

Machine Learning for Behavioral Data (CS-421)

Introduction

February 21, 2022

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/mlbd-lecture>

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

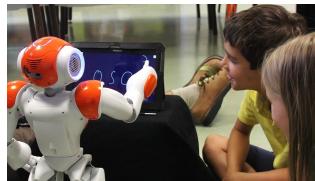
0 comments

The image shows a single message card from a 'SpeakUp Chat' platform. 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 the text '0 comments'. The entire card has a blue border.

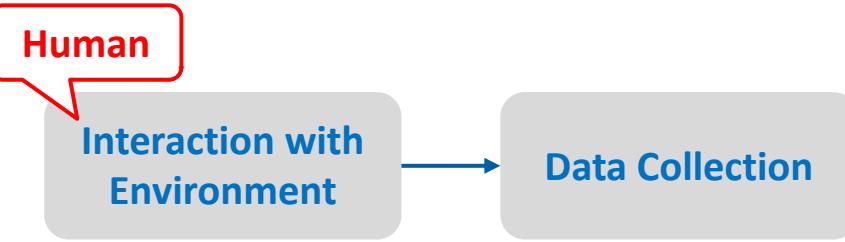
What is ML for Behavioral Data?

Human

Interaction with
Environment

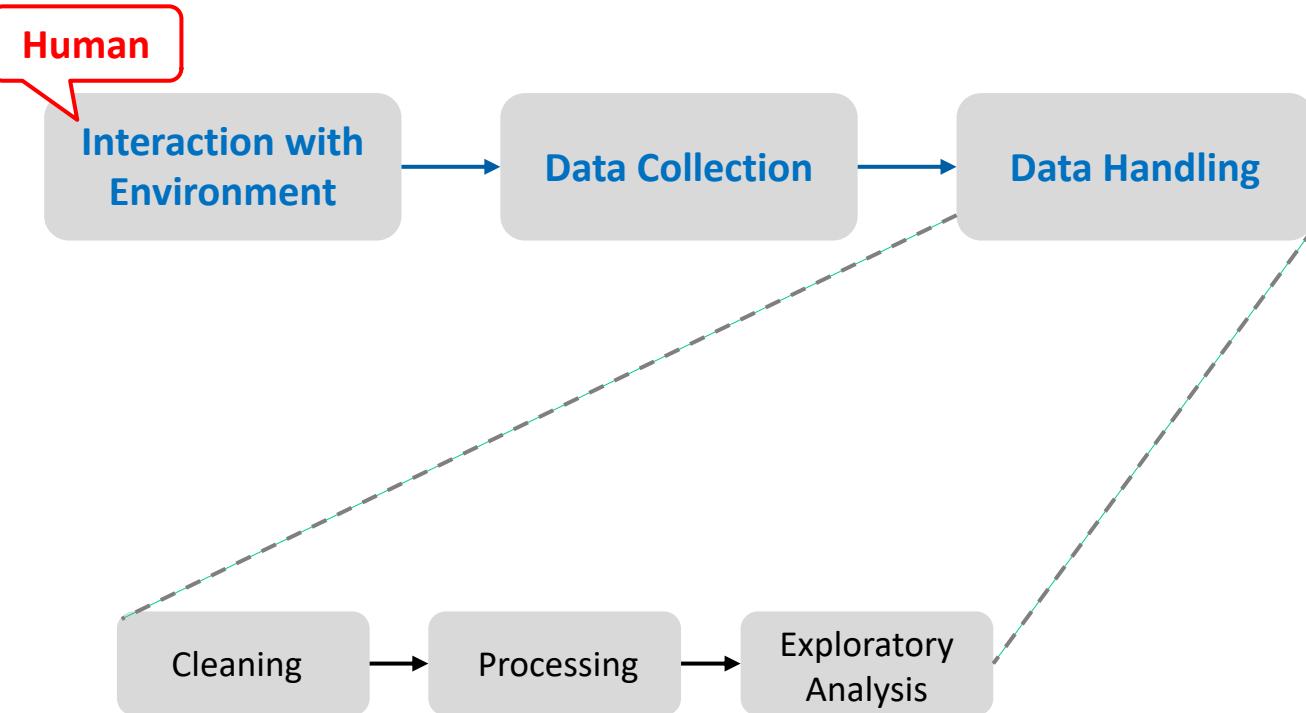


What is ML for Behavioral Data?

