MUNI
ICS

Intro to Geo/enviro AI + practical examples

Terézia Slanináková, Institute of Computer Science, Masaryk University

Outline

- About me / us
- Intro / Motivation
- Examples of AI in enviro + from us
- How to get into AI (as an enviro person)
- Conclusion

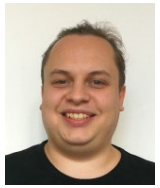
About me / us

- CERIT-SC, Institute of Computer Science, Masaryk University

Data Science @CERIT-SC representatives here*:

Terézia Slanináková

- Ph.D. student (Faculty of Informatics, Masaryk University)
  - topic: similarity search in complex data using AI
- focus at CERIT-SC: AI applied to enviro



Pavel Kraus

- Mgr. student, topic: Language models applied to criminal investigations

Tom Rebok

- Senior researcher, research group leader

*not actually here ☹

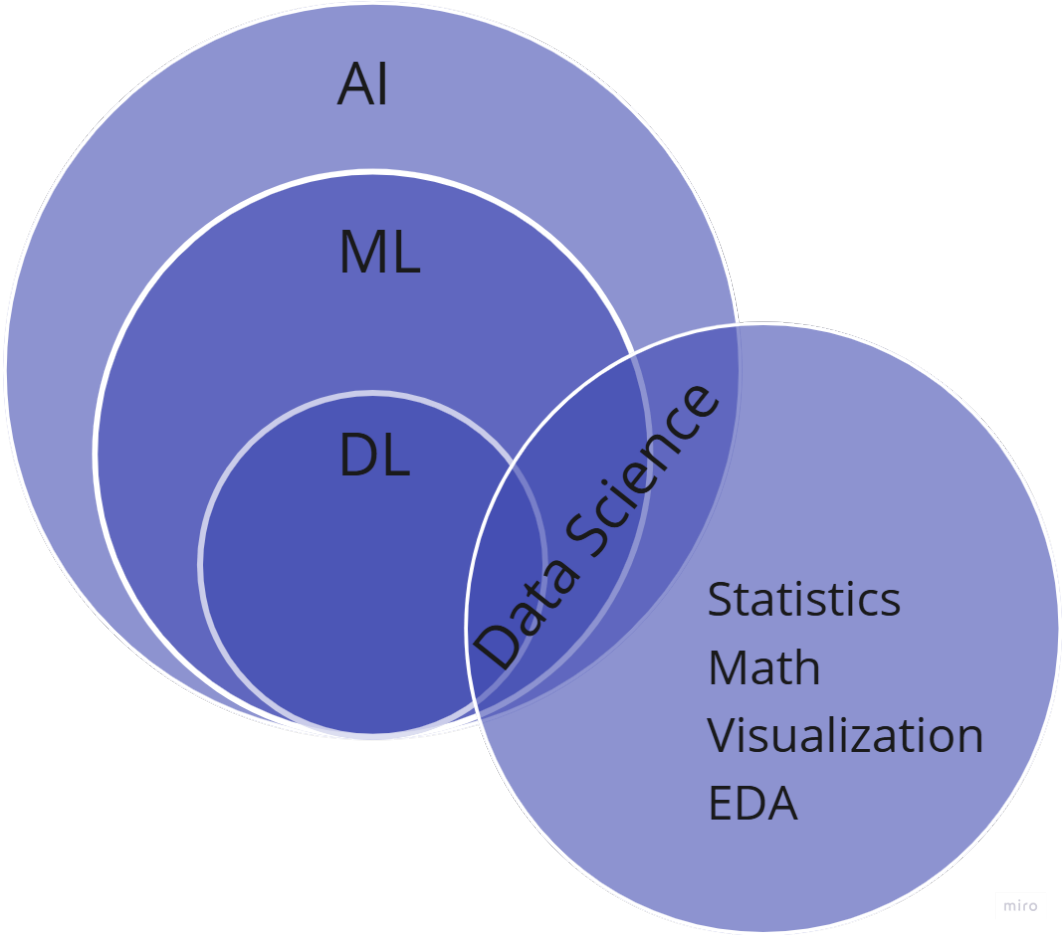
Enviro-related activities:

- Applying AI to enviro research w/ CzechGlobe (AV ČR)
- EnviLab (https://envilab.cerit-sc.cz/) w/ CDV (ČZU)
- eLTER: Environmental Research infrastructure
  - Tom's presentation

