRIPTION: The goal of this project is to develop and test embedding mechanisms using mobility behavior data, such as smartphone-based mobility data or data from the national travel survey that DTU has in the Center for Transportation Analytics (https://www.cta.man.dtu.dk/transportvaneundersoegelsen). Such datasets include information about people's travel choices (where, when, how, etc.) and activity patterns (e.g. Home -> Work -> Pickup Children -> Gym -> Home). This project aims at exploiting this rich data to learn lower-dimensional representations of human activities and other categorical variables, in an analogous way to "word embeddings" (e.g. word2vec) re-represent in a semantically-meaningful way.

PREREQUISITES: Knowledge of Python and Statistics.

LANGUAGE: English.

NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

5. Deep Spatio-Temporal Models for Mobility Prediction

PROJECT ADVISORS: Assistant professor Stanislav Borysov (stabo@dtu.dk), Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Professor Francisco Pereira (camara@dtu.dk)

KEYWORDS: Deep Learning; Travel Demand Prediction; Spatio-Temporal Models

BACKGROUND: Successful implementation of new and disruptive technologies in the transportation field, such as autonomous vehicles or demand-responsive public transport, will highly depend on predictability. Failing predictions beyond acceptable rate is fatal for many scenarios of future mobility because it may deliberately mislead their optimization mechanisms to misalign supply with demand. However, prediction of mobility patterns is known to be a highly non-trivial task since small perturbations like traffic incidents or special events can substantially affect system's behaviour as a whole.

Recent interdisciplinary developments from machine learning, statistics, neuroscience and other research fields together with exponential growth of computational power and amount of data available have opened new ways to deal with such complex prediction tasks. The new framework, known as "deep learning", has made a clear breakthrough in processing high-dimensional data, like images or natural language, allowing computers to resolve underlying highly nonlinear dependencies. This project aims to bring this exciting progress into the transportation systems modelling realm.

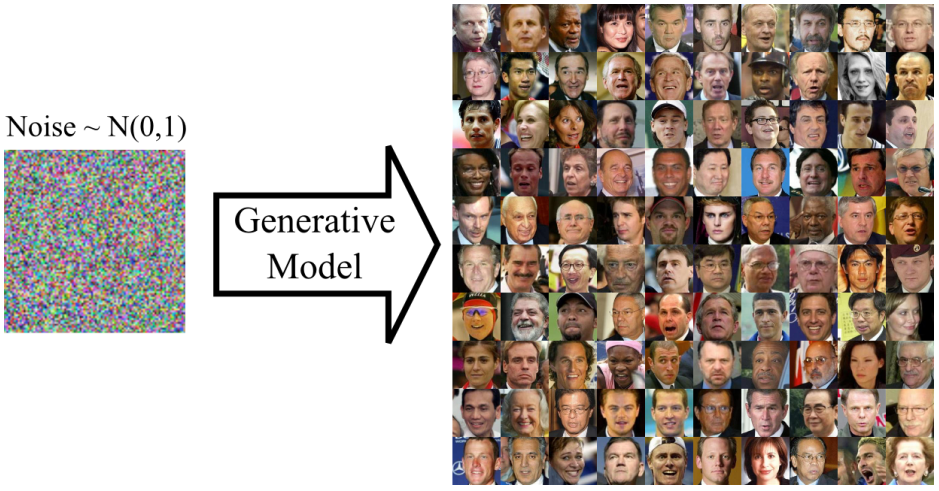
PROJECT DESCRIPTION: The main goal of the project is to develop next generation travel demand prediction models using (but not limited to) deep learning techniques, such as recurrent and convolutional neural networks, using historical data available. Particular focus will lie on analysis and prediction of spatio-temporal patterns in transport networks under various stress scenarios caused by traffic incidents, special events, extreme weather conditions, etc. This project is strongly aligned with the current research between DTU, Danish Road Directorate, Google and Copenhagen Municipality.

PREREQUISITES: Good programming skills in Python; Basic knowledge of machine learning and data mining.

LANGUAGE: English.

NUMBER OF STUDENTS: 1-2 master students (2 projects of 30-35 ECTS Credits).

6. Deep Generative Modeling for Mobility



http://hunterheidenreich.com/blog/what-is-a-gan/

PROJECT SUPERVISOR: Stanislav Borysov (stabo@dtu.dk)

PROJECT ADVISOR: Jeppe Rich (rich@dtu.dk), Francisco Pereira (camara@dtu.dk)

KEYWORDS: Machine Learning; Deep Learning; Generative Modeling; Transportation Modelling; Urban Planning

BACKGROUND: Generative modeling is a subfield of statistics and machine learning which focuses on estimating the joint probability distribution of data. The estimated joint distribution can be used for a wide range of tasks, such as generation of synthetic samples, data imputation or addressing privacy issues. Traditional generative models based on probabilistic graphical models have suffered from scalability issues, which made them applicable either to small problems or required the introduction of simplifying assumptions. In the recent years, however, deep generative models, such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs) and Recurrent Neural Networks (RNNs), have opened a path towards large scale problems. These models, which combine deep neural networks and efficient scalable inference algorithms, have proven to be effective as a mean to model high-dimensional data such as images, sound or text. This project aims at extending this progress into the transportation modeling area.

PROJECT DESCRIPTION: The main goal of the project is to apply modern deep generative modelling techniques to one of the transportation-related problems, including (but not limited to) population synthesis, generation of travel schedules, or even generation of road networks and land use. Unlike the recent focus in the field on image data, the project aims at application of deep generative models for other types of high-dimensional data represented by tables, time series or graphs. The outcomes of his project will contribute to the scalability, generalizability, and interpretability of modern transportation research and urban planning.

PREREQUISITES: Good programming skills in Python, basic knowledge in machine learning and data mining.

LANGUAGE: English.

NUMBER OF STUDENTS: 1-2 master students (2 projects of 30-35 ECTS Credits).

7. Next Generation Traffic Control Systems: Reinforcement Learning for Traffic Signal Control

PROJECT ADVISORS: Carlos Lima Azevedo (climaz@dtu.dk), Filipe Rodrigues (rodr@dtu.dk)

KEYWORDS: Dynamic Control Systems; Deep Reinforcement Learning; Traffic Simulation.

