

Machine Learning for Behavioral Data (CS-421)

Introduction

February 20, 2023

Today

- What is ML for Behavioral Data?
- Course Logistics
- Active Learning
- Projects: EdTech StartUp(s)

Today

- What is ML for Behavioral Data?
- Course Logistics
- Active Learning
- Projects: EdTech StartUp(s)



This will be an interactive course...

- More on this later
- For now: take your phone (or laptop) and join us on SpeakUp

<https://go.epfl.ch/speakup-mlbd>



About Me

- Assistant professor at EPFL since May, 2020
 - Head of the ML4ED lab
 - In the past, I was a
 - senior data scientist at the SDSC
 - postdoc at Stanford University
 - postdoc at ETH Zurich/consultant for Disney research Zurich
 - PhD student at ETH Zurich
-

Students – Shake Hands



What is ML for Behavioral Data?

SpeakUp Chat!

14/02/2022 21:25, by me

0
0 votes

0 comments

The image shows a single message card from a platform called "SpeakUp Chat". The message text is "What is ML for Behavioral Data?". To the left of the text is a thumbs-down icon. In the center is a large "0" with the text "0 votes" underneath. To the right is a thumbs-up icon. At the bottom right of the card is a link labeled "0 comments". The entire card has a blue border.

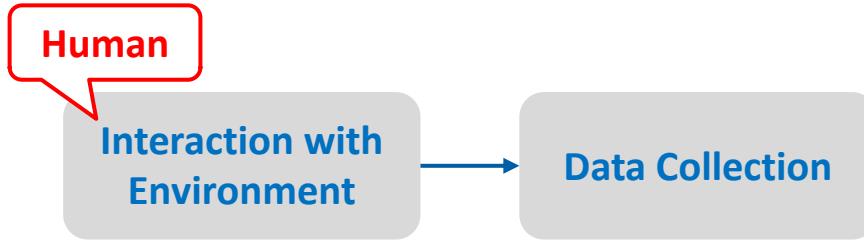
What is ML for Behavioral Data?

Human

Interaction with
Environment

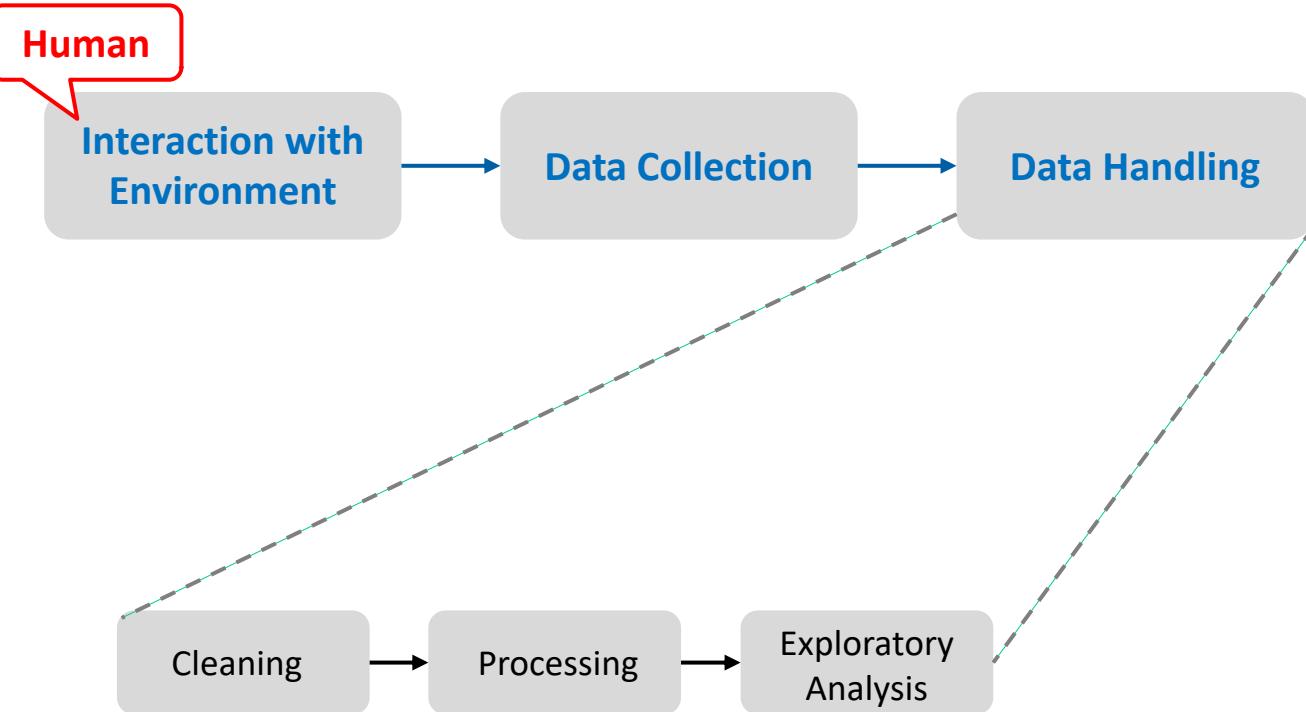


What is ML for Behavioral Data?