MUNI ICS

Introduction

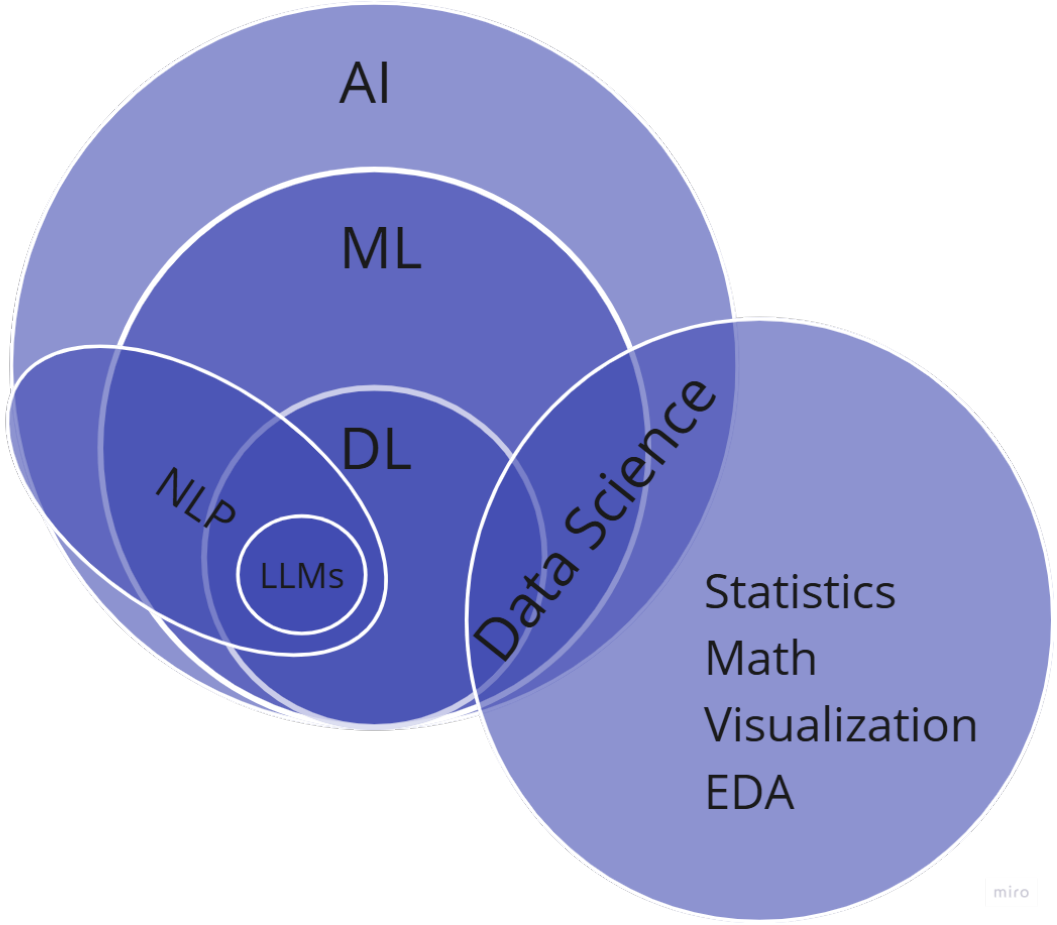
What even is AI?



MUNI ICS

Introduction

What even is AI?



MUNI ICS

Motivation

Why use AI in enviro

- Enviro systems are complex, therefore the data is complex
  - Interrelated factors: climate, biodiversity, and human activity
  - Diverse data sources: Satellites, sensors, historical records
    - High spatial, temporal, and spectral resolutions, new data every couple of days
  - We want to test a hypothesis, model a relationship, predict something
    - Previously, we might have used statistics

Motivation

Why use AI in enviro

- Enviro systems are complex, therefore the data is complex
  - Interrelated factors: climate, biodiversity, and human activity
  - Diverse data sources: Satellites, sensors, historical records
    - High spatial, temporal, and spectral resolutions, new data every couple of days
  - We want to test a hypothesis, model a relationship, predict something
    - Previously, we might have used statistics
- Statistics vs ML
  - Have complex relationships between variables?
  - Need to integrate various data sources, to process big data?
  - Not necessarily any of the two, but want a better predictive performance?

Use of AI in enviro

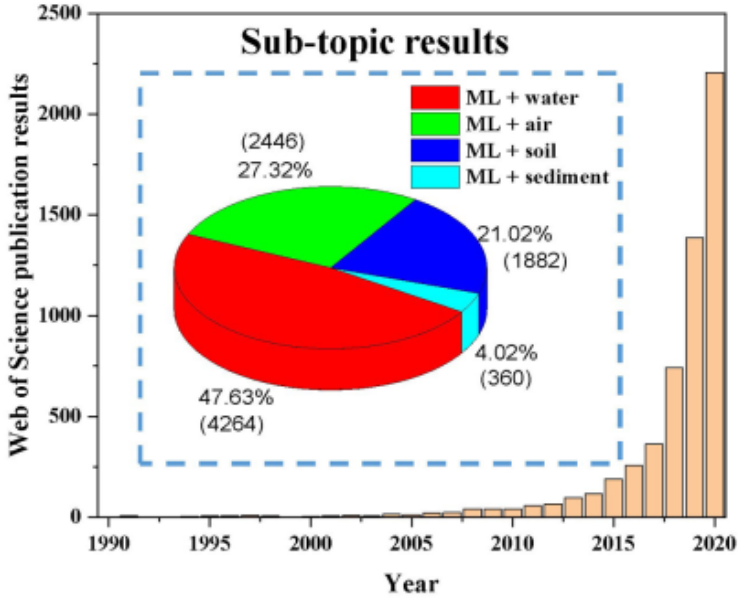


Figure 1. Number of publications in ML applications in ESE based on the Web of Science (access date 1/28/2021) with the keyword "machine learning" combined with the categories of environmental science, water resources, public environmental occupational health, environmental engineering, and environmental studies. The inset shows the subtopic results from 1990 to 2020 with the keywords specified.