BACKGROUND: Traffic control signals is an essential solution for increased traffic safety and efficiency for road, bicycle and pedestrian infrastructures. To achieve optimum efficiency, traffic signals usually measure dynamic traffic variables and adjust its own control variables. Current practice relies on the collection of dynamic traffic volumes, usually from cameras or sensors on the road infrastructure, and on heuristics with restrictive pre-defined configurations (e.g.: phases, plans, etc.). Additionally, for signals that are closely spaced, it is necessary to coordinate the green time so that vehicles may move efficiently through the set of signals. Despite the widespread of traffic signals globally, practice still uses sub-optimal frameworks from decades go, resulting in significant losses for society. Delays at traffic signals account for up to 10% of all traffic delays in the United States for example. Following current trends, the increasing number of traffic lights for handling traffic growth has raised the importance for alternative approaches for traffic control world wide. Deep reinforcement learning has given the first signs of promising results in traffic control and further developments are expected to allow near-future significant gains in our transportation system.

PROJECT DESCRIPTION: In this project you will focus on the development of coordinated traffic signal control algorithms using deep reinforcement methods and comparing with traditional fixed, actuated, adaptive and coordinated systems in a traffic simulated environment. The student can choose to focus on any (or all) of the following aspects of the project:

1. Understanding and implementing a selected adaptive traffic control algorithm from the literature and evaluate its performance in a urban traffic simulation environment;
2. Implementing a new improved deep reinforcement learning formulation for the control of a single intersection;
3. Design and develop a coordination framework for deep reinforcement learning based traffic signals and test it in a larger urban simulated network.

The student will develop all tasks in Python. Two existing models transportation systems will be used as case studies using the AIMSUN traffic microscopic simulation (www.aimsun.com). Traffic data is available for the needed assessment.

PREREQUISITES: Programing in Python. Knowledge in machine learning is also required. Understanding in simulation principles and transportation models are preferred but not required.

LEARNING OBJECTIVES:

- Understand fundamental concepts of deep reinforcement learning and its application as a control solution
- Extend traditional deep learning formulations to multi-level control problems
- Identify and apply key concepts of traffic simulation and dynamic control systems
- Compare and support the design and implementation choices with simulation based experiments
- Leverage your algorithm and coding skills along with current international research, to develop a new solution
- Be able to work, communicate and report research results orally and in practice and academic oriented written form in English.

LANGUAGE: English.

NUMBER OF STUDENTS: 1 or 2 master students (1 or 2 projects of 30-35 ECTS Credits).

8. SimDenmark: A New Agent- and Activity-based Simulation Model for the Nation

PROJECT ADVISORS: Carlos L. Azevedo (climaz@dtu.dk), Thomas K. Rasmussen (tkra@dtu.dk), Jeppe Rich (rich@dtu.dk)
KEYWORDS: Agent-Based Simulation; Transportation Modelling; Large-scale Software

BACKGROUND: With advances in communications and computational power, large-scale integrated models are becoming increasingly common and complex. These models are typically composed of several loosely interconnected models, each with a specialized purpose. They are fundamentally characterized by the generation of trips from actual activity participation of millions of agents that represent a population of a city and a detailed multi-modal representation of the transportation system dynamics. These models have allowed for unprecedented studies of smart mobility, such as the impact of the potential deployment of automated vehicles, new public transportation pricing strategies or even of behavioral shifts. Currently, integrated agent- and activity-based models can be found for a small number of pioneer cities such as Boston, Zurich and Singapore. SimMobility (https://its.mit.edu/software/simmobility) is such an integrated agent-based simulation platform, developed by the Massachusetts Institute of Technology. DTU is launching the development of a new SimMobility model for Denmark, integrating state-of-the-art models and apply it for the first time to Danish future mobility scenarios. For this project its midterm (MT) module is of key interest as it models and simulates agents behavior in terms of activity, travel plans and actions together with a multi-modal mesoscopic network simulator.

PROJECT DESCRIPTION: In this project you will be the first to join the SimDenmark team and leading the development of one of the several modules of this state-of-the-art simulation model using several data sources from Denmark:

1. **Demand Modelling:** Estimation and Simulation of a new activity-based model
2. **Network modelling:** Modelling and Simulation of a new multi-modal network model
3. **Software Engineering:** Tackling the modularity and scalability of the large-scale simulation

PREREQUISITES: Knowledge in transport modelling (for 1 and 2) and software engineering (and 3) is required. The students will be required to read and write in C++ and SQL.

LEARNING OBJECTIVES:

- Understand fundamental concepts of integrated transport simulation
- Organize, analyse and process large datasets of mobility related data
- Estimate and validate complex models using demand and network modelling techniques
- Practice in software engineering for large-scale parallel simulation architectures
- Work, communicate and report research results orally and in practice/academic oriented written form

LANGUAGE: Danish or English.

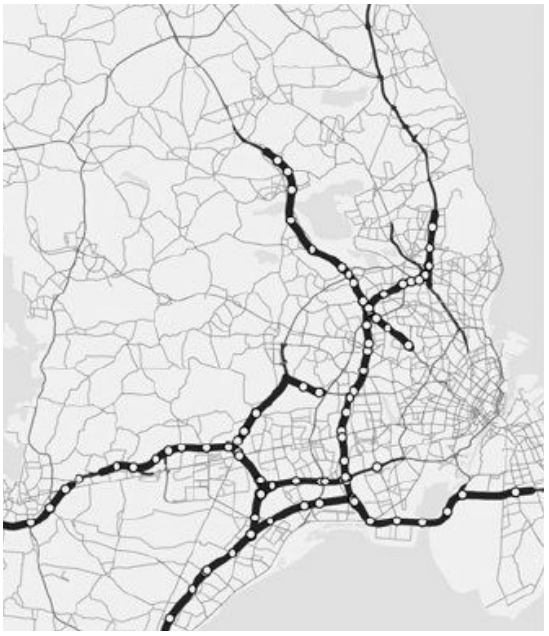
NUMBER OF STUDENTS: 1 to 5 master student (1 to 5 project of 30-35 ECTS Credits).

9. Graph Convolutional Recurrent Neural Networks for Traffic Forecasting/Bus Travel Time Prediction

PROJECT ADVISORS: Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Assistant Professor Stanislav Borysov (stabo@dtu.dk), Professor Francisco Pereira (camara@dtu.dk)

EXTERNAL ADVISOR: Niklas Petersen (ncp@moviatrafik.dk)

KEYWORDS: Deep Learning; Spatio-temporal Models; Graph Convolutions; Congestion Propagation.