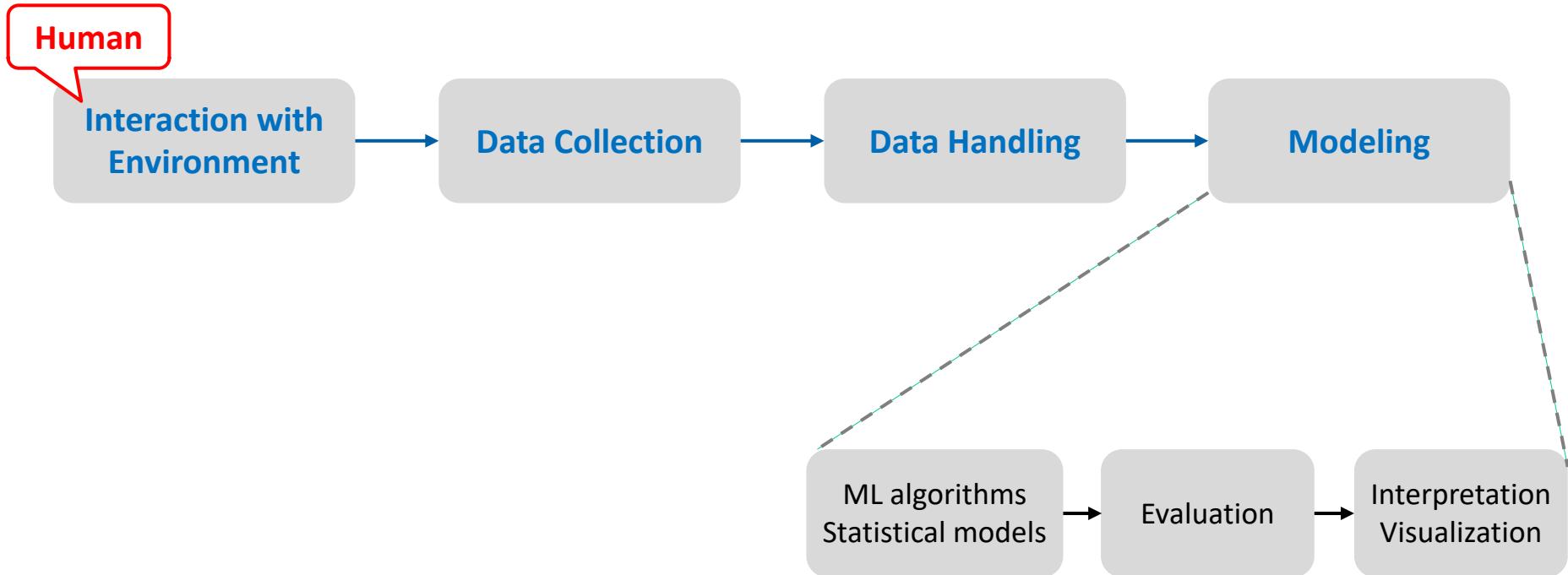


- Clickstream
- Text
- Categorical Data
- Images
- Video
- Sensor Data
- ...

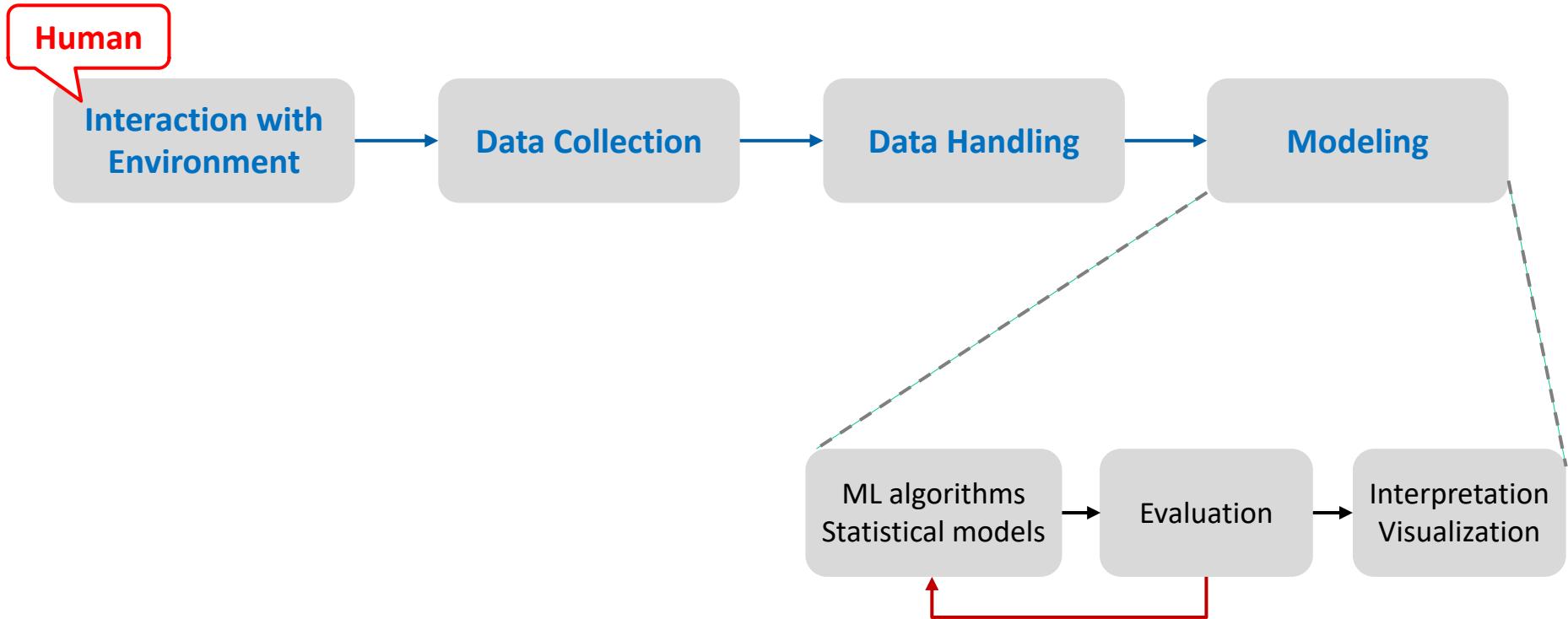
What is ML for Behavioral Data?