Zhong, Shifa, et al. "Machine learning: new ideas and tools in environmental science and engineering." *Environmental science & technology* 55.19 (2021): 12741-12754.

MUNI ICS

Use of AI in enviro

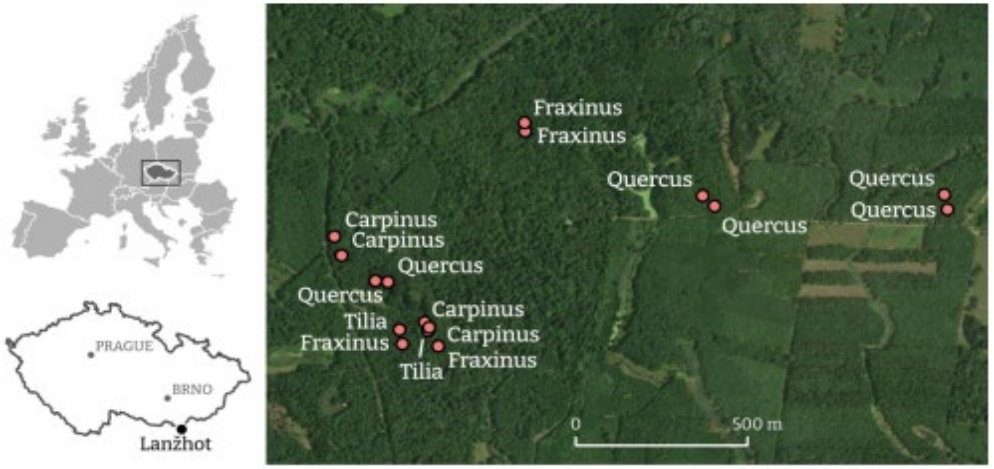
Examples

- Soil and Land Management
  Predict land use changes and soil health from satellite imagery and soil sensors
- Water Quality and Management
  Predict indicators such as pH, dissolved oxygen, and concentrations of pollutants from sensors.
- Air Quality Monitoring
  Predict pollution levels using air quality monitoring stations.
- Biodiversity and Ecosystem Monitoring
  Monitor wildlife populations and assess biodiversity from camera trap footage and acoustic recordings
- Climate and Weather Prediction
  Improve the accuracy of precipitation forecasts using deep learning

Examples from us

Vegetation parameters | Coop. w/ R. Janoutová, M. Švik, CzechGlobe

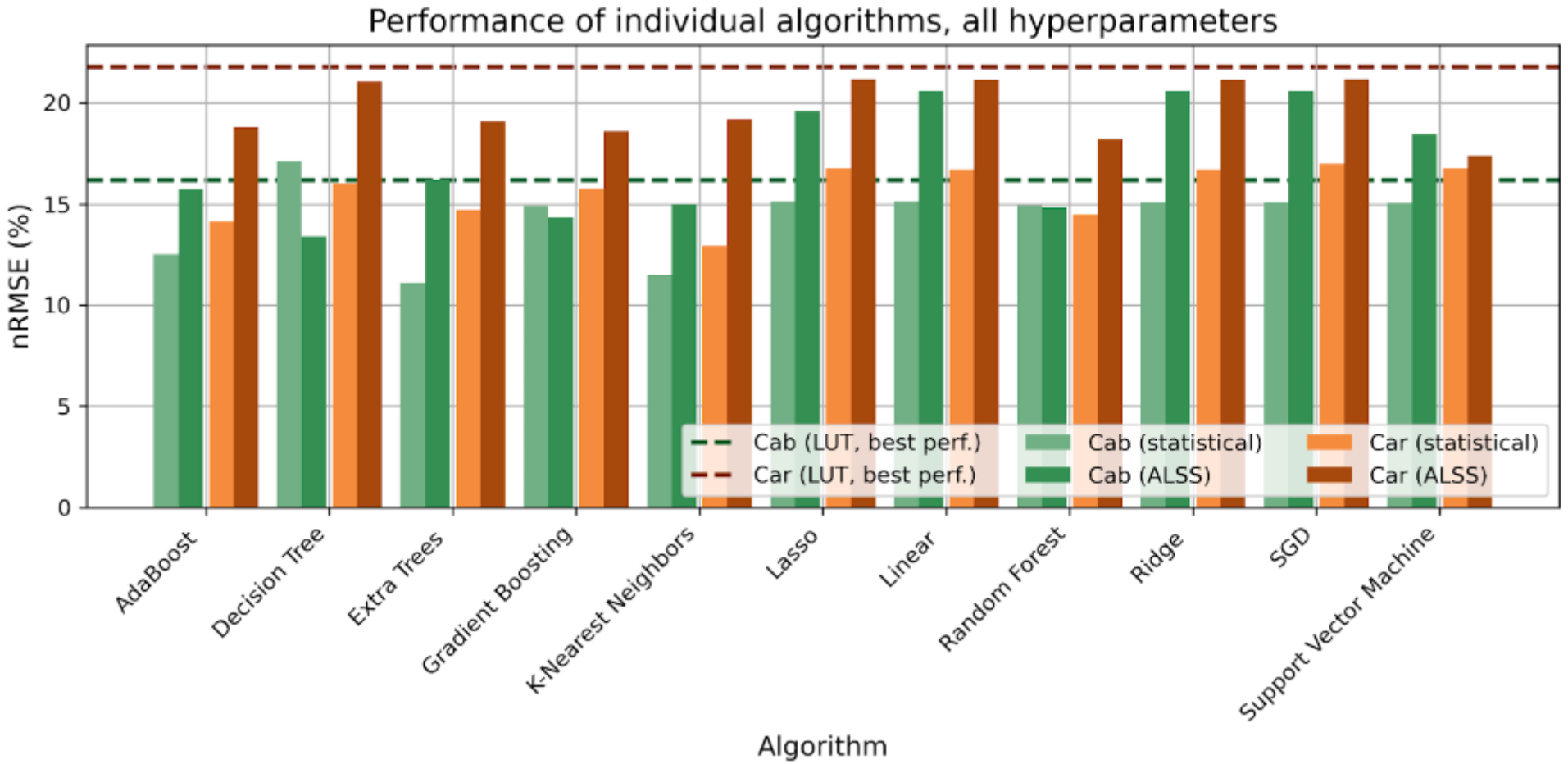
- Goal: Predict the amount of chrolophyll and carotenoids in vegetation using real and simulated data
- Data:
  - Satellite (S2, HLS) image, Lanžhot
  - “Labels”:
    - in-situ acquisition, 69 trees
    - Similar simulated spectra using DART



