Introduction to ML & DL L+Pr

Béla J. Szekeres, PhD

Lecture 1

ELTE Faculty of Informatics, Szombathely, Hungary

About the instructor

- Béla J. Szekeres
- BSc, MSc (2012), PhD (2017) in applied mathematics at ELTE
- special addiction to algorithms, artificial intelligence
- Mathematics expert in ML and DL (2022)
- Szombathely since 2018
- `szekeres@inf.elte.hu`, room: B106

About the course

Deep learning/Machine learning

- Canvas + `https://github.com/szbela87/ml_22_elteik`
- Grade: homeworks (20%) + practice (20%) + mini-projects (60%) - 4 + 1 mini-projects

Schedule

- 1st half of the semester
  - Machine learning algorithms, reminder + XGBoost, hyperparameter tuning - Optuna, Hyperopt
  - Google Colab, NumPy, Pandas, Matplotlib, Scikit-Learn
- 2nd part
  - Neural networks, deep learning, computer vision, time series
  - PyTorch

Important questions

- What is machine learning?
- How does machine learning work?
- When to use machine learning?
- Types of learning.

How machine learning works

- **Decision proces**: In general, machine learning algorithms are used to make a prediction or classification. Based on some input data, which can be labelled or unlabeled, your algorithm will produce an estimate about a pattern in the data.
- **Error function**: An error function serves to evaluate the prediction of the **model**. If there are known examples, an error function can make a comparison to assess the accuracy of the model.
- **Optimization process**: fitting the model parameters.

Types of learning

Supervised learning

- Labeled dataset.

Unsupervised learning

- Unlabeled dataset.

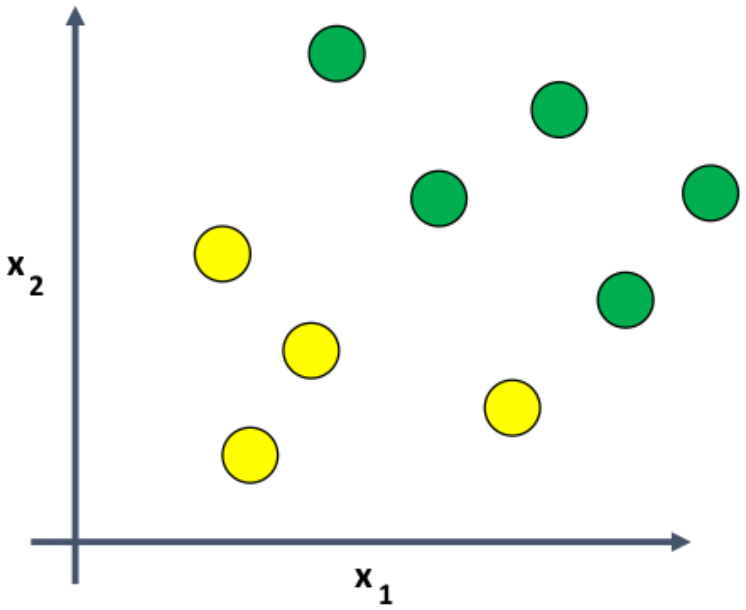
Reinforcement learning

- No dataset, interactions with the environment.

1. Supervised learning

We have a dataset: with samples and labels as well.

Most of the machine learning techniques rely heavily on dataset.



- We

give the algorithm the right answers!

- During the **training procedure**:

$\rightarrow$ the

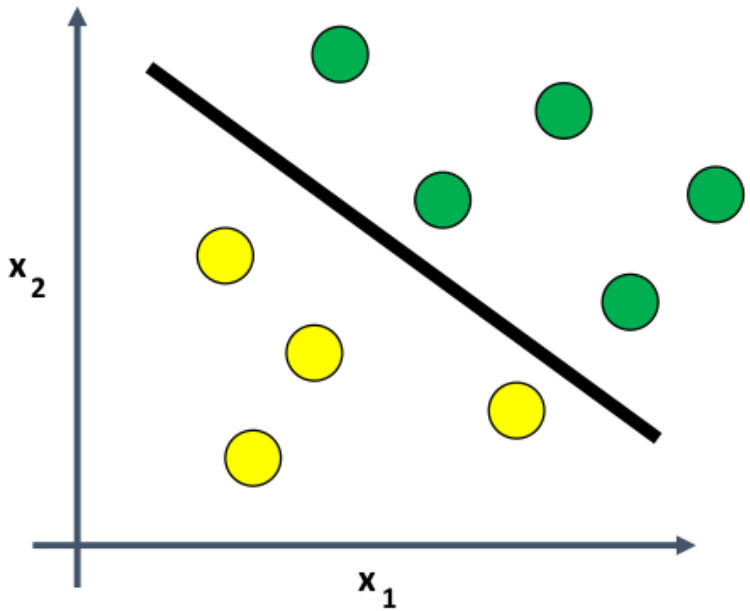
algorithm will find out what is the

difference between the two classes

1. Supervised learning

We have a dataset: with samples and labels as well.

Most of the machine learning techniques rely heavily on dataset.



- We

give the algorithm the right answers!