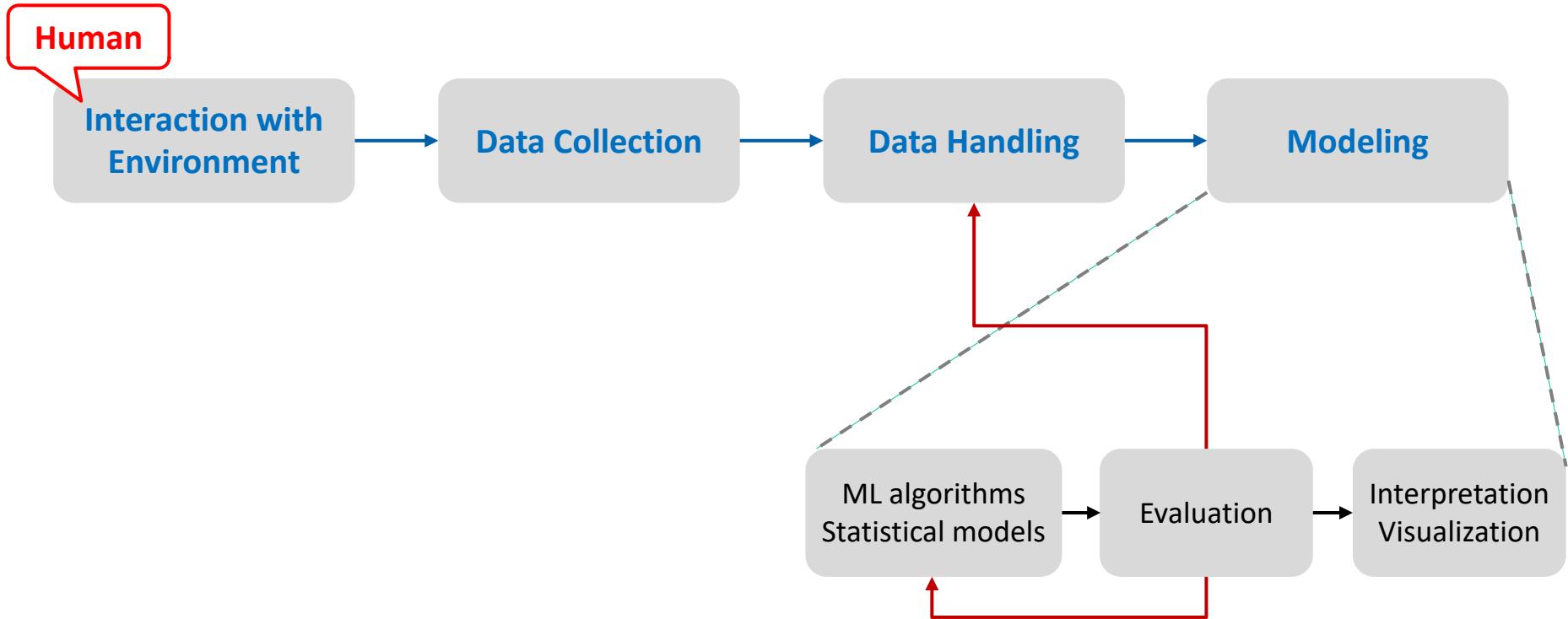
What is ML for Behavioral Data?



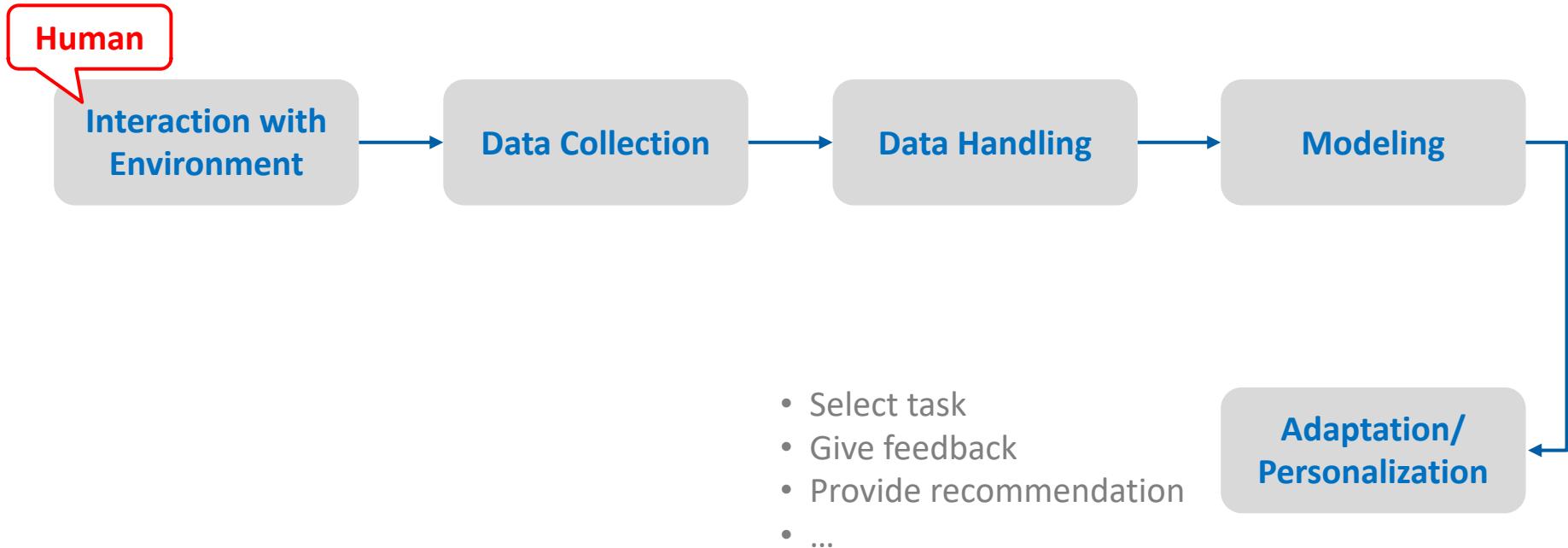
What is ML for Behavioral Data?