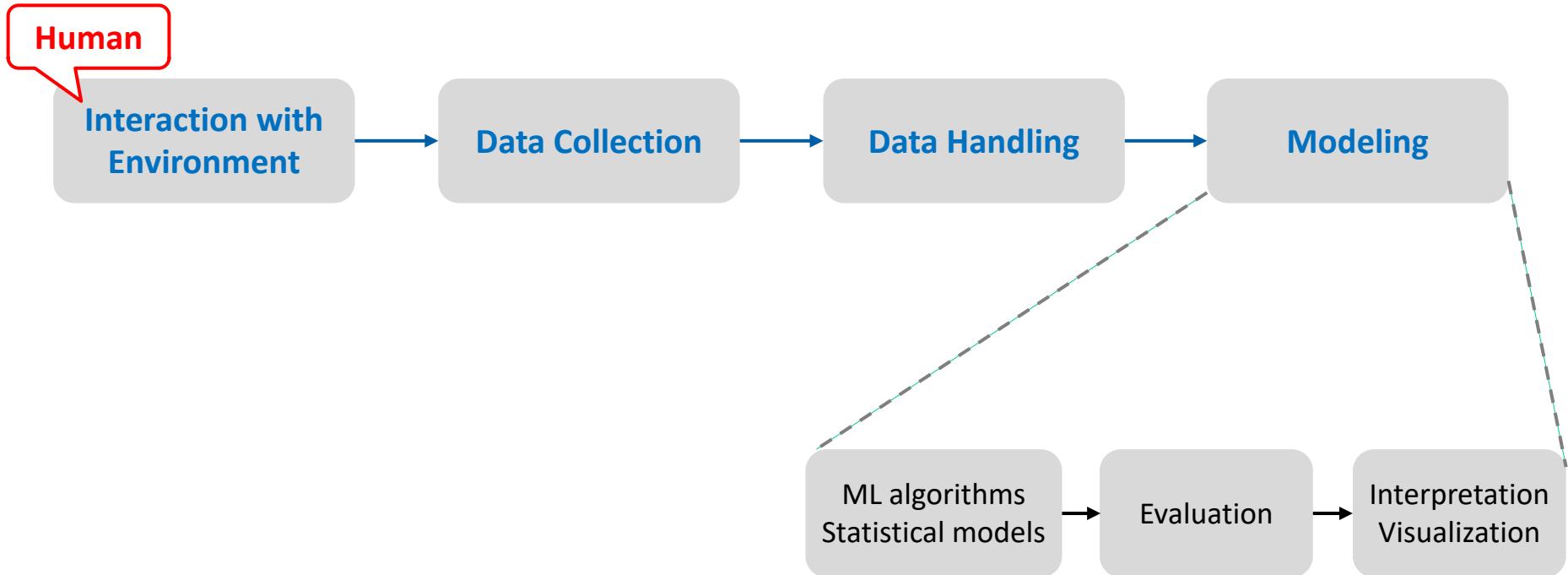


- Clickstream
- Text
- Categorical Data
- Images
- Video
- Sensor Data
- ...

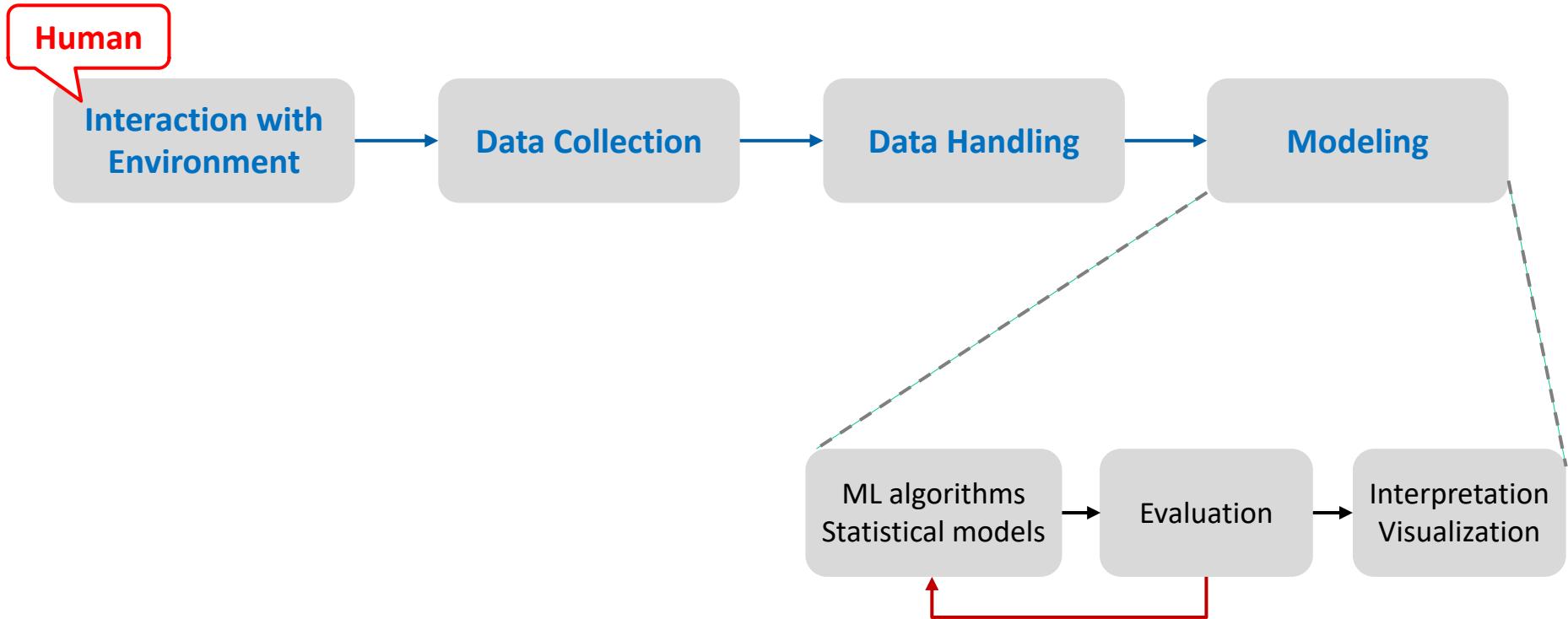
What is ML for Behavioral Data?