BACKGROUND: Forecasting traffic conditions is essential for developing intelligent transportation systems, because it provides real-time and anticipatory information about the performance of the network. However, the complex spatio-temporal dependencies over the road network and inherent stochasticity of traffic make it very challenging to forecast, particularly when considering longer-term horizons. For similar reasons, the related problem of obtaining accurate bus travel time predictions can also be extremely challenging. However, addressing that problem is crucial for improving bus operations, managing fleets, and providing reliable information to the users so that they can make better decisions about their travel.

For both these problems, a key challenge consists of capturing spatio-temporal correlations (over the network and over time) in the data. Due to recent advances in machine learning and modern deep learning techniques, there already exist powerful models than can capture such king of dependencies. Recurrent Neural Networks (RNNs) are great for modelling temporal data, and Convolutional Neural Networks (CNNs) are great at capturing spatial patterns (e.g. between pixels in an image). However, these structures do not generalize trivially to road networks (directed graphs). Very recently, there have been some research efforts in generalizing these ideas to graphs (e.g. Graph CNNs and Diffusion CNNs), which have the potential to model congestion propagations patterns in the road network.

PROJECT DESCRIPTION: The goal of this project is to explore these novel deep neural networks architectures based on graphs (or develop new improved ones that are specially tailored for the particular application) in order to model mobility data. Namely, we suggest two possible applications:

- Traffic forecasting (e.g. speeds or traffic flows) using probe-vehicle data from INRIX [1] or Google for Copenhagen.
- Bus travel time prediction using data from the bus operator in Copenhagen - Movia [2]. In this case, the project with be in close collaboration with Movia and, if successful, it has the potential of eventually being integrated in the actual system used by Movia to provide arrival time estimates at bus stops in Copenhagen.

PREREQUISITES: Good programming skills in Python; Basic knowledge of machine learning and data mining.

LANGUAGE: English.

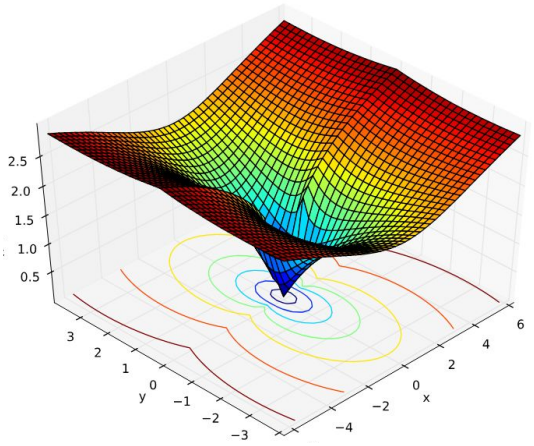
NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

[1] http://inrix.com/
[2] https://www.moviatrafik.dk/

10. Bayesian Optimization for Simulation Calibration

PROJECT ADVISORS: Carlos Lima Azevedo (climaz@dtu.dk)
Filipe Rodrigues (rodr@dtu.dk)
EXTERNAL ADVISOR: Tamara Djukic (taamaara@gmail.com)

KEYWORDS: Bayesian Optimization; Traffic Simulation; Calibration; Machine Learning



BACKGROUND: Each time a new data set is available to transportation modellers, an existing or new model needs to be fine-tuned. For example, when a new metro station is being designed and a traffic model is developed to assess the impact the station will have on road traffic, the model needs to be calibrated for the existing conditions first, allowing to proof its capability of data replication; similarly, when a new traffic management system, such as traffic lights, is being designed calibrated simulations are again used as the main support tool. The calibration of such transport simulation models still relies on a manual process in practice, resulting in an overwhelming inefficiency of skilled human resources. The research community has stepped away from manual processes by proposing several algorithms (such simulated annealing, gradient approximation, genetic algorithms, etc). Yet these methods still have computational, implementation and transferability constraints that have limited its applicability to the market.

PROJECT DESCRIPTION: In this project we target two main goals: (1) to develop a Bayesian Optimization algorithm for the generic calibration of traffic simulation models and (2) demonstrate its applicability through a proof of concept focused on the development of a flexible interface to run and assess the calibration process.
The project has the following tasks (to be selected depending on the number of students):

4. Understanding and implementing an existing calibration algorithm from the literature;
5. Building a flexible a GUI interface to operate the algorithm with two case-study simulation platforms models;
6. Develop a new Bayesian Optimization based algorithm for the calibration of traffic models, integrate it in the developed interface and compare with the existing algorithm.

The student will develop both tasks in Python and Java(Script). For this project, two existing models of Danish transportation systems in two different simulation applications will be used as case studies: an intersection model in the AIMSUN traffic microscopic simulation (https://www.aimsun.com/), and a highway model in the VISSIM traffic microscopic simulation (https://www.ptvgroup.com/en/solutions/products/ptv-vissim/). Traffic data for both sites is also available for this project. Additional info: https://arxiv.org/abs/1502.05700.

PREREQUISITES: Comfortable programing in Python and basic Java(Script). Knowledge of optimization algorithms and Bayesian statistics is helpful. Knowledge of simulation principles and transportation models is preferred but not required.

LEARNING OBJECTIVES:

- Understand the fundamental framework of calibration as an optimization problem, the modelling structure of a traffic simulation tool and its space of uncertainty
- Demonstrate the capacity to design, plan and build commercial level applications based on complex algorithms
- Leverage your algorithms and coding skills along with current international research, to develop a new solution
- Be able to work, communicate and report research results orally and in practice and academic oriented written form in Danish and English.

LANGUAGE: Communication: English. Thesis: English or Danish.
NUMBER OF STUDENTS: 1 or 2 master student (1-2 project of 30-35 ECTS Credits).