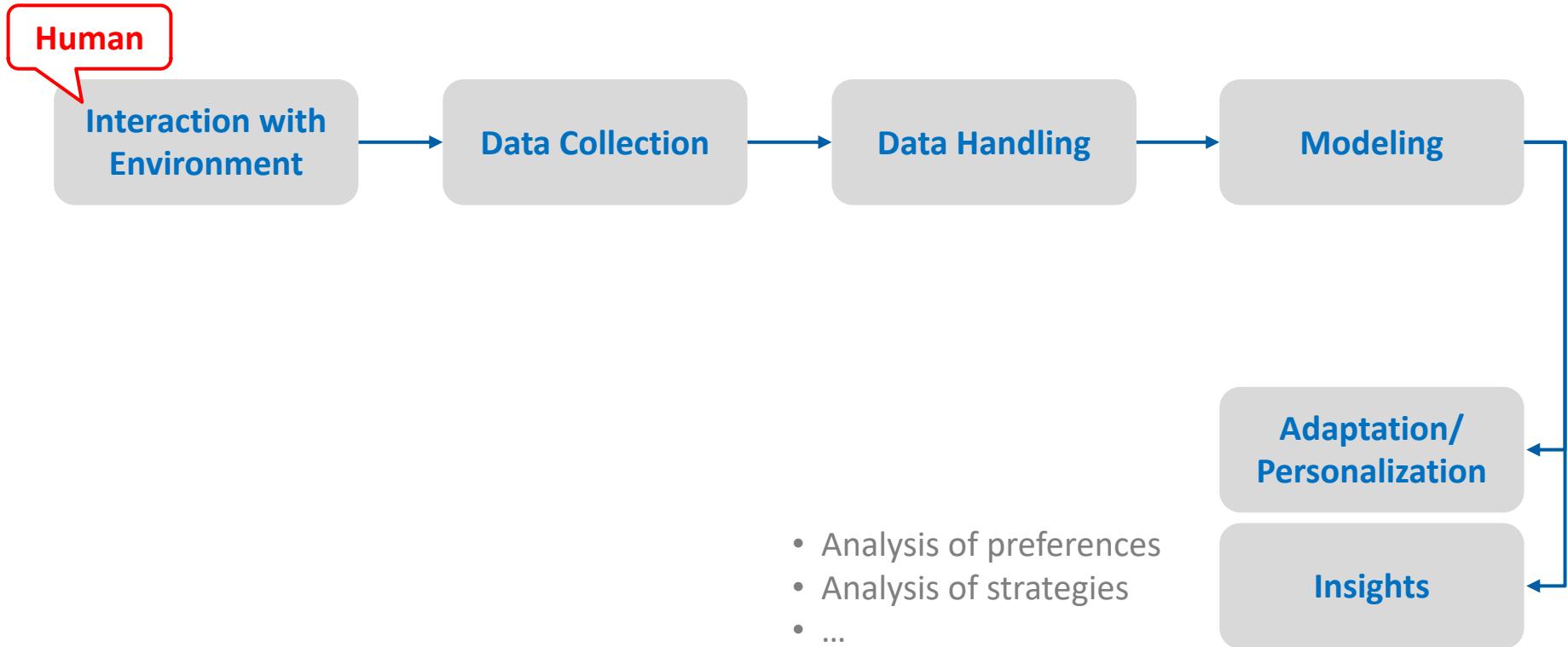
What is ML for Behavioral Data?