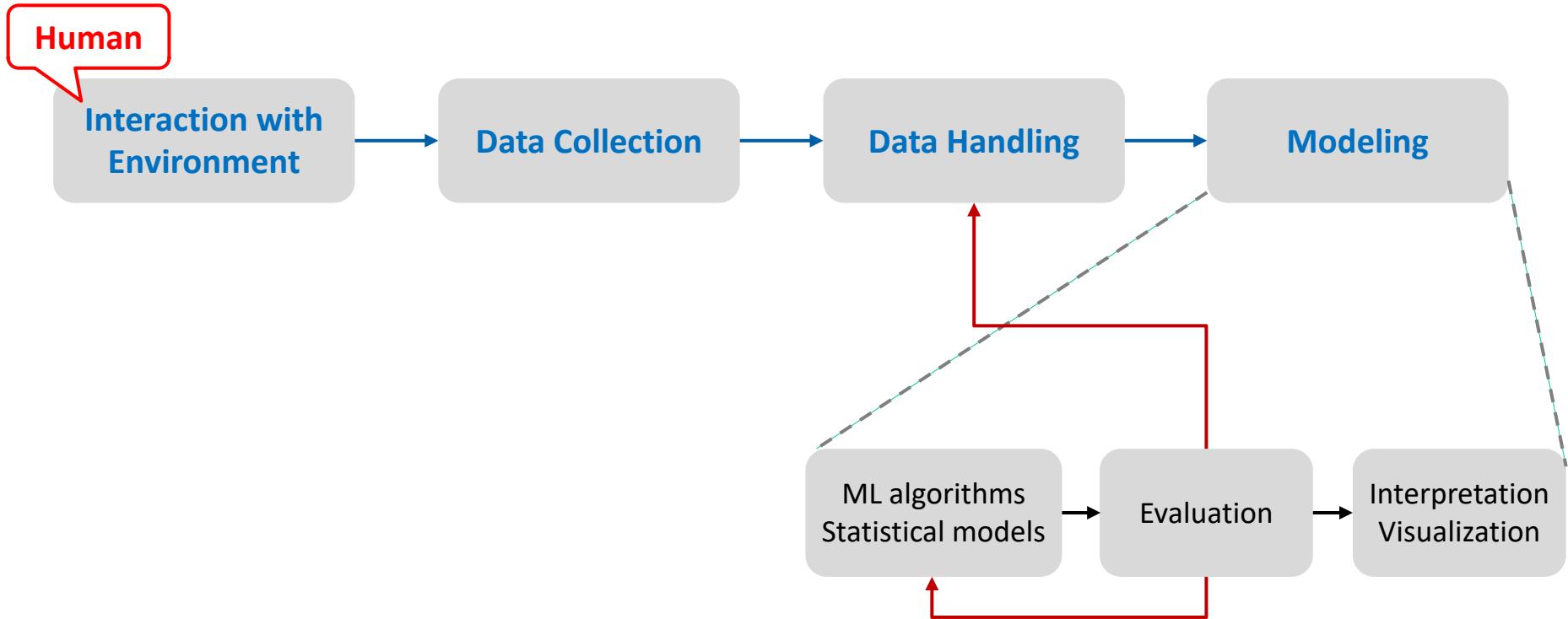
What is ML for Behavioral Data?