11. Multi-modal Visualization for Extracting Mobility Insights

PROJECT ADVISORS: Professor Francisco Pereira (camara@dtu.dk), Assistant Professor Stanislav Borysov (stabo@dtu.dk), Assistant Professor Filipe Rodrigues (rodr@dtu.dk)

KEYWORDS: Data Visualization; Uncertainty Visualization; Unexpected Behaviour

BACKGROUND: For a large city with a very vibrant and lively atmosphere such as Copenhagen, understanding how mobility demand changes with different factors such as weather conditions, road works or special events (e.g. music concerts, sports games, festivals, etc.) and how the transport systems reacts to these demand changes is of extreme importance to both transport operators/planners and users. However, analysing the complex patterns in mobility data can be a challenging task. Without proper tools, various important aspects of mobility can be hard to detect and communicate efficiently.

PROJECT DESCRIPTION: Aiming at addressing some of the aforementioned issues, this project proposes the development of a visualization for integrated analysis of mobility data from various sources and transportation modes (e.g. bus, car, bicycle, train, etc.). The main focus of this visualization will then be on extracting important insights regarding urban mobility and expressing then in a meaningful manner. For this purpose, it will have support for integrating mobility data from different transportation modes, weather and event data, and also data from predictions models, where particular emphasis will be put on visualizing prediction uncertainty and deviations from normal conditions. By doing so, the proposed visualization aims at allowing transport operators/planners to pinpoint extraordinary/unexpected situations and provide cues to the possible causes. By integrating data from multiple transportation models, it will allow for a unique possibility for extracting valuable insights regarding their interactions and how disruptions in one mode propagate to the others. Furthermore, it will provide city planners with an interactive tool for detecting problematic areas and conditions.

PREREQUISITES: Good programming skills, good knowledge of visualization techniques and GIS.

LANGUAGE: English.

NUMBER OF STUDENTS: 1 master student (1 projects of 30-35 ECTS Credits).

12. Machine Learning Approaches for Explaining and Anticipating Mobility Disruptions Caused by Special Events

PROJECT ADVISORS: Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Professor Francisco Pereira (camara@dtu.dk)

KEYWORDS: Deep Learning; Data Fusion; Mobility Disruptions

BACKGROUND: Special events like music concerts, sports games, festivals and parades are known to be able to cause severe impacts in transportation. Depending on the type of events, venue, popularity of the performers involved and other factors, the impact caused by a special event can vary in terms of magnitude, affected areas and transportation modes (e.g. bus delays, overcrowded subway or traffic congestion). Therefore, the inability to properly forecast the impact of special events well in advance can be very problematic for transport systems and, more importantly, for their users.

PROJECT DESCRIPTION: With this in mind, this project proposes to develop specialized prediction models for event areas that are able to predict in advance potential disruptions caused by special events. Based on recent advances on deep learning and multi-task learning, the proposed model will combine information from multiple events sources on the Web, together with data from different transportation modes. Namely, using bus data from Movia [1], floating car data from INRIX [2] and loop detector data from Copenhagen Municipality, this project aims at giving transport operators and planners the opportunity to mitigate disruptions caused by upcoming events (e.g. using nudging and specialized campaigns to promote mode-shift), by providing them with reliable forecasts of the transport conditions ranging different transportation models. As a proof of concept, the developed model will be focused on study areas in Copenhagen that are known to host events capable of causing significant disruptions, e.g. Royal Area, Telia Parken, DR Koncerthuset and Forum Copenhagen.

PREREQUISITES: Good programming skills in Python, knowledge of statistics, data mining and machine learning.

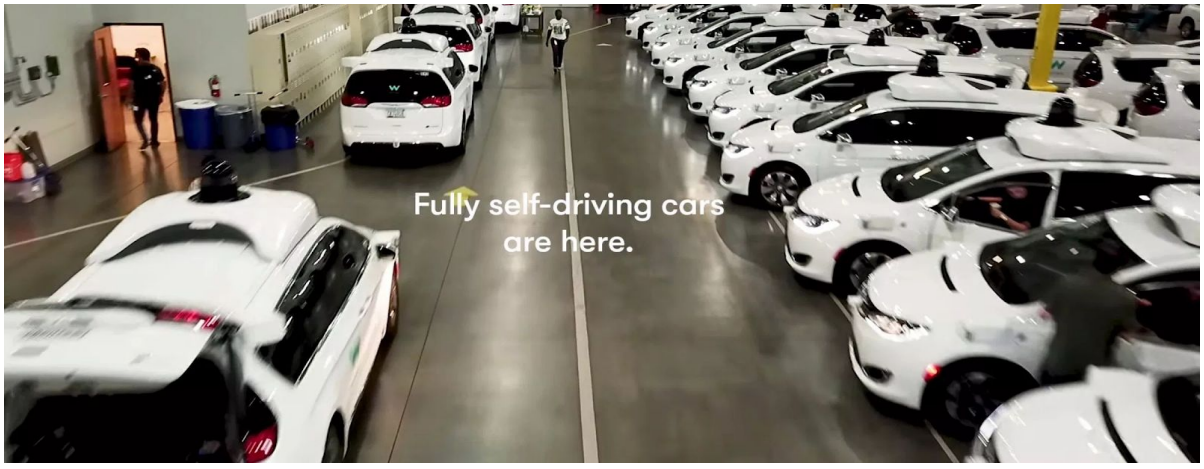
LANGUAGE: English

NUMBER OF STUDENTS: 1 or 2 master students (each can focus on a different transport mode).

[1] https://www.moviatrafik.dk/

[2] http://inrix.com/

13. On Demand Peer-to-Peer vs. Fleet-Based Automated Vehicles: A Simulation Approach



imagesource: Waymo

PROJECT ADVISORS: Carlos Lima Azevedo (climaz@dtu.dk), Jeppe Rich (rich@dtu.dk)

KEYWORDS: Automated Vehicles; Simulation; Fleet Management

BACKGROUND: The world of shared mobility is changing. Tesla announced that the owners of its future automated vehicles (AV) will be able to rent their cars in a shared on-demand service and current Teslas are already available along with other cars, by the hour, in Turo[1]. On the other hand, automated mobility giants such as Waymo or Delphi-Nutonomy are pushing for fleet based deployments of AV and have rolled out pilot tests in a few cities across the globe. While the mobility paradigm has clearly shifted towards having more and more options available to the end consumer, understanding the impacts of the different deployments AV allow is crucial, not only for the design and development of the new solutions, but also for policy assessment. Analysis on the changes in travel patterns and its consistent impacts on the performance of the solutions proposed and on the overall transportation system are of much need.

