

Machine Learning I Course

Practicalities for this course

Class Format

Flipped Class format

Offline

Watch videos

Quiz

1 hour

*You can watch videos between 9h15-10am or before.
But you must watch the videos before coming to class!*

*The quizzes are optional – they are meant to prepare you for
the final exam that will be a multiplechoice.*

Class Format

10h15-11h00

Interactive lecture

Live *in class room CE2*

Live *on zoom*

11h15-13h00

Regular exercise session

Live *in class room CE2*

Live *on zoom & discord*

**3
hours**

Class Format

Interactions during interactive lecture

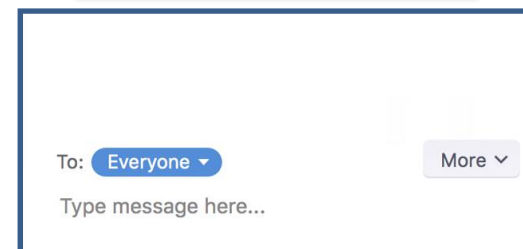


Ask live!



Type it in zoom chat box

Teaching assistant will reply live



Class Format

Interactions during exercises

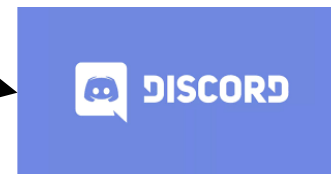


Raise hand and wait for an available TA



Zoom will be used only for presenting the solutions

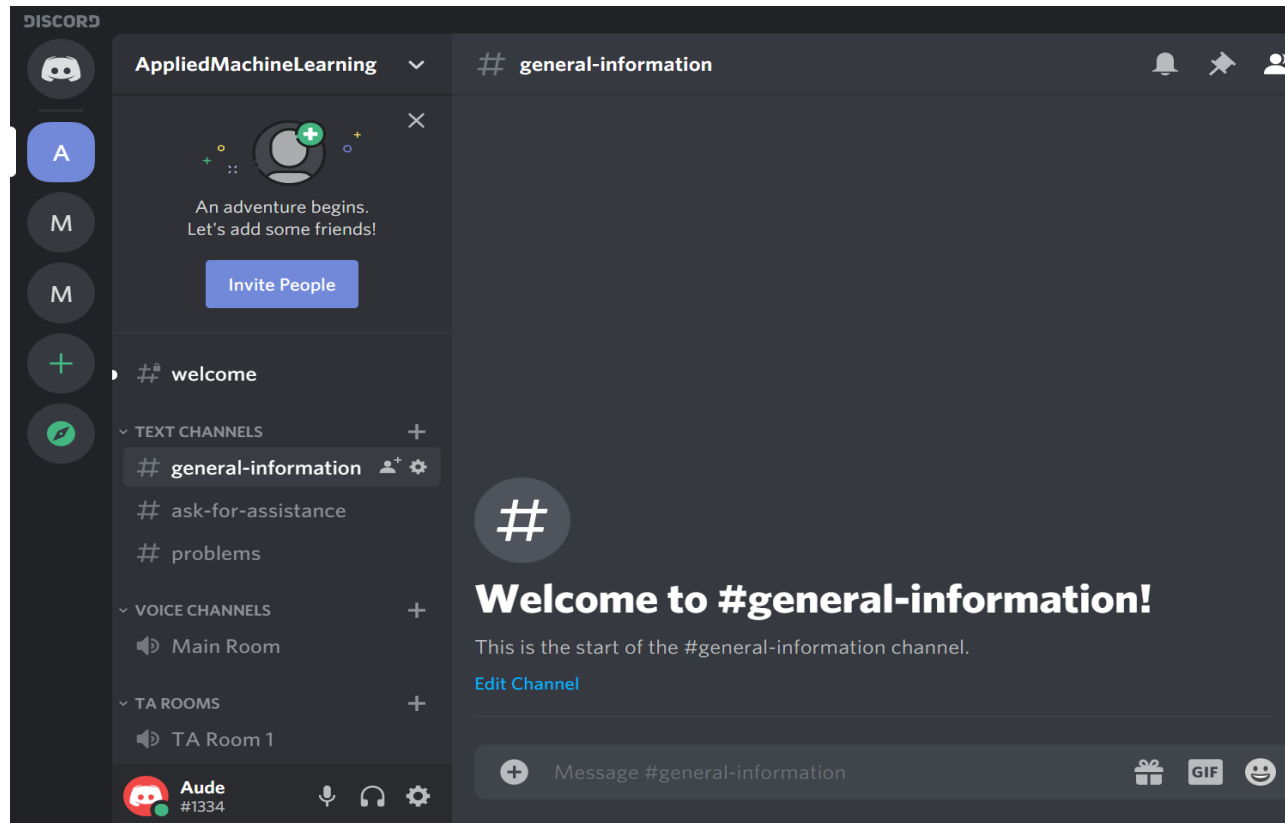
Solutions presented on zoom will be recorded and made available on moodle



**Students can work individually or in group
→ create their own room**

Go to the TA room or ask a TA to joint your room.