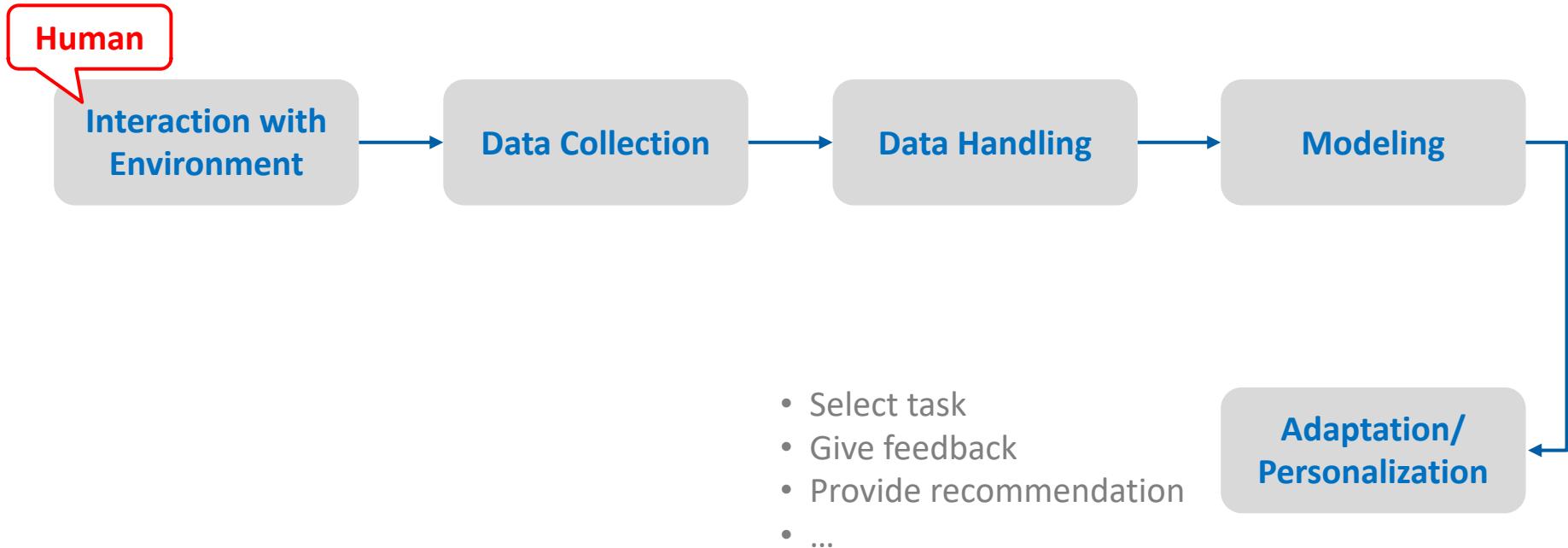
What is ML for Behavioral Data?



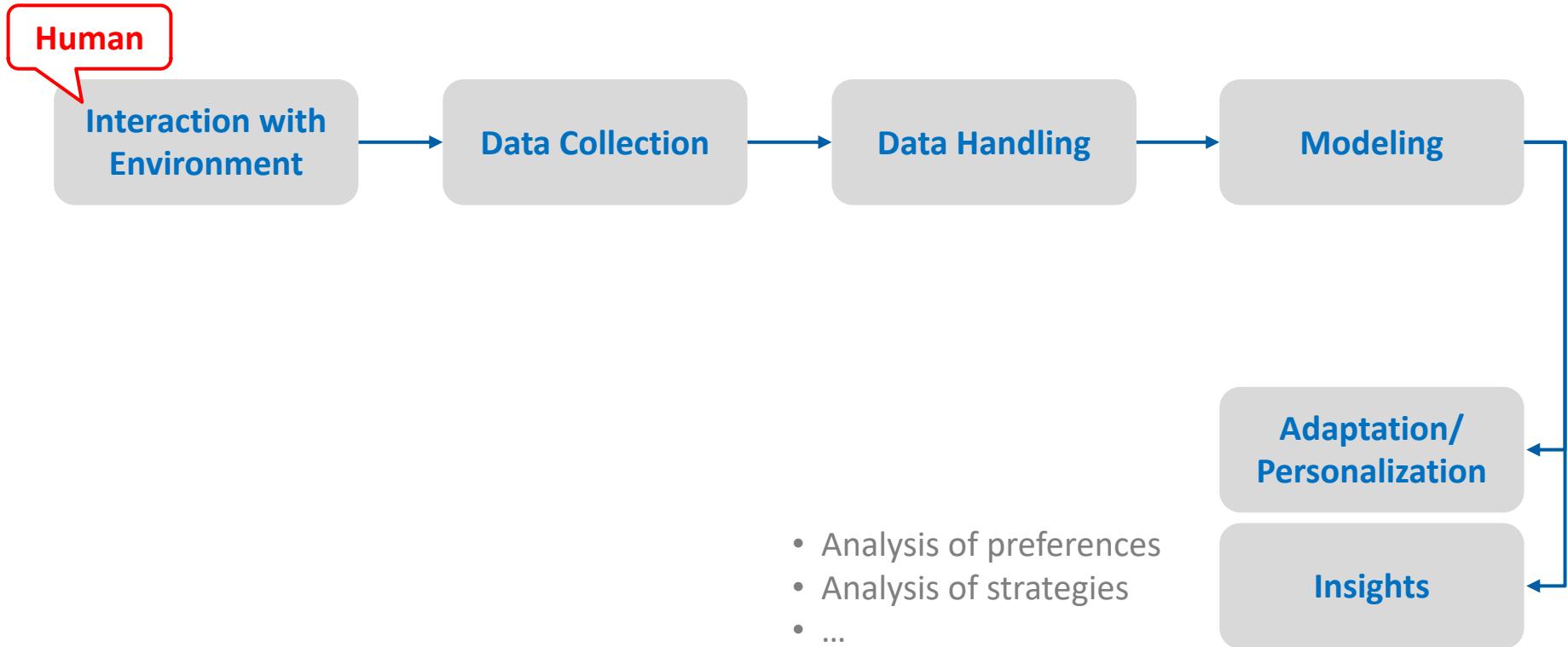
What is ML for Behavioral Data?