MUNI ICS

Examples from us

Vegetation parameters

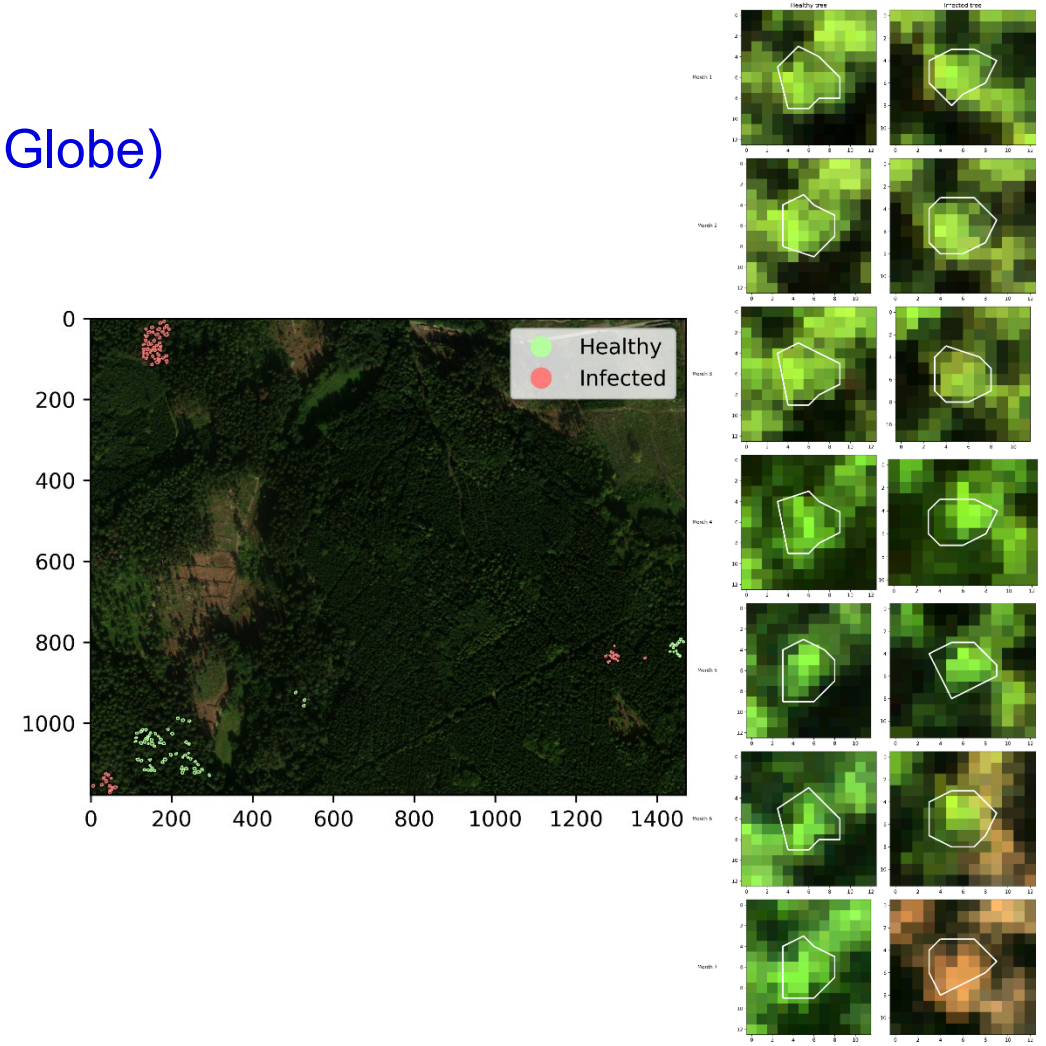


MUNI ICS

Examples from us

Bark beetle prediction | Coop w/ Vojta Bárta (CzechGlobe)