PROJECT DESCRIPTION: The main goal of the project is to design and develop a comprehensive multi-agent simulation tool for the flexible analysis of AV deployment scenarios. The student will have to grasp demand and supply aspects of smart mobility systems, including travelers and vehicle owners preferences, travel patterns, network modelling (inc. congestion) and system control management. Simple models for each component, including the operational scenario, will be designed based on existing literature and a framework for integration will be developed. A selected set of operational scenarios will then be used for testing and impact assessment. The outcomes of his project will contribute to the understanding of operational settings of AVs and tool development to support the design of future systems.

PREREQUISITES: Knowledge in network modelling, demand modelling or optimization is required. The students will be required to code in Python.

LEARNING OBJECTIVES:

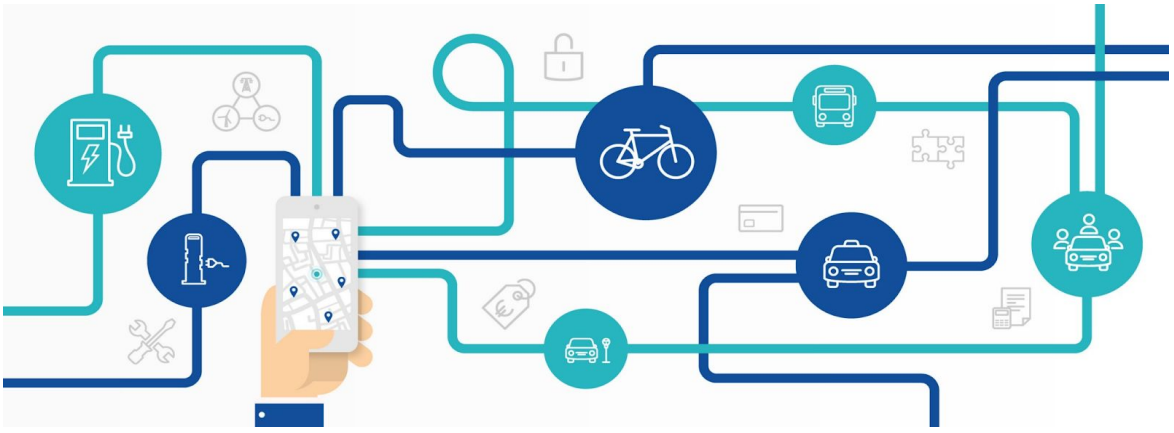
- Understand the different components of a smart mobility system and its complex interactions
- Demonstrate the capacity to design, develop and test integrated models for mobility solution assessment
- Leverage your algorithms and coding skills along with current international research, to develop a new solution
- Be able to work, communicate and report research results orally and in written form in both Danish and English

LANGUAGE: Communication: English. Thesis: English or Danish.

[1] https://turo.com/rentals/cars/tesla

NUMBER OF STUDENTS: 1 or 2 master student (1-2 project of 30-35 ECTS Credits).

14. Mobility Tokens and Blockchain: Development of a Smart Mobility Management System



imagesource: Medium

PROJECT ADVISORS: Carlos L. Azevedo (climaz@dtu.dk)

KEYWORDS: Blockchain; Mobility Management; Software Development

BACKGROUND: Smartphones are taking over the mobility spectrum. Having a wide range of mobility services at your fingertips has transformed the way we move and solutions for coordination of demand and supply as in much need. How to coordinate ride sharing? How to match trip requests with service vehicles? how to make sure mobility demand can be assigned in the most efficient way? How can information be coordinated, validated and accessible in real-time? Blockchain has been proposed as a game changer and presented a viable solutions to similar questions in many other sectors. However, its design and deployment in the mobility realm is still to be explored. As every trip using a mobility service can be viewed as a transaction, the potential of blockchain development for transportation demand-supply coordination is still unknown.

PROJECT DESCRIPTION: This project aims to build a first prototype of a mobility management solution based on the blockchain technology. Pushing for a smartphone-based implementation of blockchain, a system for mobility transactions data based on tokens will be designed as a proof of concept. Tokens can be viewed as electronic credits used in the mobility transaction exchange and could be implemented as incentives (for more mobility efficient transactions, i.e. causing less congestion or emissions) or permits that need to be spent when using a mobility service. A prototype would then be developed by the student. The development shall be based on traditional programming languages used in blockchain and mobile environment compatible.

PREREQUISITES: Knowledge of C++, JavaScript or Python is required; familiarity or interest in self-learning in mobile environment is expected.

LEARNING OBJECTIVES:

- Understand fundamentals of blockchain technology and how to develop them
- Recognise key components of an on-demand mobility service and its control logic
- Demonstrate the capacity to design, develop and prototype a commercial level solution
- Leverage your algorithms and coding skills along with current international research
- Be able to work, communicate and report research results orally and in written form in both Danish and English

LANGUAGE: Communication: English. Thesis: English or Danish.
NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

15. Deep Learning Approaches for Joint Activity and Location Modelling

PROJECT ADVISORS: Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Valentino Servizi (valse@dtu.dk), Professor Francisco Pereira (camara@dtu.dk)

KEYWORDS: Trip Purpose Detection; Next Location Prediction; Deep Learning; Data Fusion; transfer learning; Smartphone-based Travel Surveys.

BACKGROUND: In order to study human travel behaviour, there is a widespread consensus that novel smartphone-based data collection and survey technologies (e.g. FMS [1]) provide a unique potential to unlock our understanding how people move, why they move, what travel modes they choose, what factors influence their choices, etc. This potential is further increased by the possibility of harvesting better results by the fusion of multiple datasets (Transport Network, Points-of-Interest, Social Media, Land Use, etc.). The consequence is that datasets are becoming bigger and more complex, and so are the resulting challenges. For such complex types of data, Machine Learning approaches, and in particular modern Deep Learning and Data Fusion techniques present themselves as excellent candidates for overcoming a number of the current limitations.

PROJECT DESCRIPTION: The goal of this project is to review, develop and test machine learning algorithms for trip purpose detection and next location prediction, by building joint models of activity (what people do) and location (where they choose to do it). The idea is that by modelling the temporal/sequential and spatial dimensions of the data jointly, while eventually also levering information from other data sources (data fusion), we can be able to better predict the next activity of an users and its location. Addressing such problem would be of tremendous value for numerous transport-related applications (e.g. car sharing/car pooling solutions, Mobility-on-Demand services, travel recommendation/assistant apps, etc.), as well as for improving our understanding of human travel behaviour.