- During the **training procedure**:

$\rightarrow$ the

algorithm will find out what is the

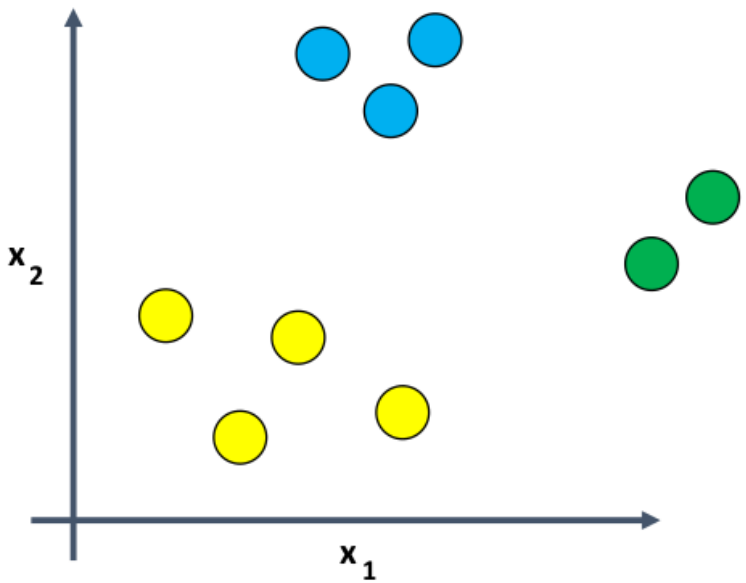
difference between the two classes

2. Unsupervised learning

We have a dataset: samples without labels.

The algorithm will find some patterns in this unlabeled dataset.

For example: clustering algorithms.



- We don't give the algorithm the right answers!
- During the **training procedure**:

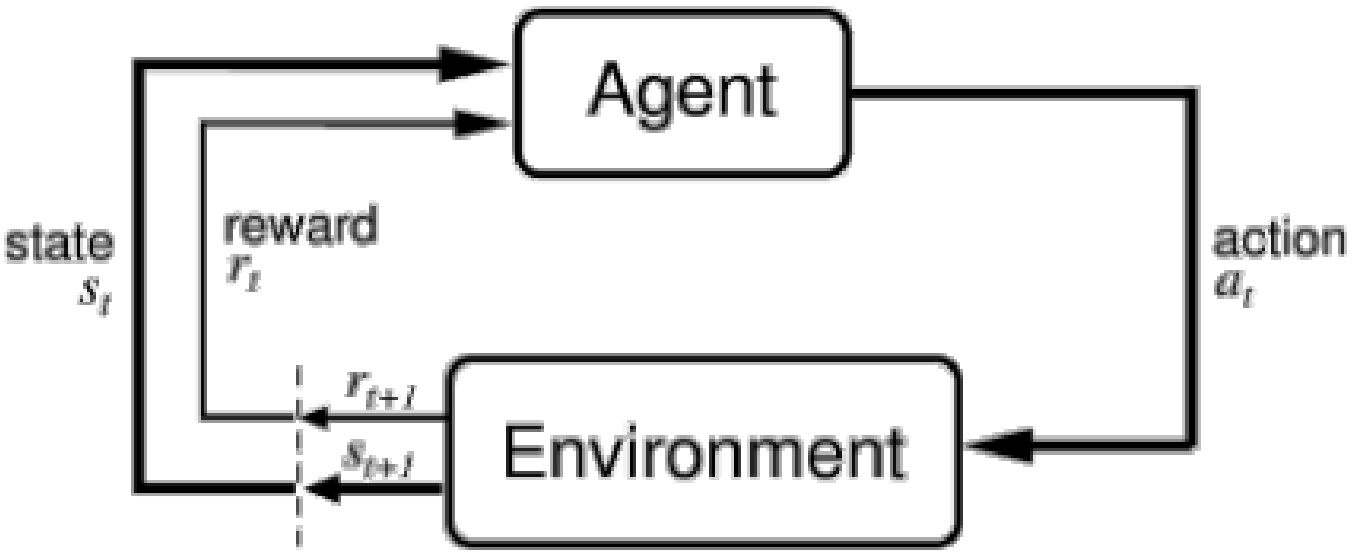
$\rightarrow$ the algorithm will find some relevant info about the features to make the classification.

3. Reinforcement learning

We don't have any dataset at all.

The algorithm (agent) will interact with the environment (states) and figure out what to do (actions).

Upon doing actions the agent receives a reward/penalty.



3. Reinforcement learning

0		X
	0	
		X

- reinforced learning

can be used to learn playing tic-tac-toe game

- **environment:** cells on the board
- the **agent** (computer player)

makes a move according to the board states

- **action**: where to put the X
- eventually the game will end: agent will recieve either **reward** (win) or **penalty** (lose).

The agent initially plays poorly: but after training procedure (so playing a lot) it will able to choose the right actions.

Examples of applications

- Face recognition, Image classification
- Game playing
- Speech recognition
- Text-to-scpeech generation
- Handwritin transcription
- Machine translation
- Medical diagnosis
- Self-driving cars
- Digital assistants
- Ads, social reccomendations

Topics in the next week

Lecture

- Linear and logistic regression,
- Naive Bayes Classifier, KNN
- Overfitting/Underfitting

Practice

- Pandas.

Today's Practice

We will use the following libraries across the whole semester: Numpy, Pandas, Matplotlib, Scikit-learn

Reminders about

- NumPy
- Pandas
- Matplotlib

1st Homework - Canvas/Assignments

- Notebooks
- Due date: next friday, till 23.59 pm.