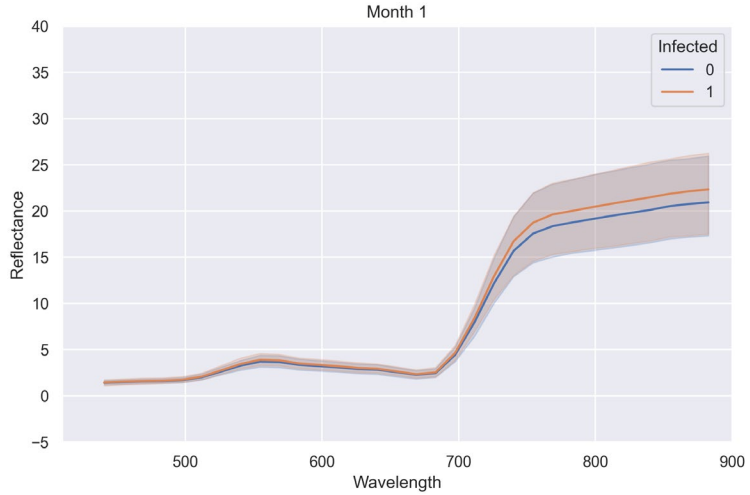
- Goal: Predict the state (healthy or not) of a tree in a region infested with the bark beetle
- Data:
  - Airborne (CASI) image, VNIR spectrum (400-1000 nm)
  - „labels“: 7 acquisitions nearby Brno (Jedovnice), 150 trees
    - manual annotation
- Methods
  - Explorative data analysis (EDA)
  - Supervised binary classification