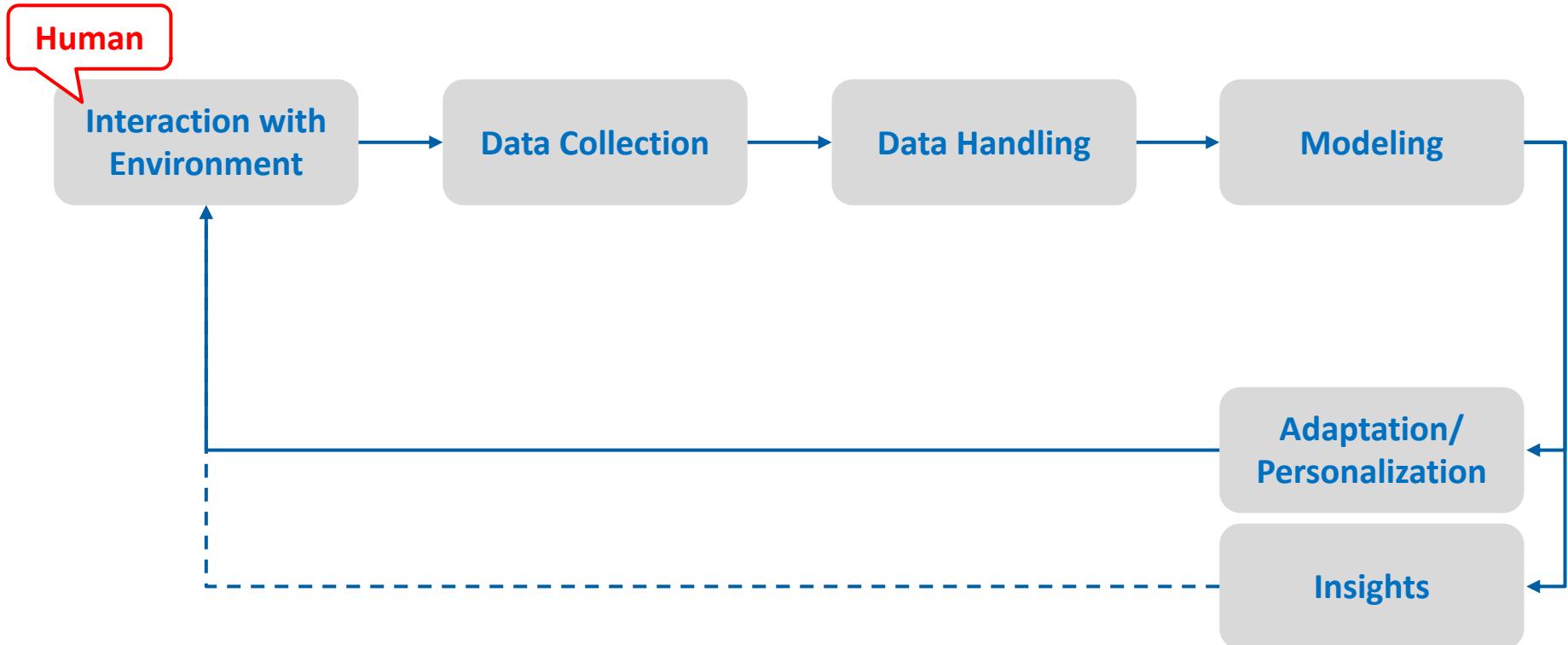
What is ML for Behavioral Data?



What is ML for Behavioral Data?

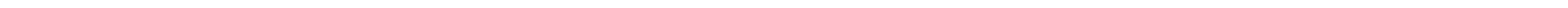


What is ML for Behavioral Data?



Lecture Syllabus

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction



Lecture Syllabus

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction

- Exploring & visualizing data
- Time Series Exploration

Lecture Syllabus

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction

- Generalized Linear Models
- Mixture Models
- Regression for time series

Lecture Syllabus

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction

- Random Forest, nearest neighbors, etc.
- Classifying time series data

Lecture Syllabus

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction

- Cross validation, bootstrap, information scores
- Error metrics & visualization

Lecture Syllabus

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction

Complete pipeline for one use case:

- Data exploration
- Prediction
- Model evaluation

Lecture Syllabus

Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction

Supervised learning on time series:

- Probabilistic graphical models
- Neural networks: LSTM, GRU, etc.

Lecture Syllabus

Week	Lecture/Lab
8	Spring Break
9	Time Series Prediction
10	Unsupervised Learning
11	Unsupervised Learning
12	Ethical Machine Learning
13	Ethical Machine Learning
14	Project Presentations
15	Whit Monday

Lecture Syllabus

Week	Lecture/Lab
8	Spring Break
9	Time Series Prediction
10	Unsupervised Learning
11	Unsupervised Learning
12	Ethical Machine Learning
13	Ethical Machine Learning
14	Project Presentations
15	Whit Monday

- 
- K-Means, Spectral Clustering
 - Choosing the optimal K*
 - Clustering time-series data

Lecture Syllabus

Week	Lecture/Lab
8	Spring Break
9	Time Series Prediction
10	Unsupervised Learning
11	Unsupervised Learning
12	Ethical Machine Learning
13	Ethical Machine Learning
14	Project Presentations
15	Whit Monday

- 
- Fairness
 - Explainability

Lecture/Lab

- Monday, 15:15 – 18:00
 - INR 113
 - Lecture + practice session
 - Slides will be uploaded to our GitHub
 - Jupyter Notebooks will be uploaded to our GitHub
 - Recording: we will make the recordings from the past year available
-

Project

- Teams of 3 people
 - We will provide the data sets (from EdTech Start-Ups)
 - We will provide example research questions
 - You will suggest an additional analysis/extension to the selected research question
 - We will give feedback during the semester (see milestones)
 - You will do a **poster presentation** in the penultimate week of the semester
 - Final project (Code + Report) delivered by **June 9, 2023 23:59 CET**
-

Project (Office) Hours

- Wednesday, 9:15-10:00
- INM 10
- Content:
 - Introduction to project tasks
 - Individual feedback meetings with teams
 - Drop-in office hours for questions regarding the lecture or project