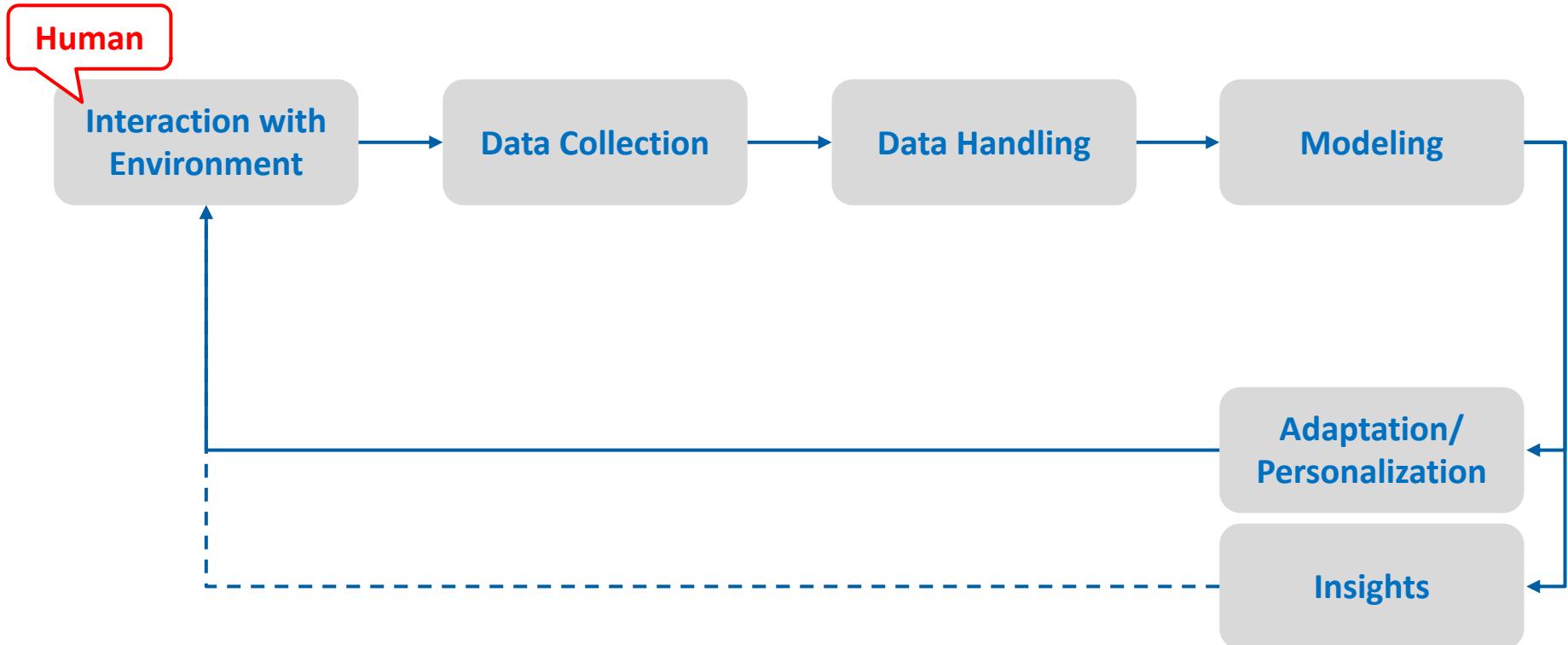
What is ML for Behavioral Data?