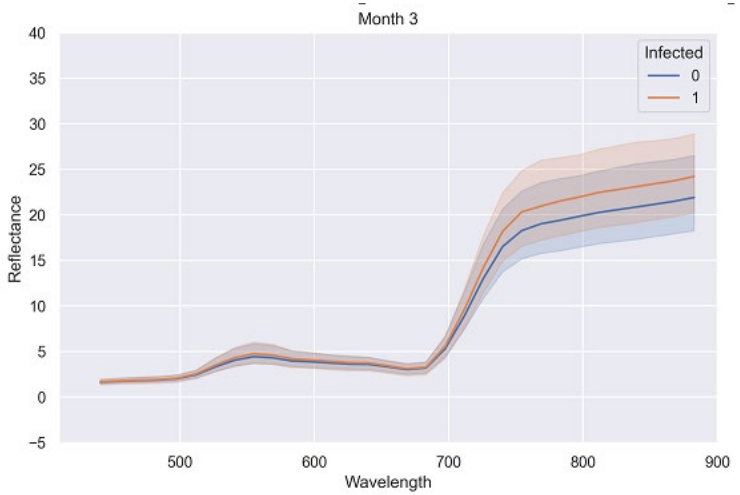


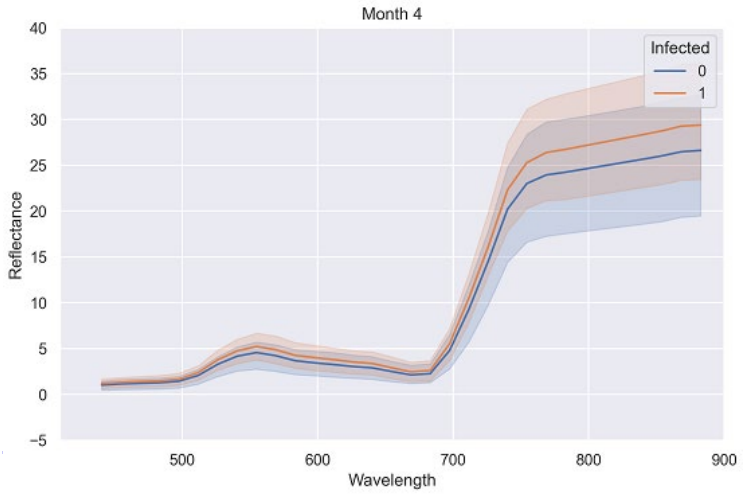
MUNI ICS

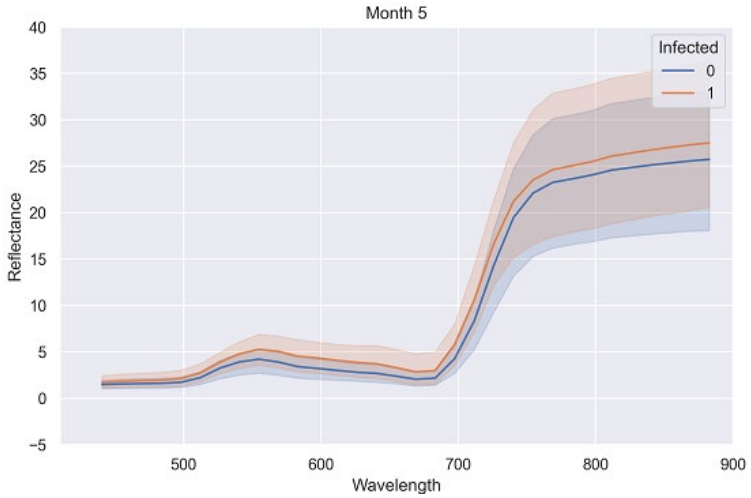
Examples from us

Bark beetle prediction [Work in progress] | Explorative data analysis





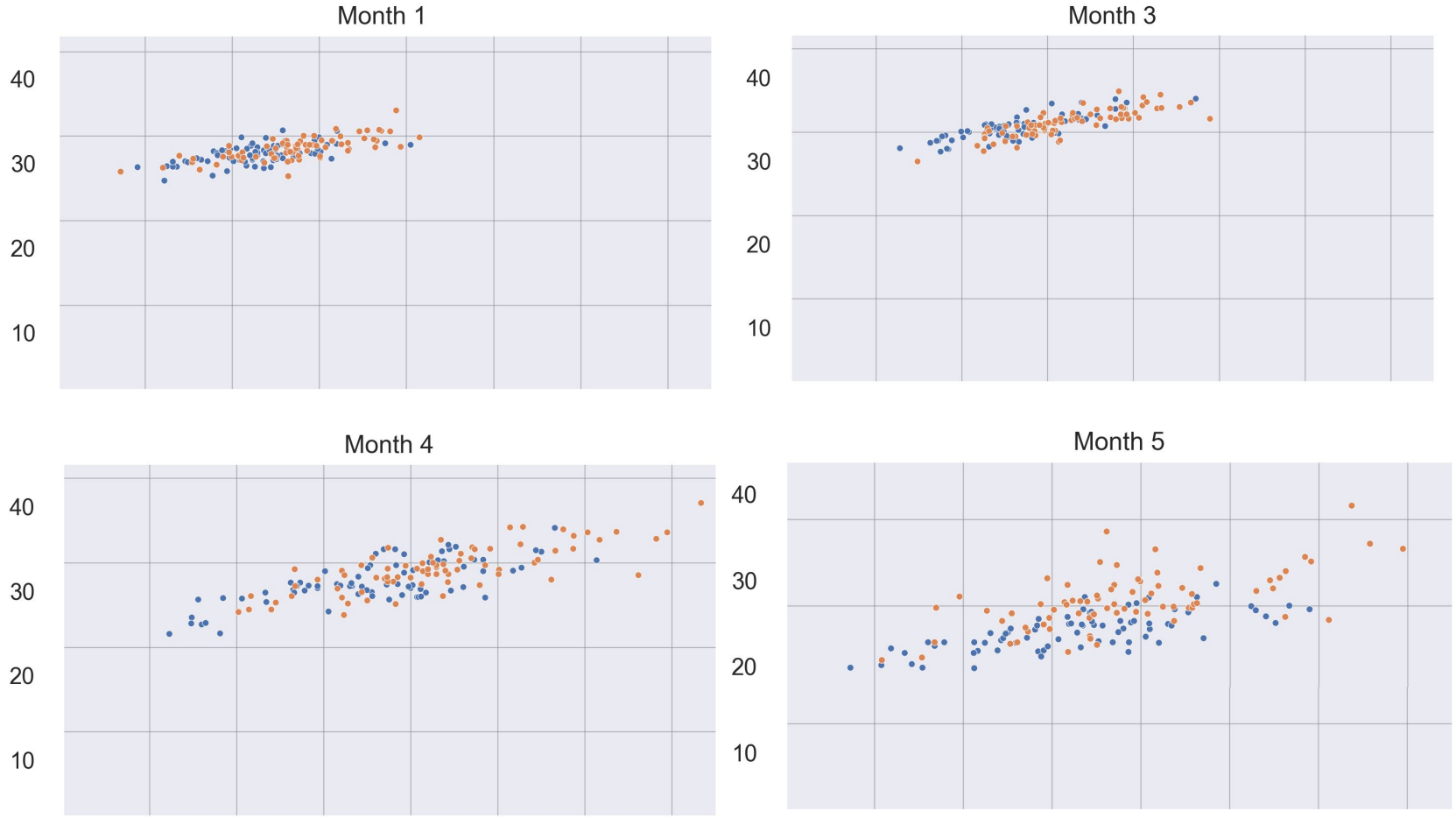




MUNI ICS

Examples from us

Bark beetle prediction [Work in progress] | Explorative data analysis



MUNI ICS

Examples from us

Bark beetle prediction [Work in progress]

- Preliminary results:

MUNI ICS

How to use AI in enviro research?