Project Schedule

Week	Project Hours	Milestones
1	Environment setup	-
2	Introduction to tasks for M2	<i>M1: preferences on team members and start-up</i>
3	Office hours	
4	Introduction to tasks for M4	<i>M2: individual exploration of selected data set</i>
5	Office hours	<i>M3: selection of research question and approach</i>
6	Individual discussion with teams	
7	Office hours	
8	Spring Break	

Project Schedule

Week	Project Hours	Milestones
9	Team Coaching	
10	Office hours	<i>M4: submission of results for first research question M5: ideas for extension (+ approach)</i>
11	Individual discussion with teams	
12	Office hours	
13	Team Coaching	
14	Poster Presentations	<i>M6: poster session (in person, on campus)</i>
15	Office Hours	
16		<i>M7: Hand in report and code base</i>

Grading

- **50% Project**
 - Teams of 3 people
 - All milestones are mandatory
 - All individual feedback meetings and team coaching meetings are mandatory
 - 15% individual exploration (M2), 25% supervised learning (M4), 20% presentation (M6), 40% final results (M7)
 - **50% Final Exam (exam session)**
 - Individually, at the laptop
 - Mix of conceptual and coding questions
-

Course Goals

- Explain the main machine learning approaches to personalization, describe their advantages and disadvantages and explain the differences between them
 - Implement algorithms for these machine learning models
 - Apply them to real-world data
 - Assess / evaluate their performance
-

Which ML courses have you taken?

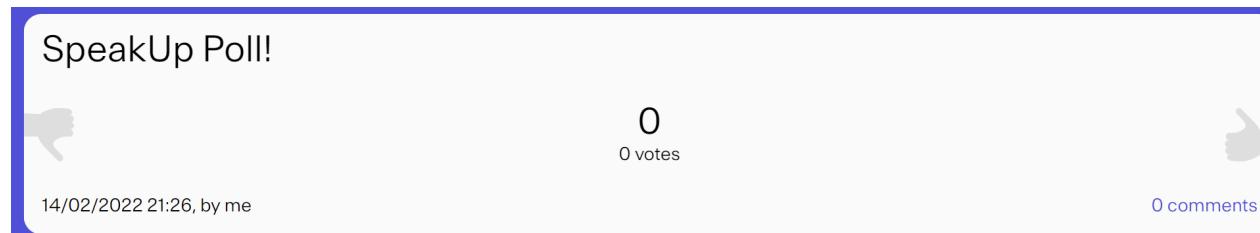
- A. Introduction to Machine Learning
- B. Machine Learning
- C. Applied Data Analysis
- D. Other

SpeakUp Poll!

0
0 votes

14/02/2022 21:26, by me

0 comments



Course Prerequisites

- Probabilities and statistics
- Programming:
 - Project: Python
 - Exam: Python
- Foundations of machine learning



Important Websites

- Moodle: <https://moodle.epfl.ch/course/view.php?id=16434>
 - Contains all important information
 - Use forum for questions
 - For more personal questions contact teaching assistants
- Project:
 - GitHub: <https://github.com/epfl-ml4ed/mlbd-2023>
 - EPFL Noto: <https://noto.epfl.ch/>

Team

Instructor

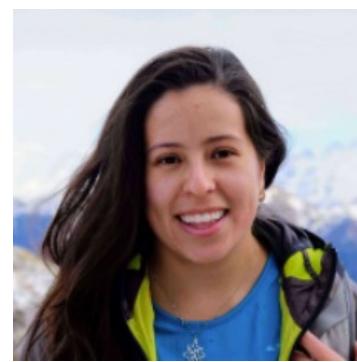


Tanja Käser
tanja.kaeser@epfl.ch

Teaching Assistants



Vinitra Swamy, Paola Mejia
vinitra.swamy@epfl.ch, paola.mejia@epfl.ch



Feedback

- We are committed to providing the best possible version of the course
- If you want to give us feedback, there will be a link on Moodle:

Feedback

We are fully committed to providing the best possible version of the course and we appreciate all constructive feedback.
We are looking forward to reading your comments and improving based on them.

[Feedback link \(anonymous\)](#)

Questions?



Today

- What is ML for Behavioral Data?
- Course Logistics
- **Active Learning**
- Projects: EdTech StartUp(s)

Active learning – what is it?

SpeakUp Chat!



0

0 votes



14/02/2022 21:25, by me

0 comments

Active learning – what is it?

- Activities that students do to construct knowledge and understanding
 - Read
 - Write
 - Explore
 - Discuss
 - ...