PREREQUISITES: Knowledge of Python, Statistics and Machine Learning.

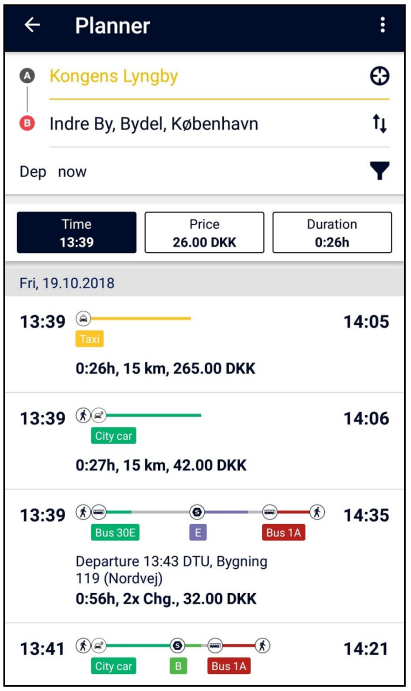
LANGUAGE: English.

NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

[1] https://its.mit.edu/future-mobility-sensing

16. Dynamic Travel Behavior under Smart Mobility

PROJECT ADVISORS: Carlos Lima Azevedo (climaz@dtu.dk)
Thomas K. Rasmussen (tkra@dtu.dk)
Francisco Pereira (camara@dtu.dk)
KEYWORDS: Demand modelling; Dynamic behaviour; Smart mobility.



BACKGROUND: Transportation systems are dynamic by nature. The system conditions are constantly changed by endogenous and external factors, affecting all travelling decision making. Incidents, congestion, unplanned events or even delays in performed activities ultimately affect trip departure time, mode, route and actual trip making choices. Smart mobility brings even more dynamic factors into the individual decision making process, with multiple responsive services and real-time information accessible at your fingertips through smartphone apps. Modelling dynamic choices in transportation has recently become more popular and has been implemented to different time scales. Decisions can be divided into three different categories of horizons of dynamic responses: (1) long- term, (2) mid-term, and (3) short-term. Long-term dynamics include lifestyle changes, usually formulated as a result of resource availability — such as owning a car or subscribing to a smart mobility service. Mid-term horizons are based on day-to-day learning of the effects of activities performed throughout a day and their effect, such that "successful" patterns are reinforced. Short-term dynamics are concerned with travel and activity-making responses to real-time conditions of travel networks and activity availabilities. Understanding how individuals change their behavior from their initial plan is essential to creating the informative tools of the future for policy- and decision-makers which generally still rely on static assumptions.

PROJECT DESCRIPTION: In this project you will focus on the development and understanding of the dynamics in individual decision making related to transportation. Both econometric classical methods, based on discrete choice modelling, **or** emerging machine learning techniques will be combined with stochastic process formulations to model such dynamics. Besides the choice of fundamental methods to use, the students can choose to focus on any of the following individual choice dimensions when introducing a new smart mobility service in an urban area:

1. Modelling long-term dynamics in terms of individual and household level car ownership and mobility subscriptions and its complex relationships;
2. Modelling the day-to-day dynamics of individual preferences towards modes given its dynamic performances;
3. Modelling the activity and travel planning process of individuals and its observed choices within a day, where the dynamic attributes of the transportation system are key variables to consider

This project will be based on the existing Danish National Travel Survey (TU), an interview survey where Danish residents are asked about their travel activities on a specific day (https://www.cta.man.dtu.dk/english/tvu).

PREREQUISITES: No knowledge of any particular software is required, but interest and basic experience in R or Python is recommended. Knowledge in behavioural modelling is also a plus.

LEARNING OBJECTIVES:

- Understand fundamental concepts of dynamic behaviour
- Formulate, estimate and validate complex models using discrete choice, stochastic processes and/or machine learning techniques
- Leverage your modelling skills along with current international research
- Be able to work, communicate and report research results orally and in practice and academic oriented written form in Danish and English.

LANGUAGE: Danish or English.
NUMBER OF STUDENTS: 1 to 3 master student (1 to 3 project of 30-35 ECTS Credits).

17. Smartphone Data Fusion and Machine Learning for Travel Mode Detection

PROJECT ADVISORS: Valentino Servizi (valse@dtu.dk), Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Professor Francisco Pereira (camara@dtu.dk)

KEYWORDS: Transport Mode Detection; Machine Learning; Data Fusion; Transfer Learning; Smartphone-based Travel Surveys.

BACKGROUND: In order to study human travel behaviour, there is a widespread consensus that novel smartphone-based data collection and survey technologies (e.g. FMS [1]) provide a unique potential to unlock our understanding how people move, why they move, what travel modes they choose, what factors influence their choices, etc. This potential is further increased by the possibility of harvesting better results by the fusion of multiple datasets (Transport Network, Points-of-Interest, Social Media, Land Use, etc.). The consequence is that datasets are becoming bigger and more complex, and so are the resulting challenges. Machine Learning approaches such as Deep Learning or Transfer Learning might be some of the right tools for overcoming a number of the current limitations.

In particular, there are several machine learning fields of study oriented towards the automatic generation of travel diaries which focus on inferring travel mode and travel mode chains from smartphone-based sensor data.

PROJECT DESCRIPTION: The goal of this project is to review, develop and test machine learning algorithms for transport mode detection on smartphone-based sensor. The idea is to leverage data from multiple sources (data fusion), as well as information from multiple users (transfer learning), in order to build robust and personalized travel mode detection methods, which are essential for making sense out of the collected smartphone-based data.

The data preparation and the experiments' design will include multiple datasets, some are available already, others might be discovered by the student. The measures resulting from the designed tests will lead the development.

LEARNING OBJECTIVES:

- Understand fundamental concepts of deep learning and its application for classification problems;
- Extend traditional deep learning formulations to multiple datasets both alternatives (transfer learning) or complementaries (data fusion);
- Compare and support design and implementation choices for Big Data based experiments;
- Leverage your algorithm and coding skills along with current international research, to develop a new solution;
- Be able to work, communicate and report research results orally and in practice and academic oriented written form in Danish and English.