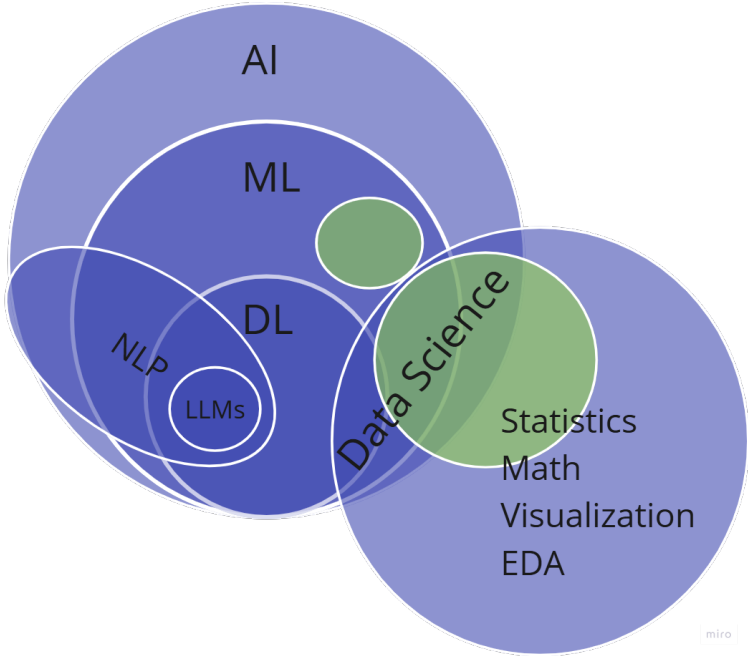
2 Options:

- 1) Find a partner to outsource ML to
  - Not as “0 work” as it sounds
  - Need a lot of contact with you to understand the problem, to gather feedback on results
- 2) Learn ML
  - Not necessarily “all of ML”

How to use AI in enviro research?

2 Options:

- 1) Find a partner to outsource ML to
  - Not as “0 work” as it sounds
  - Need a lot of contact with you to understand the problem, to gather feedback on results
- **2) Learn ML**
  - Not necessarily “all of ML”



MUNI ICS

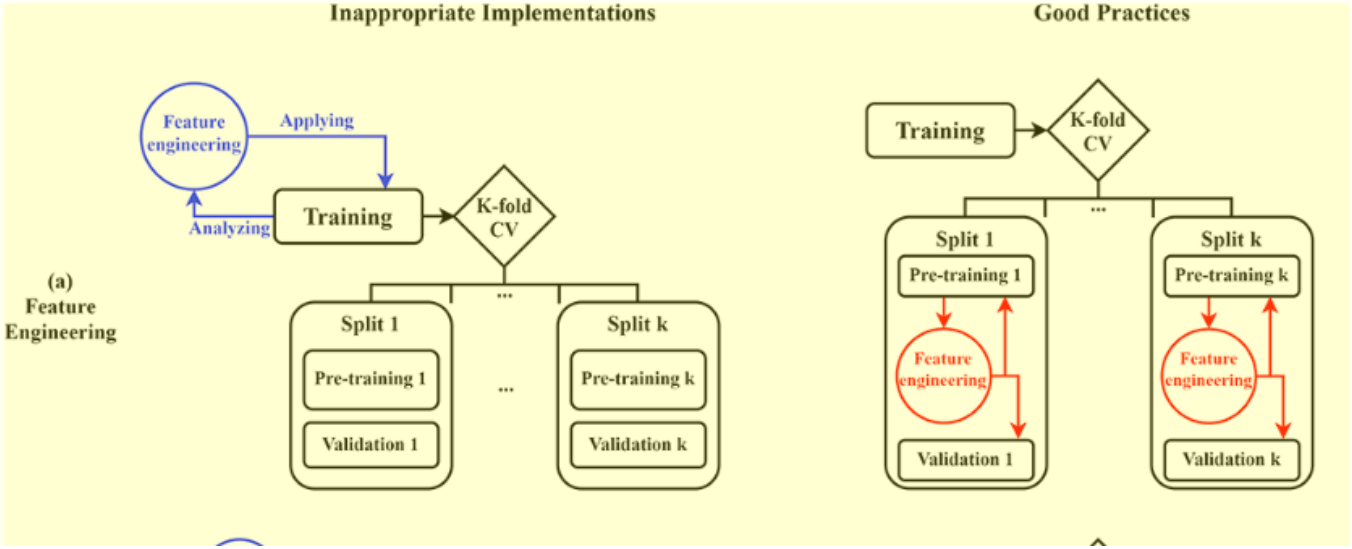
How to use AI in enviro research?