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
8	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
8	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
8	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
8	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
8	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
8	Time Series Prediction

Complete pipeline for one use case:

- Data exploration
- Prediction
- Model evaluation

Lecture Syllabus

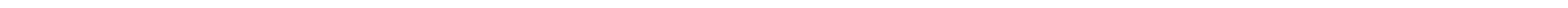
Week	Lecture/Lab
1	Introduction
2	Data Exploration
3	Regression
4	Classification
5	Model Evaluation
6	Time Series Prediction
7	Time Series Prediction
8	Time Series Prediction

Supervised learning on time series:

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

Lecture Syllabus

Week	Lecture/Lab
9	Spring Break
10	Guest Lecture: Neuroscience
11	Unsupervised Learning
12	Unsupervised Learning
13	Ethical Machine Learning
14	Ethical Machine Learning
15	Project Presentations



Lecture Syllabus

Week	Lecture/Lab
9	Spring Break
10	Guest Lecture: Neuroscience
11	Unsupervised Learning
12	Unsupervised Learning
13	Ethical Machine Learning
14	Ethical Machine Learning
15	Project Presentations

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

Lecture Syllabus

Week	Lecture/Lab
9	Spring Break
10	Guest Lecture: Neuroscience
11	Unsupervised Learning
12	Unsupervised Learning
13	Ethical Machine Learning
14	Ethical Machine Learning
15	Project Presentations

- Data Ethics
- Fairness
- Explainability

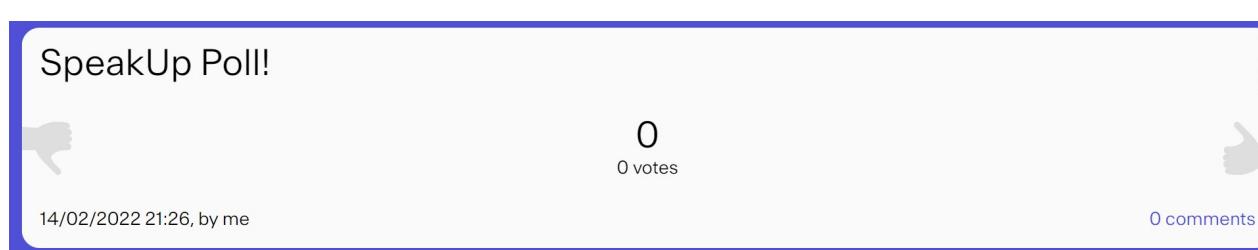
Lecture/Lab

- Monday, 15:15 – 17:00
- INR 113
- Lecture slides will be uploaded to Moodle
- Demo Jupyter Notebook will be uploaded to Moodle



Should we record the lecture?

- We can record the live sessions (without interactions) and upload it to Moodle



Project

- Teams of 3 people
 - We will provide the data sets
 - 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 **short presentation** in the last week of the semester
 - Final project (Code + Report) delivered by **June 12, 2022 23:59 CET**
-

Project (Office) Hours

- Wednesday, 8:15-10:00
- 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	Detailed project presentation	-
2	Introduction to tasks for M2	<i>M1: Preferences on team members and start-up</i>
3	Office hours	
4	Office hours	
5	Introductions to tasks for M4	<i>M2: Individual exploration of selected data set</i>
6	Individual discussion with teams	<i>M3: selection of research question and approach</i>
7	Office hours	
8	Office hours	

Project Schedule

Week	Project Hours	Milestones
9	Spring Break	
10	Office hours	
11	Individual discussion with teams	<i>M4: submission of results for first research question M5: ideas for extension (+ approach)</i>
12	Office hours	
13	Office hours	
14	Office hours	
15	Project Presentations	<i>M6: Final Presentation</i>
16		<i>M7: Hand in report and code base</i>

Grading

- **50% Project**
 - Teams of 3 people
 - 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?

- Introduction to Machine Learning
- Machine Learning
- Applied Data Analysis
- ...

SpeakUp Poll!

0
0 votes

14/02/2022 21:26, by me

0 comments

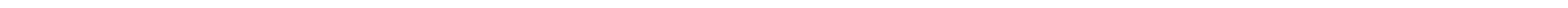
Course Prerequisites

- 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 head assistants
- Project:
 - GitHub: <https://github.com/epfl-ml4ed/mlbd-2022>
 - 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

Student Assistant



Francesco Salvi
francesco.salvi@epfl.ch

Feedback

- Course is still new -> please give feedback
- 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

The image shows a single message card from a digital platform called "SpeakUp Chat". The message content is "Active learning – what is it?". On the left side of the message, there is a thumbs-down icon. On the right side, there is a thumbs-up icon. Below the message, it says "0 votes" and "0 comments". The timestamp "14/02/2022 21:25" and the author "by me" are also visible. The entire message card is enclosed in a blue border.