PREREQUISITES: Knowledge of Python, Statistics and Machine Learning.

LANGUAGE: English..

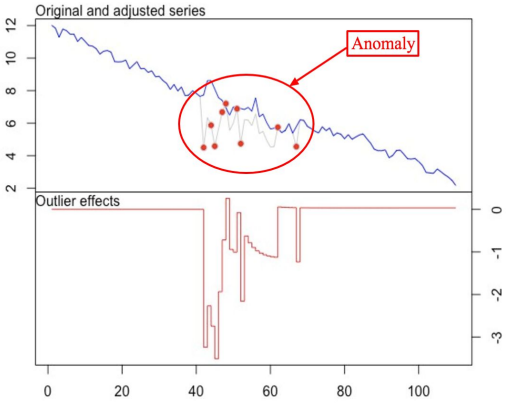
NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

[1] https://its.mit.edu/future-mobility-sensing

18. Anomaly Detection in Mobility Data

PROJECT ADVISORS: Francisco Pereira (camara@dtu.dk), Carlos Lima Azevedo (climaz@dtu.dk)
KEYWORDS: Big data; data processing; sensor data

BACKGROUND: Managing traffic in real-time is a well-known challenge, particularly in highly dynamic cities such as Copenhagen. While the transport system may be carefully planned, it is never possible to cater for all eventualities such as traffic incidents, special events, system disruptions (e.g. public transport breakdown), or harsh weather. A first challenge in such situations is detecting these events from the data.
Techniques such as network sciences, time series analysis or simply outlier detection can help detect such situations if designed carefully. They demand some level of transport network knowledge and statistics, and basic machine learning skills.



Source: https://blog.statsbot.co

PROJECT DESCRIPTION: In this project, we will analyse a dataset with traffic sensing counts from the city of Copenhagen in three major phases:

1. Descriptive statistical analysis – apply common statistical tools (mean, standard deviation, quantiles, histograms, etc.) to understand both the "normal behaviour" and the "extreme behaviour"
2. Network analysis – Estimate transport and network science measures (flow, density, speed, centrality, node degree, etc.) to further characterise normal and extreme behaviour
3. Classification modelling – Using the measures from 1 and 2, train a classification algorithm (e.g. decision tree, logistic regression, neural network, deep learning) that detects, at each moment, if there is a relevant event in the network.

PROJECT PURPOSE: The purpose of this thesis is primarily to gain insight into the traffic sensing data from Copenhagen, and its potential for real time event detection. A successful project will eventually be proposed for the traffic management center.

PREREQUISITES: Statistics, traffic modeling and operations

NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

19. Extracting Contextual Explanations of Abnormal Mobility Patterns from the Web

PROJECT ADVISORS: Professor Francisco Pereira (camara@dtu.dk)**,** Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Assistant Professor Stanislav Borysov (stabo@dtu.dk)

KEYWORDS: Web Mining; Information Retrieval; Information Extraction

BACKGROUND: Managing traffic in real-time is a well-known challenge, particularly in highly dynamic cities such as Copenhagen. While the transport system may be carefully planned, it is never possible to cater for all eventualities such as traffic incidents, special events, system disruptions (e.g. public transport breakdown), or harsh weather. A key factor in treating situations is detecting why a specific disruption is happening. A congestion due to special events will require a very different treatment than a traffic incident or flood.Often, the explanation for such disruptions can be found on the web, in dedicated platforms (e.g. Twitter, Facebook), news websites or querying search engines. The internet has been shown as a very valuable source to understand mobility patterns and problems. The challenge is then how to find the right information, this is known as an Information Retrieval problem.

PROJECT DESCRIPTION: The objective of this project is to develop a methodology for automatically finding contextual explanations online about previously detected traffic events (information retrieval), which were observable in traffic sensing data (e.g. vehicle counters) or public transport records (e.g. Rejsekort data). The project will build on earlier work already implemented that automatically queries for event explanations for taxi demand peaks in Singapore.

The project has three general phases:

1. Selection of relevant event case studies from the available datasets
2. Manual information retrieval, where the student needs to understand, for each case, the best possible explanatory content (thus creating a “ground truth” database)
3. Automatic information retrieval, where the student will adapt the previous work to our context

PROJECT PURPOSE: The purpose of this thesis is primarily to gain insight into the possibilities of using internet data for the Copenhagen context, and its potential for real time event explanation. A successful project will eventually be proposed for the traffic management center.

PREREQUISITES: Programming languages.

NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

20. Benchmark Tests for Data-Driven Transportation Modelling

PROJECT SUPERVISOR: Assistant professor Stanislav Borysov (stabo@dtu.dk)
PROJECT ADVISOR: Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Professor Francisco Pereira (camara@dtu.dk)

KEYWORDS: Data Mining; Data Science; Machine Learning

BACKGROUND: Recently, data-driven approaches have gained a lot of attention in the transportation community thanks to the progress in machine learning field combined with exponential growth of data and computational resources available. However, comparing to the other data-intensive fields, such as computer vision or natural language processing, transportation still lacks unified and well-understood benchmark datasets. Creation of such datasets along with design of various test cases based on them are crucial not only to compare performance of the proposed models but also to study their limitations and transferability.

PROJECT DESCRIPTION: The main goal of the project is to provide the transportation research community with well-documented and well-understood benchmark tests based on existing open datasets, for example, Bike Sharing Dataset [1], New York City Taxi Trip Data [2] or Taxi Service Trajectory ECML PKDD [3]. Extension of the existing datasets might be also considered, for example, using weather, special events or incidents data collected from online sources. A particular focus lies on designing a number of various test cases for both classification and regression tasks using different data representations, processing and aggregation techniques. The project also involves application of various machine learning algorithms to understand their strengths, weaknesses and conceptual limitations for the test cases designed. The project will hopefully provide guidance towards more controllable, reliable and transferable data-driven transportation modelling instead of ad hoc black-box approaches widely used nowadays.

The project consists of the three main steps:
1. Collection and analysis of data available (statistics, visualization, handling missing data, etc).
2. Design of different test cases for both regression and classification tasks.
3. Study of the performance of different models on the proposed tests.

The developed tests, their detailed description and basic modelling results will be made publicly available.