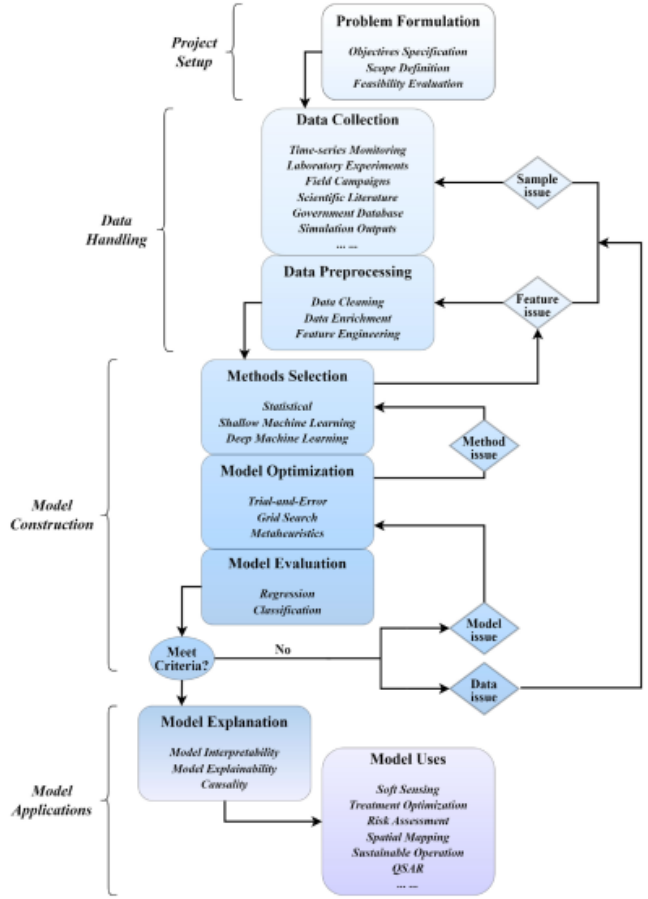
Learn ML - sources

Machine Learning in Environmental Research: Common Pitfalls and Best Practices

Jun-Jie Zhu,* Meiqi Yang, and Zhiyong Jason Ren*

Cite This: *Environ. Sci. Technol.* 2023, 57, 17671−17689 Read Online





MUNI ICS

How to use AI in enviro research?

Learn ML - sources

Machine Learning in Environmental Research: Common Pitfalls and Best Practices

Jun-Jie Zhu,* Meiqi Yang, and Zhiyong Jason Ren*

Cite This: *Environ. Sci. Technol.* 2023, 57, 17671−17689 Read Online

- Python, scikit-learn, some kind of AI pair-programmer
- Kaggle.com
  - E.g. https://www.kaggle.com/datasets/akshatgupta7/crop-yield-in-indian-states-dataset

Agricultural Crop Yield in Indian States Dataset

Crop yields of Indian States and UTs from year 1997-2020

Data Card | Code (11) | Discussion (0) | Suggestions (0)

About Dataset

This dataset encompasses agricultural data for multiple crops cultivated across various states in India from the year **1997 till 2020**. The dataset provides crucial features related to crop yield prediction, including crop types, crop years, cropping seasons, states, areas under cultivation, production quantities, annual rainfall, fertilizer usage, pesticide usage, and calculated yields.

Usability ⓘ
10.00

License
CC BY-SA 4.0

Expected update frequency

MUNI ICS

How to use AI in enviro research?

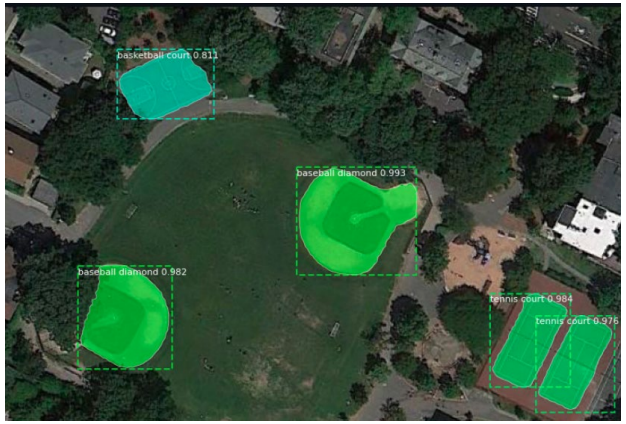
Learn ML - sources

Machine Learning in Environmental Research: Common Pitfalls and Best Practices

Jun-Jie Zhu,* Meiqi Yang, and Zhiyong Jason Ren*

Cite This: *Environ. Sci. Technol.* 2023, 57, 17671−17689 Read Online

- Python, scikit-learn, some kind of AI pair-programmer
- Kaggle.com
- Geospatial machine learning from Microsoft:
  - https://www.microsoft.com/en-us/research/project/geospatial-machine-learning/, various datasets: https://torchgeo.readthedocs.io/en/stable/api/datasets.html



MUNI ICS

Conclusion

- AI is a useful tool for modelling relationships and generating predictions
  - The results are only as good as the data is: Data curation, sharing, open repositories, adherence to FAIR principles matters
- To apply AI properly, you need to master the basics
  - Many tutorials, resources
- Getting into programming has never been easier