Active learning in this course

SpeakUp

Collecting Ideas

Polls

Think – Pair - Share

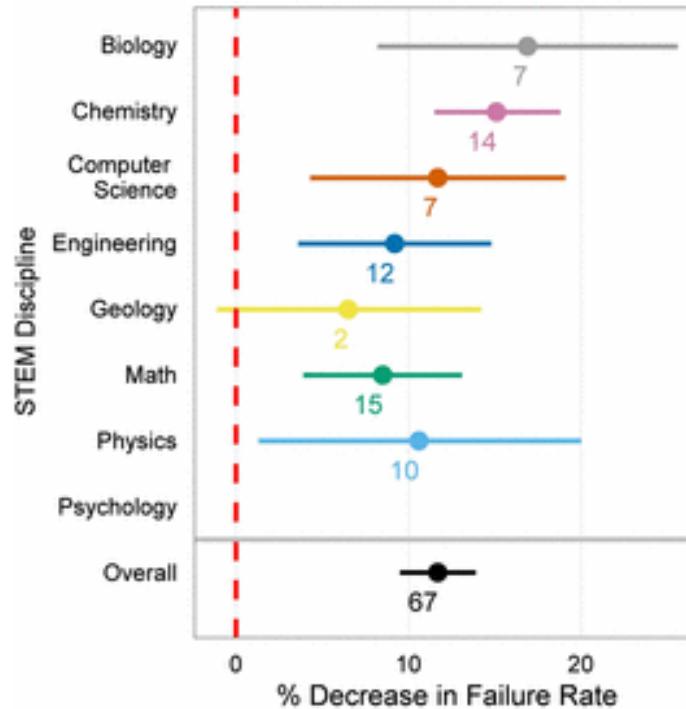
Jupyter Notebook

Demonstration

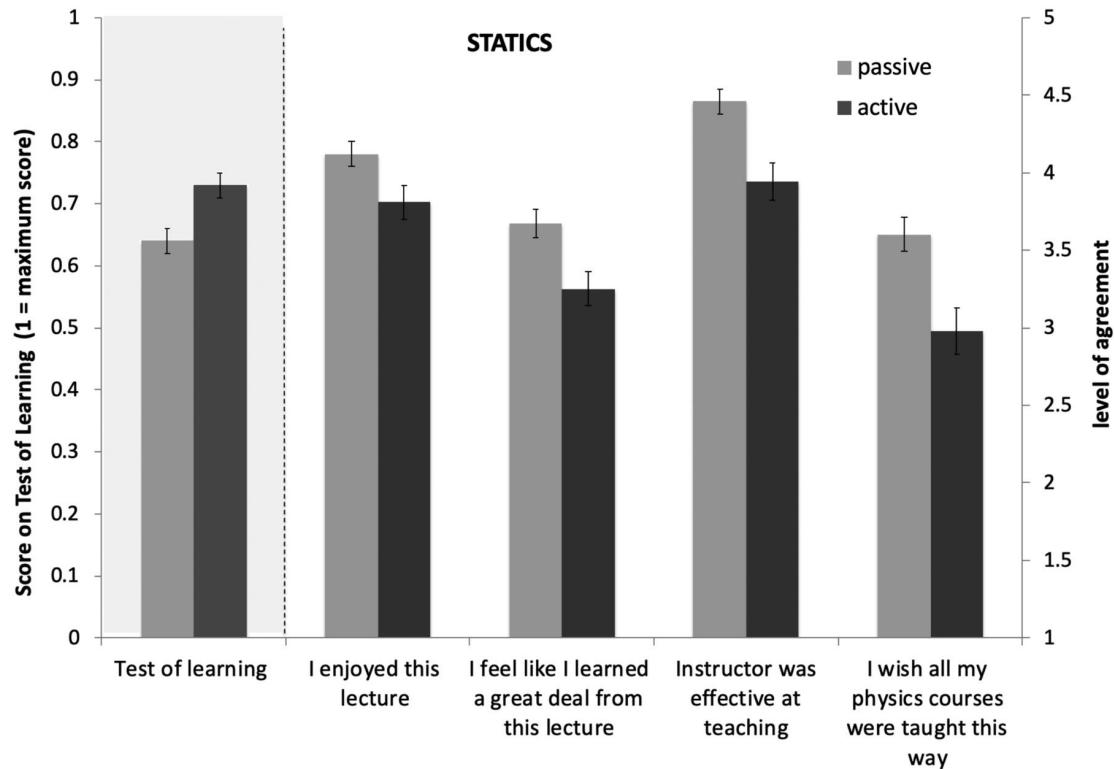
Worked examples

Small coding tasks

Active learning increases performance



Watch out: Feeling-of-Learning can deceive you!



The lecture will be interactive, thus

- we expect you to attend the lecture
- we expect you to participate in all the activities

Important: bring your laptop !

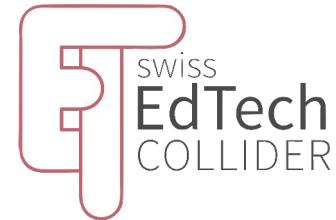
Questions?



Today

- What is ML for Behavioral Data?
- Course Logistics
- Active Learning
- **Projects: EdTech StartUp(s)**

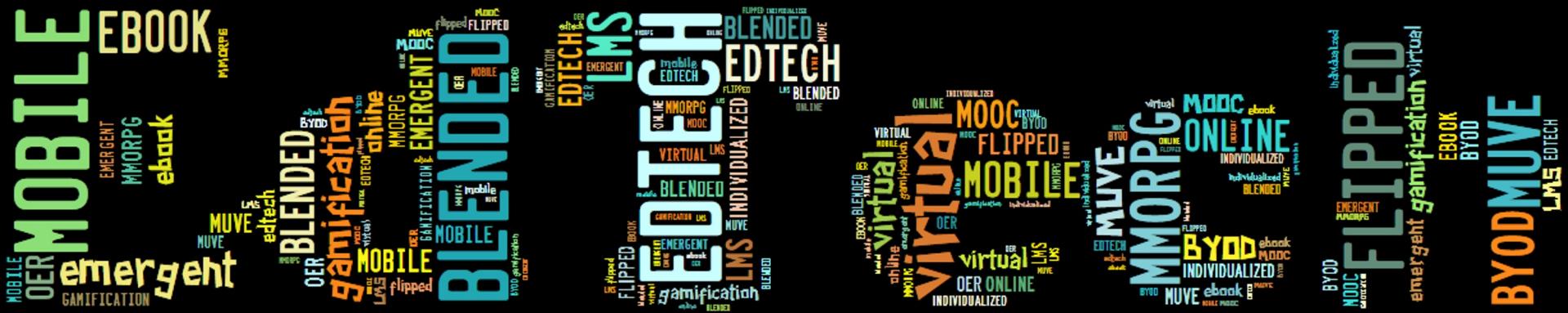
Swiss “EdTech” Hub



Why a Swiss “EdTech” Hub?