DISCORD: <https://discord.gg/TwTfKkv>



Questions posted on Discord will be answered solely during class exercises' hours.

Class Format

Practice Sessions on Computer

On site in classrooms BC 07 – 08

Live Online

Teacher
presentation



Type question in
zoom chat box

To: Everyone More

Type message here...

Questions &
Interactions



Students can work
individually or in group →
create their own room

Ask for a TA who
will join your room
(5 min. max)

4
hours

Material for the class

← → × ⌂ https://moodle.epfl.ch/course/view.php?id=14821

EPFL

Microengineering (MT) / MT - Master

Machine learning I

▼ General Informations

Course's general objective and approach

Real-world engineering applications must cope with a large dataset of dynamic variables, which cannot be well approximated by linear models. This course gives an overview of methods of Machine Learning for the analysis of non-linear, highly noisy, and multi-dimensional data. We will see methods for clustering, classification, and nonlinear regression.

Because machine Learning can only be understood through practice, by using the algorithms, the course is accompanied by machine learning algorithms with real-world data.



- Videos
- Quizzes
- Slides
- Exercises
- Solutions
- Lecture Notes

← → ↺ ⌂ https://tube.switch.ch/channels/7e86d16d

EPFL

LASA - Machine Learning Courses

Aude Billard, École polytechnique fédérale de Lausanne (EPFL)

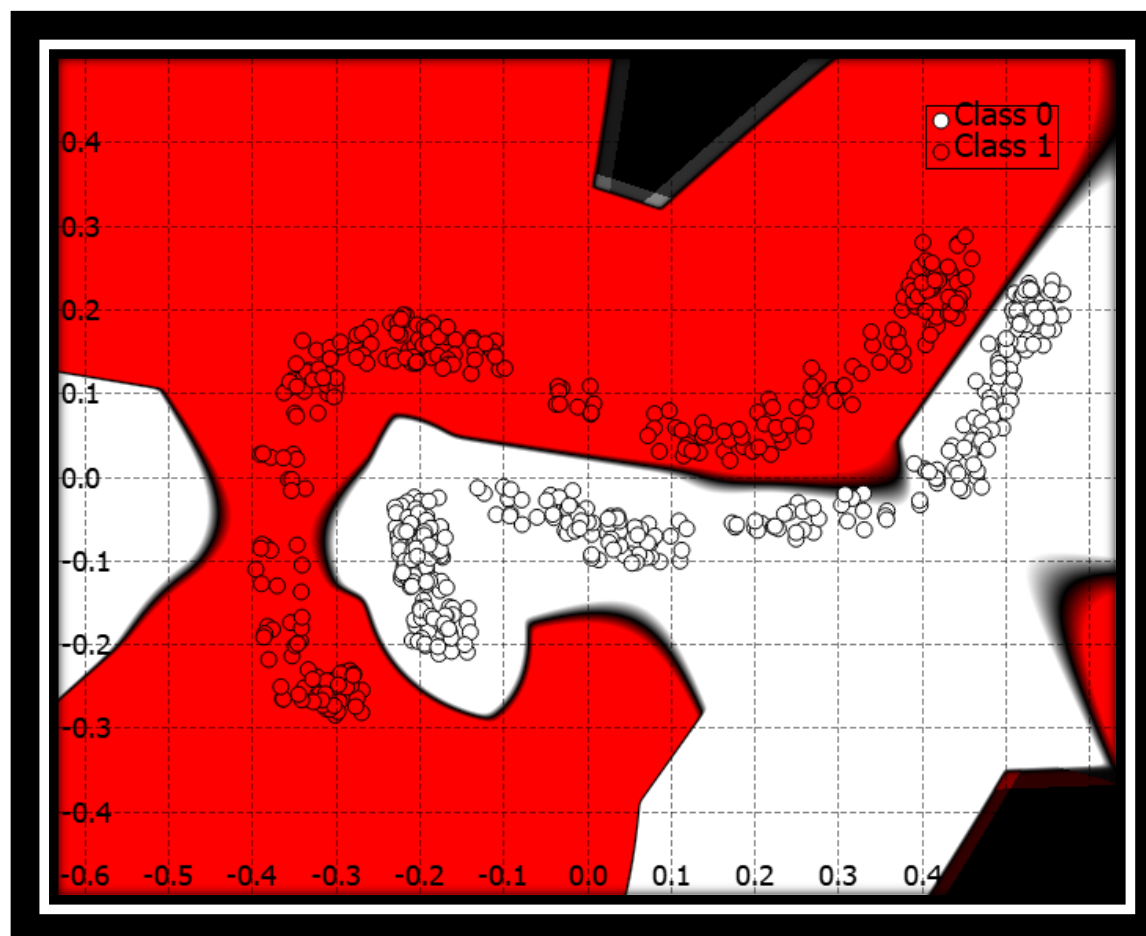
<https://tube.switch.ch/channels/7e86d16d>



