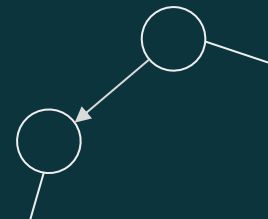


TECHNISCHE
UNIVERSITÄT
DARMSTADT

Hands-on: Causality for Machine Learning

Matej Zečević

7th Int'l Summer School on Data Science (SSDS 2022),
Day 4, 13th October 2022



Overview of Today's Programme:

- ~~☐ Done: *Causality for Machine Learning 1*, where we explore the basic concepts from causality that are being used in current ML research~~

~~9 am — 10:30 am (including Q&A)~~

- ~~☐ Done: *Causality for Machine Learning 2*, where we discuss some selected publications and on-going work on the interface of causality and ML~~

~~10:45 am — 12:15 am (including Q&A)~~

- ☐ **Now:** *Hands-on session: Causality for Machine Learning*, where you get to experience some code for using causal ML first hand

1:30 pm — 3 pm

master

[SSDS-2022](#) / [Day-4](#) / [Causality-for-Machine-Learning](#) / [Tutorial](#)

Go to file

Add file

...



zecevic-matej Added Causality for ML Material

d710578 10 hours ago History

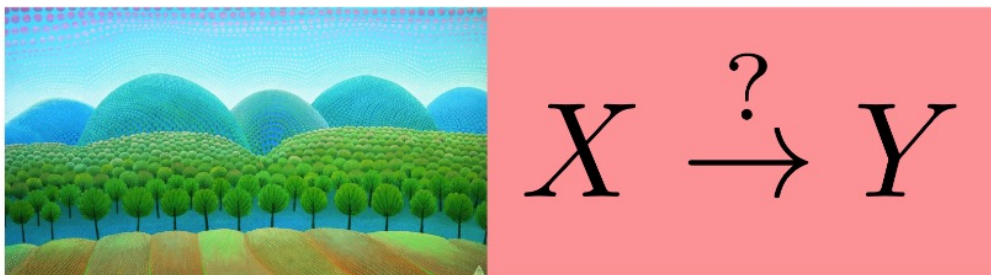
..

Banner-Tutorial.png	Added Causality for ML Material	10 hours ago
Banner-Tutorial.svg	Added Causality for ML Material	10 hours ago
Dataset-Example-Phenotype-and-Genes-Nonlinear.csv	Added Causality for ML Material	10 hours ago
README.md	Added Causality for ML Material	10 hours ago
SSDS_2022_Hands-on_Causality_for_ML.ipynb	Added Causality for ML Material	10 hours ago
linear.py	Added Causality for ML Material	10 hours ago
utils.py	Added Causality for ML Material	10 hours ago

README.md



SSDS 2022











Explainable AI and Causal Learning for Data Science

Hands-on session: Causality for Machine Learning

Open in Google Colab

by searching for repository in
Google Colab



	zecevic-matej Added Causality for ML Material
..	
	Banner-Tutorial.png
	Banner-Tutorial.svg
	Dataset-Example-Phenotype-and-Genes-Nonlinear.csv
	README.md
	SSDS_2022_Hands-on_Causality_for_ML.ipynb
	linear.py
	utils.py

CO

Welcome to Colaboratory

File Edit View Insert Runtime Tools Help

Table of contents

Getting started

Data science

Machine learning

More resources

Featured examples

Section

Code

Text

Copy to Drive

1

2

3

Examples

Recent

Google Drive

GitHub

Upload

Enter a GitHub URL or search by organisation or user

☐ Include private repos

https://github.com/SSDS-Croatia/SSDS-2022/tree/master/Day-4/Causality-for-Machine-Learning/Tutorial

Repository: SSDS-Croatia/SSDS-2022 Branch: master

Path

Day-4/Causality-for-Machine-Learning/Tutorial/SSDS_2022_Hands-on_Ca...

New notebook

Cancel



SSDS 2022



Explainable AI and Causal Learning for Data Science

Hands-on: Causality for Machine Learning

This tutorial combines elements from two previously existing tutorials. One of them first authored by [Alexandre Drouin](#) with contributions from [Philippe Brouillard](#) and [Thibaud Godon](#)

Abstract: This tutorial will consist of a practical introduction to the estimation of causal effects. We will experiment with the concepts of average treatment effect, randomization, covariate adjustment, and inverse probability weighting to derive common estimators from the literature. We will also see where machine learning models fit into such estimators. Formal derivations will be presented and supported by extensive visualizations.

Outline

- Exercise 1: Simpson's paradox (15 min)
- Exercise 2: Identification and estimation via parent adjustment (20 min)
- Exercise 3: From parent to back-door adjustment (20 min)
- Exercise 4: Estimation via machine learning (15 min)
- Exercise 5: Causal Structure Learning (10 min)

5 Exercises

That's what it
looks like

SSDS_2022_Hands-on_Causality_for_ML.ipynb

File Edit View Insert Runtime Tools Help Cannot save changes

Files

sample_data
Dataset-Example-Phenotype-and-G...
linear.py
utils.py

Utility functions

Run these functions to bootstrap the notebook.

```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns

from google.colab import data_table
data_table.enable_dataframe_formatter()

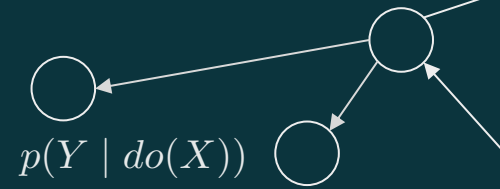
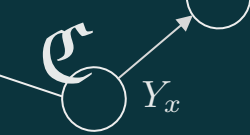
%pip install igraph
%pip install causal-learn
import random
import utils
from string import ascii_uppercase
from sklearn.linear_model import LinearRegression
from causallearn.search.ConstraintBased.BIC import BIC
from causallearn.search.ConstraintBased.BIC import BIC
from linear import notears_linear
```

Looking in indexes: <https://pypi.org/simple>
Collecting igraph
Downloading igraph-0.10.1-cp37-cp37m-macosx10.9.pkg | 3
Collecting texttable>=1.6.2
Downloading texttable-1.6.4-py2.py3-none-any.whl (10 kB)
Installing collected packages: texttable, igraph
Successfully installed igraph-0.10.1 texttable-1.6.4

Don't forget to upload the
GitHub repository files (3 files)
to the Runtime of the Google Colab

Hands-on Algorithm

1. You setup the Collab
2. We go over to the next exercise task, if none, then go to Step 5
3. Then you start exploring yourself, “learning by doing”
4. Finally, we look at the solution together and go back to Step 2
5. **We are done and have learned a lot, yeah!**



Let's Code!

See you in Google Colab.

matej.zecevic@tu-darmstadt.de | <https://www.matej-zecevic.de>

Further, my gratitude and thanks go out to Kristian Kersting, Devendra Dhami, all our collaborators, the AIML lab @ TU Darmstadt and finally the organizers of SSDS 2022