To support the digital transformation in **ED**ucation
with **TECH**nological solutions

EdTech Market – Highly Fragmented



Large diversity in the use of technology-enhanced solutions



Early Childhood Education



Compulsory Education



**Upper Secondary /
Higher Education**



University/VET



**Continuous Training &
Education**

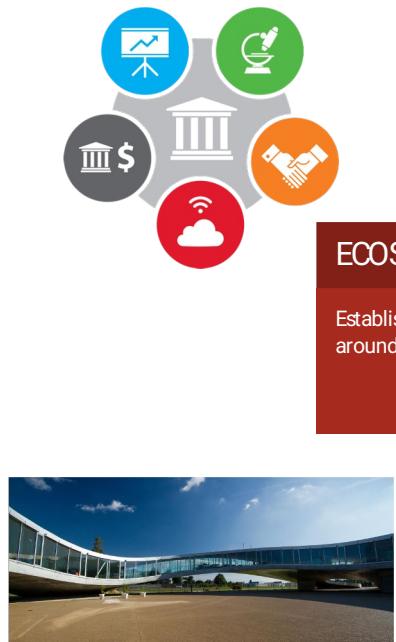


Corporate Training & Learning

Mission and Vision

- Bring players in EdTech together in one place in order to create a market place around Education and EdTech
- Focus on future learning solutions / future skills
- Long-term partnerships (not a short-term incubator)
- Help to accelerate growth
- Sustainability / long term positive impact on society

EdTech Collider - Facts



SPACE
Physical co-working space / Virtual space



Independent Association/NPO (04/17)
Four EPFL Professors
93+ Members (EdTech startups)
Support: EPFL, Jacobs Foundation, Credit Suisse

VISIBILITY
Create an image / branding and reputation Swiss EdTech Collider



PROXIMITY EPFL

Close to EPFL Research (Digital Learning / Learning Sciences)

LEARN - Center for Learning Sciences EPFL

COLLISIONS

Nourish the 'collision' of knowledge / sharing ideas / creating co-operations / collaborations (members/external)



+175 Collisions/Events

93 StartUps (status: 01.2022)



Follow Us

Twitter: [@SwissEdTech](#)

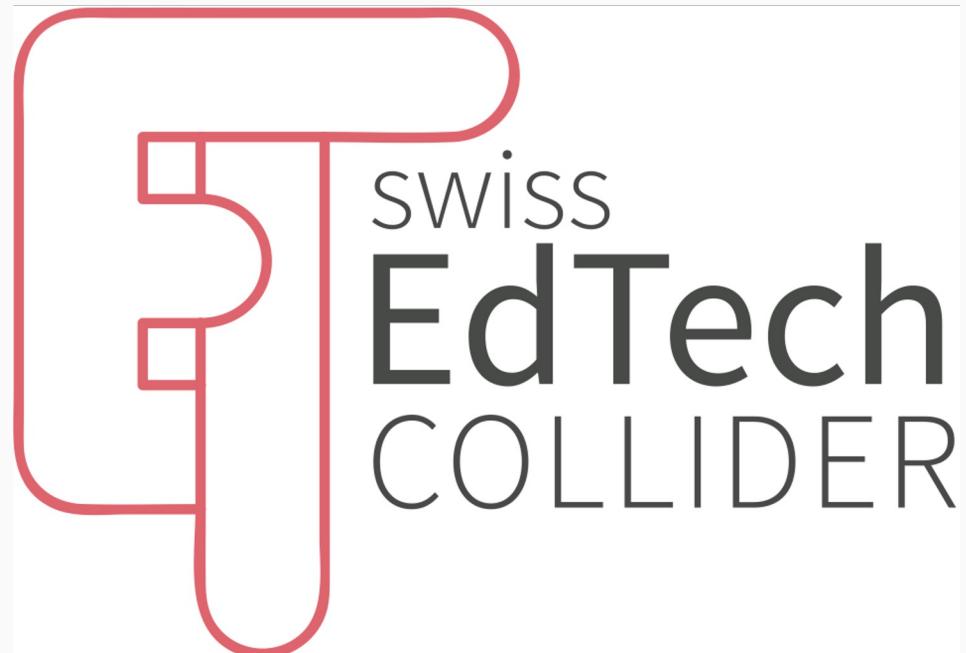
LinkedIn: [Swiss-EdTech-Collider](#)

Instagram: [Swiss_EdTech_Collider](#)

Facebook: [SwissEdTech](#)

contact@edtech-collider.ch

www.edtech-collider.ch



The two participating StartUps

- Dybuster Alemira (Marco Bär)
- Taskbase (Anette Hunziker)



Up next...

- Detailed information regarding the project: milestones, guidelines, grading, data sets, etc. [lab session today]
- Setting up GitHub and Jupyter notebook for the lecture and project [lab session on Wednesday]

Remember

- Register for the course on IS Academia
- Bring your laptop!
- You find everything on...

Moodle:

<https://moodle.epfl.ch/course/view.php?id=16434>