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

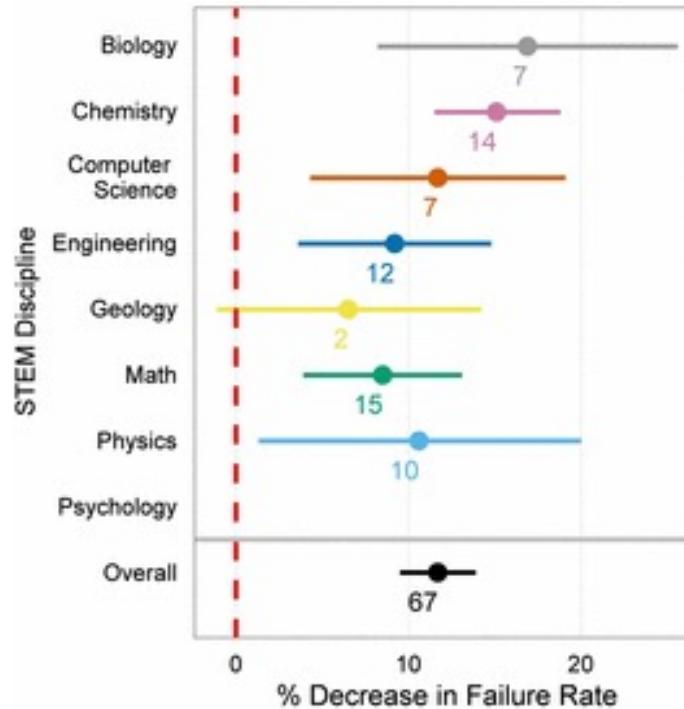
Jupyter Notebook

Demonstration

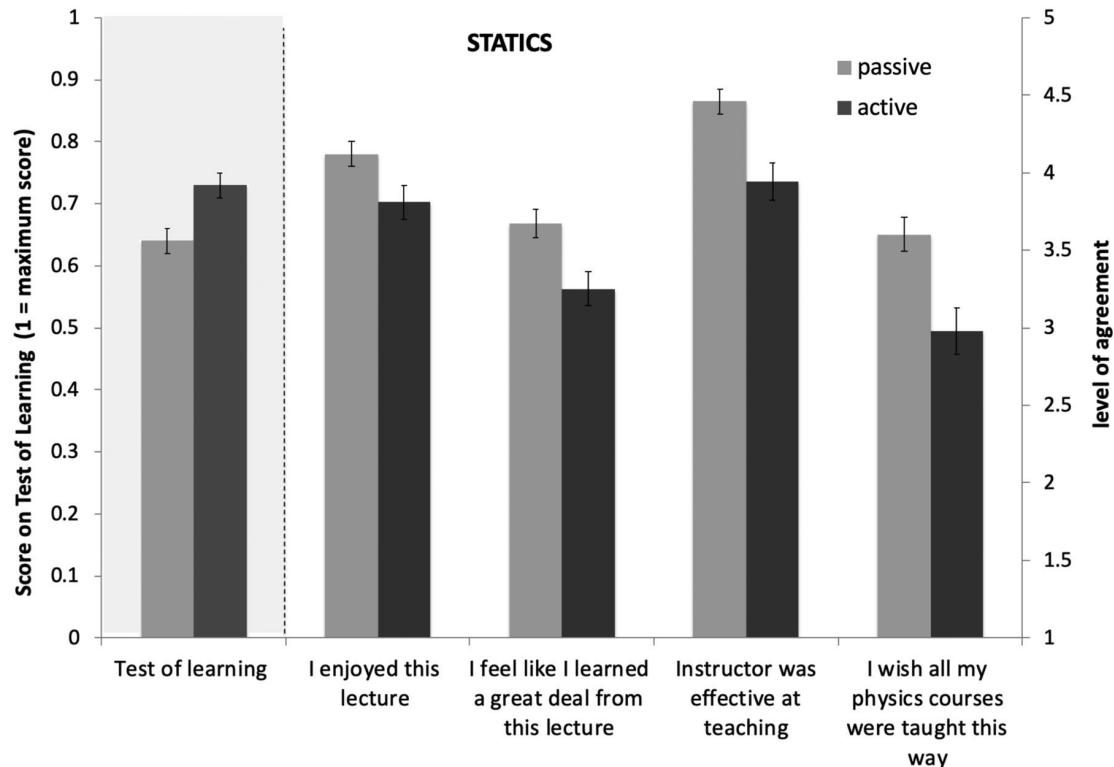
Worked examples

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 !

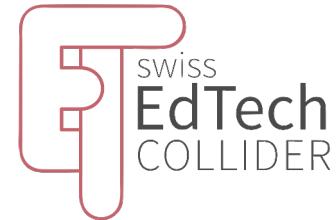
Questions?



