

# DATA SCIENCE WITH SOCIETAL APPLICATIONS

Michael Mathioudakis Digital Humanities Summit Helsinki, 11-December-2020



#### Welcome to the Algorithmic Data Science research group!

The goal of our group is to develop efficient and easily manageable data science systems and use them for novel applications. Towards this end, we conduct research at all stages of data science, from data management and processing to model inference and applied data

# RESEARCH GROUP Algorithmic Data Science helsinki.fi/algorithmic-data-science since 2018 Department of Computer Science University of Helsinki

#### People

Read more about our group members.

### Publications

Articles in scientific journals and conference proceedings.

#### Research

We conduct research under two broad themes: (i) data science applications with societal interest, (ii) end-to-end data science systems. To...



Michael Mathioudakis Assistant Professor Group Leader Since 2018



Arpit Merchant Doctoral Student Since 2019



Sachith Pai Doctoral Student Since 2020



Yanhao Wang Postdoctoral Researcher Since 2020



Ananth Mahadevan Doctoral Student Since 2020

Collaborations with Aalto U. (Finland), Aarhus U. (Denmark), ISI (Italy), KTH (Sweden), NUS (Singapore), UPF (Spain) Laura Huuskonen Doctoral Student Social Sciences Since 2020 Francesco Fabbri Visiting Doctoral Student UPF

Eeva-Maria Laiho Madhav Narendran <sup>Master's students</sup>

Previous members Corinna Hertweck (now Zurich HAW) Riku Laine (now Social Sciences @ UH)

## Efficient algorithms

# Applied domain

Probability

Computing

## Applications of societal interest

Our connection with HELDIG affirmative action policies, evaluation of automated decisions, social interactions on the web, social network analysis

## Design of fair decision policies Affirmative action policies for university admissions

Mathioudakis, M., Castillo, C., Barnabo, G., & Celis, S. (2020, March). Affirmative action policies for top-k candidates selection: with an application to the design of policies for university admissions. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing* (pp. 440-449).

# Evaluation of automated decisions Over biased data

Laine R., Hyttinen A., Mathioudakis M. (2020) Evaluating Decision Makers over Selectively Labelled Data: A Causal Modelling Approach. In: Appice A., Tsoumakas G., Manolopoulos Y., Matwin S. (eds) Discovery Science. DS 2020. Lecture Notes in Computer Science, vol 12323. Springer, Cham. Best paper award.



## <u>Setting</u>

Students obtain a set of scores during high school (e.g., high-school grades, SAT exams, etc) and apply to university programs

Assigned total admission score (e.g., average of subject scores) Highest admission scores are admitted

> Socioeconomic attributes (gender, income)

## **Task**

Design bonus policy to favor under-respresented groups (e.g., give 5 bonus points to females who apply for engineering school)

### Aim for:

(1) equal in-group admission rates
(e.g., % admitted female/male applicants)
(2) excellence

optimize measure of policy quality: high admission scores + low admission disparity



## Applications: Engineering School, Santiago, 2017





## **CHALLENGE**

Computer decides BAIL where Human had decided JAIL: What would have happened if the defendant had been set free?

## **ALGORITHM**

Guess missing outcomes in historical data.

SOLUTION Counterfactual analysis





Similar situation in other decision scenarios: E.g., Loan or No-Ioan, Medicine or No medicine Automated decisions must be evaluated <u>BEFORE</u> actual use Data do not contain the outcome for decisions that were not made Algorithmic Data Science Research Group

Research

Driven by societal applications Work on computational challenges that arise in them

Vision

Enable efficient, automated, easily administered Data Science for a variety of applied domains.



#### Welcome to the Algorithmic Data Science research group!

The goal of our group is to develop efficient and easily manageable data science systems and use them for novel applications. Towards this end, we conduct research at all stages of data science, from data management and processing to model inference and applied data

RESEARCH GROUP Algorithmic Data Science helsinki.fi/algorithmic-data-science since 2018 Department of Computer Science University of Helsinki

#### People

Read more about our group members.

### Publications

Articles in scientific journals and conference proceedings.

#### Research

We conduct research under two broad themes: (i) data science applications with societal interest, (ii) end-to-end data science systems. To...