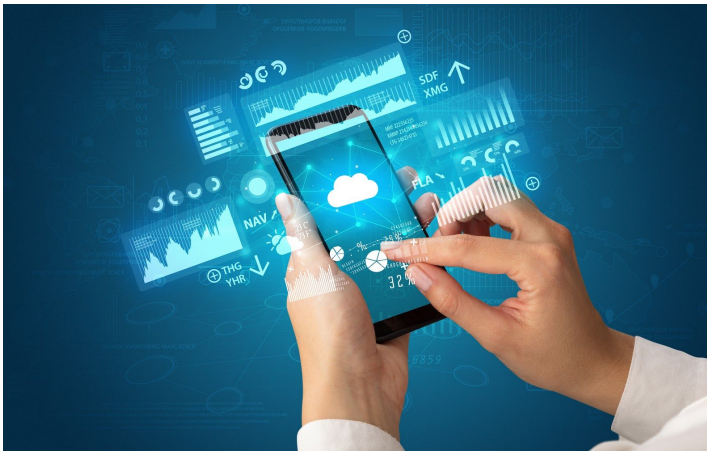
[1] https://archive.ics.uci.edu/ml/datasets/bike+sharing+dataset
[2] http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml
[3] http://www.geolink.pt/ecmlpkdd2015-challenge/dataset.html

PREREQUISITES: Good programming skills in Python, basic knowledge or interest in machine learning and data mining.

LANGUAGE: English.

NUMBER OF STUDENTS: 1-2 master students (2 projects of 30-35 ECTS Credits).

21. Post-Travel Analytics on Smartphone-based Activity Diaries



PROJECT SUPERVISOR: Professor Francisco Pereira (camara@dtu.dk)

PROJECT ADVISORS: Valentino Servizi (valse@dtu.dk), Hjalmar Christiansen (hjalc@dtu.dk)

KEYWORDS: Smartphone-based Travel Surveys; Data Analytics; Data Mining.

BACKGROUND: In order to achieve a smarter mobility paradigm and improve our experience as users while reducing our footprint on the environment, there is a large consensus about the smartphone-based travel surveys potential of revealing new awareness on the human interaction with the transport system, such as the motivation behind (i) the individual mobility choices, (ii) the commute patterns in different contexts, (iii) the departure time decisions, and so on. However, even after the introduction of technologies claiming to have solved both the battery consumption problems and the concerns about privacy, despite the encouraging results of the ongoing research concerning the provision of the right incentives to motivate people to share their travel behaviour data, smartphone-based travel surveys are still far from being a widespread technology and there is still a considerable amount of research to be done.

PROJECT PURPOSE: The purpose of this project is to evaluate the data quality of smartphone-based travel diaries by developing a range of such post-travel analytics measures, estimate them with available technology and evaluate their reliability.

PROJECT DESCRIPTION: The project will comprise the following three components.

- A small multi-day data collection campaign (potentially based on DTU Transport students and staff, some of which recording a personal log for later ground truth);
- Development of post-travel analytics measures, and estimate them from the available data;
- Evaluate their reliability by comparing with ground truth.

PREREQUISITES: Basic skills in statistics are as necessary as the interest in survey design and in smartphone-based technologies.

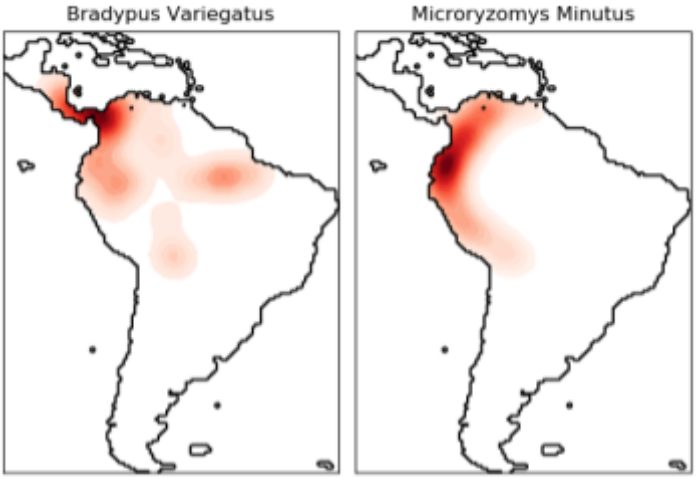
LANGUAGE: English.

NUMBER OF STUDENTS: 1 master student (1 project of 30-35 ECTS Credits).

22. Spatio-temporal conditional density estimation

PROJECT ADVISORS: Assistant Professor Filipe Rodrigues (rodr@dtu.dk), Sergio Garrido (shgm@dtu.dk)

KEYWORDS: Deep learning, Bayesian inference, Density estimation, Spatio-temporal data, Data visualization.



Source: "Maximum entropy modeling of species geographic distributions" S. J. Phillips, R. P. Anderson, R. E. Schapire - Ecological Modelling, 190:231-259, 2006.

BACKGROUND: Density estimation aims to model a distribution of an outcome variable "y" given a set of covariates "x": P(y|x). We are particularly interested in tasks where the covariates "x" include both the spatial and temporal domain. Therefore, the estimated distribution can be used to predict densities for several tasks related with cities (which is the focus of this research). For example, we could predict the density of crime or the "demand" of shared mobility vehicles in a certain area (with specific characteristics) at a certain time in a specific month of the year. There are several difficulties associated with estimating this probability distribution: data can have high sparsity, data may contain selection bias or be censored, models could overfit to the data that it is trained on, etc.

PROJECT PURPOSE: The aim of this project is to develop a robust method to do spatio-temporal conditional density estimation with an application to a mobility problem. We would explore the generative modeling literature and the deep generative modeling literature. The outcome of this project will have direct policy implications and might be sent for conferences or scientific journals. As an aside, we are also interested in the visualization of spatio-temporal data.

RELEVANT LITERATURE:

- Conditional Density Estimation with Neural Networks: Best Practices and Benchmarks (2019) (https://arxiv.org/pdf/1903.00954.pdf)
- Masked Autoregressive Flow for Density Estimation (2018) (https://arxiv.org/pdf/1705.07057.pdf)

PREREQUISITES: Good programming skills in Python, knowledge of probability, statistics and (ideally) machine learning.

LANGUAGE: English.

NUMBER OF STUDENTS: 1-2 master students (2 projects of 30-35 ECTS Credits).