Today

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

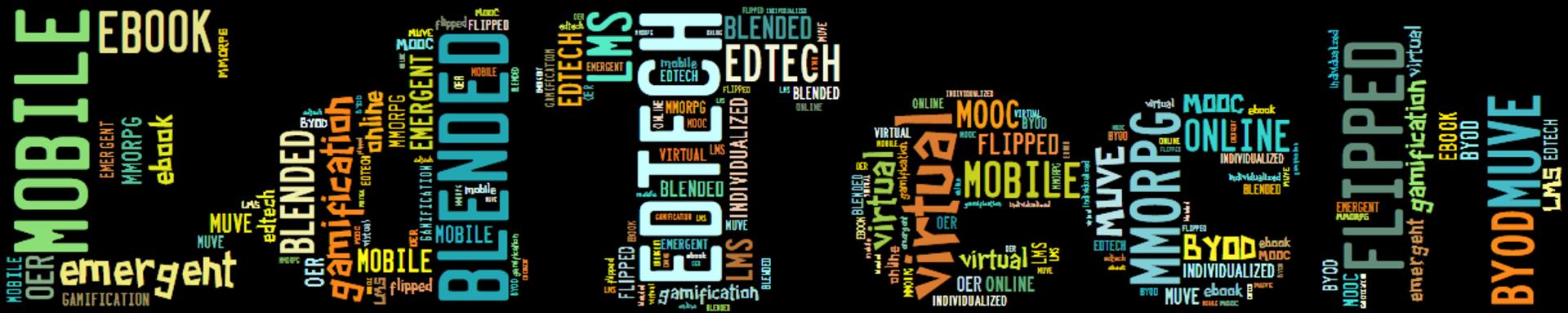
Swiss “EdTech” Hub



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**

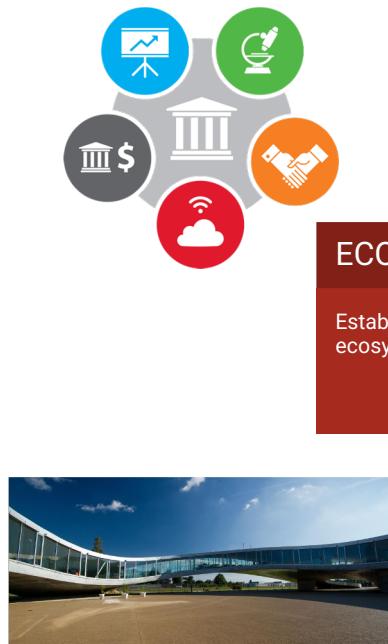


Corporate Training & Learning

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



93 StartUps (status: 01.2022)



Follow Us

Twitter: [@SwissEdTech](#)

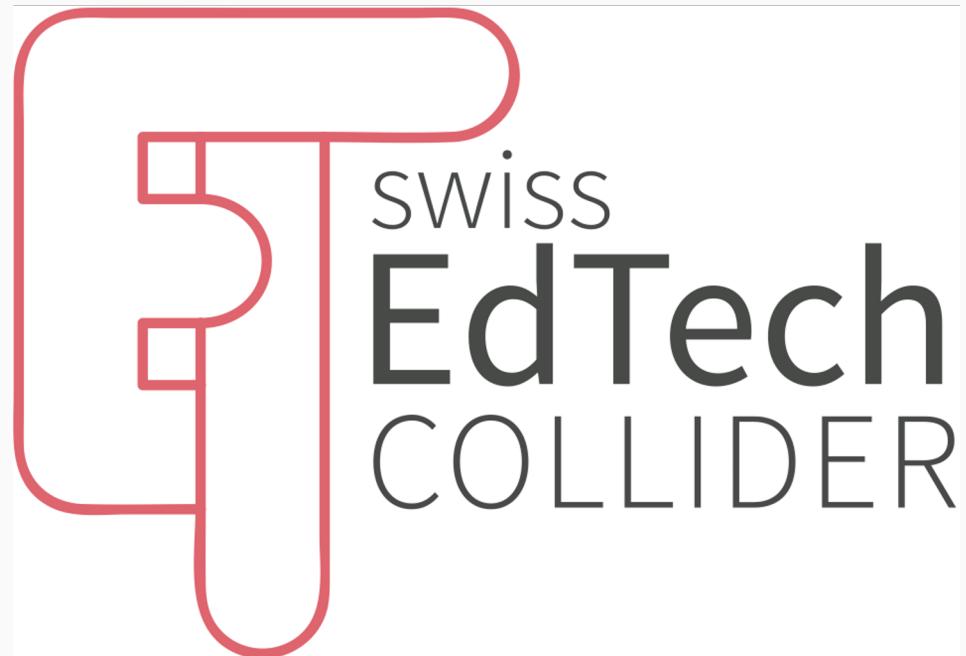
LinkedIn: [Swiss-EdTech-Collider](#)

Instagram: [Swiss_EdTech_Collider](#)

Facebook: [SwissEdTech](#)

contact@edtech-collider.ch

www.edtech-collider.ch



Two participating StartUps

- Classtime (Jan Rihak)
- Taskbase (Samuel Portmann)



Up next...

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