Complete repository
of all videos

Software

<http://mldemos.epfl.ch>



Software

Instructions to install the software on moodle


<https://moodle.epfl.ch/course/view.php?id=14821>


80%






MOODLE

[FR](#) | [EN](#) | [D](#)

=====

Software *MachineLearningDemos*

A suite of algorithms has been implemented for you in the form of a user-friendly program that allows you to play with data and to study how each method performs in different tasks of classification, clustering and regression. The software runs on Windows and requires the .Net Framework 3.0 (which should be installed in your machine already, but if it isn't, pick it up [here](#)).

Download MLDemos [here](#)

=====

Grading Scheme

Midterm (0%) – On-line Quiz – Self-evaluation

Written Exam (100)% of the grade

Multiple-Choices

3 hours long

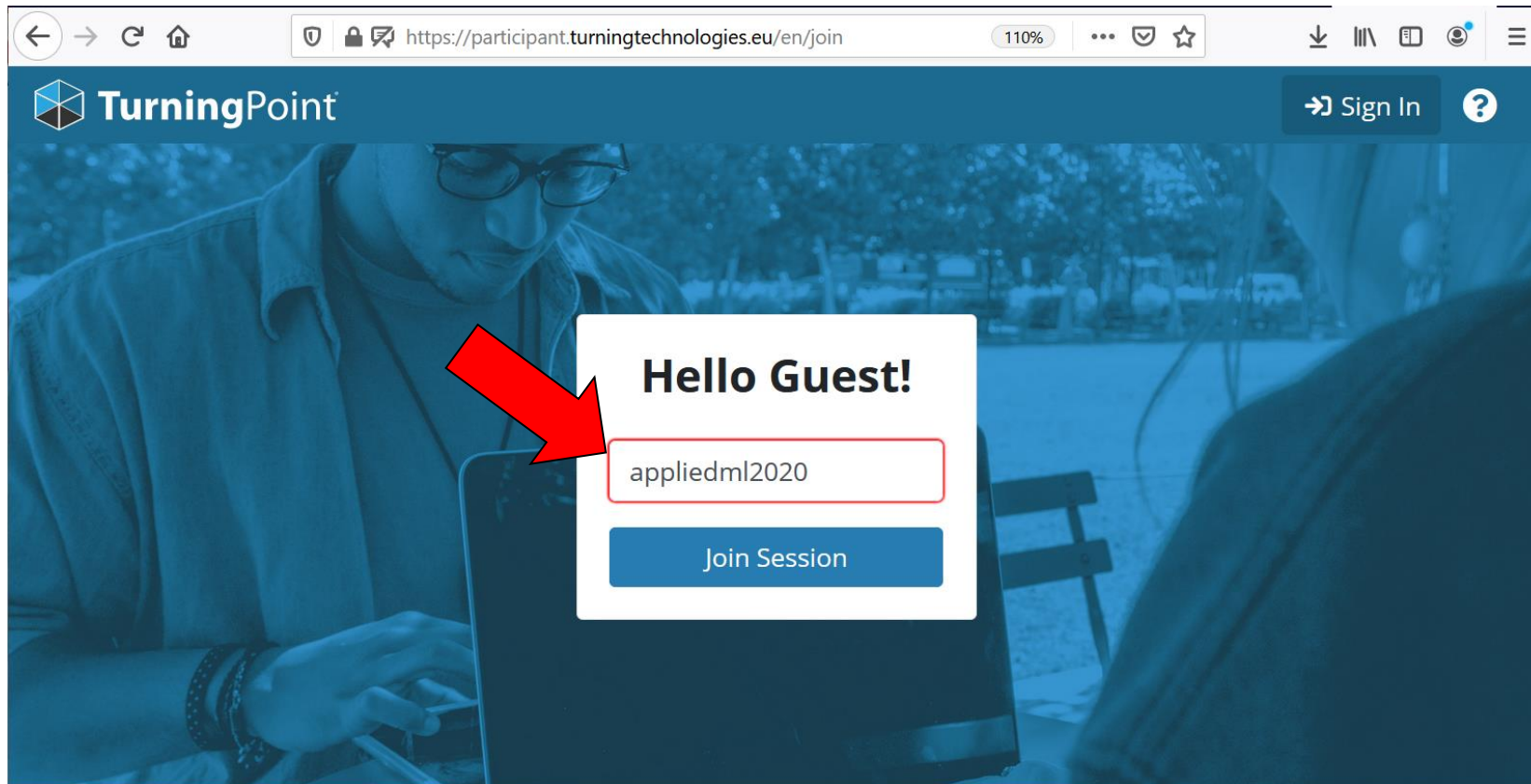
Closed book

Allowed 1 A4 recto-verso page with *handwritten notes*

Test polling system

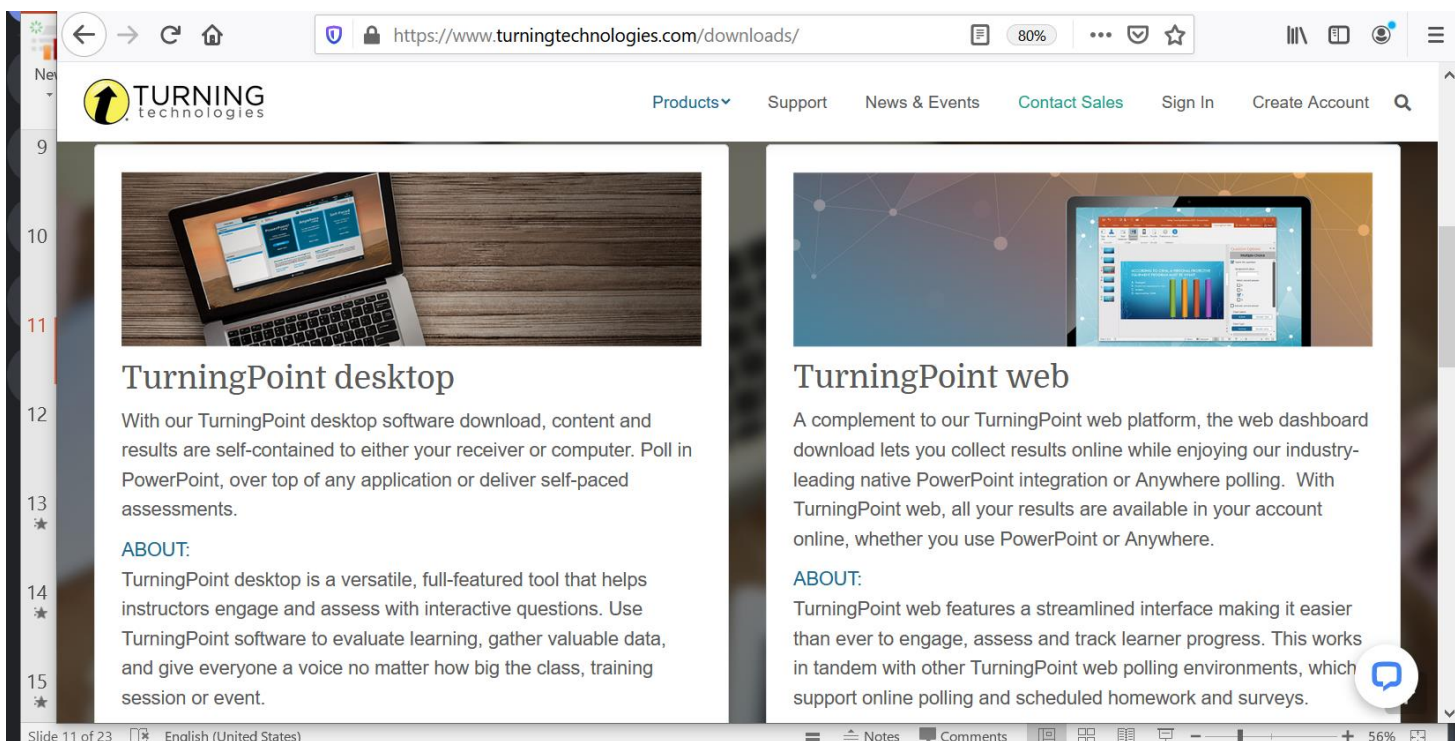
<https://participant.turningtechnologies.eu/en/join>

Access as GUEST and enter the session id: *appliedml2020*



Test polling system

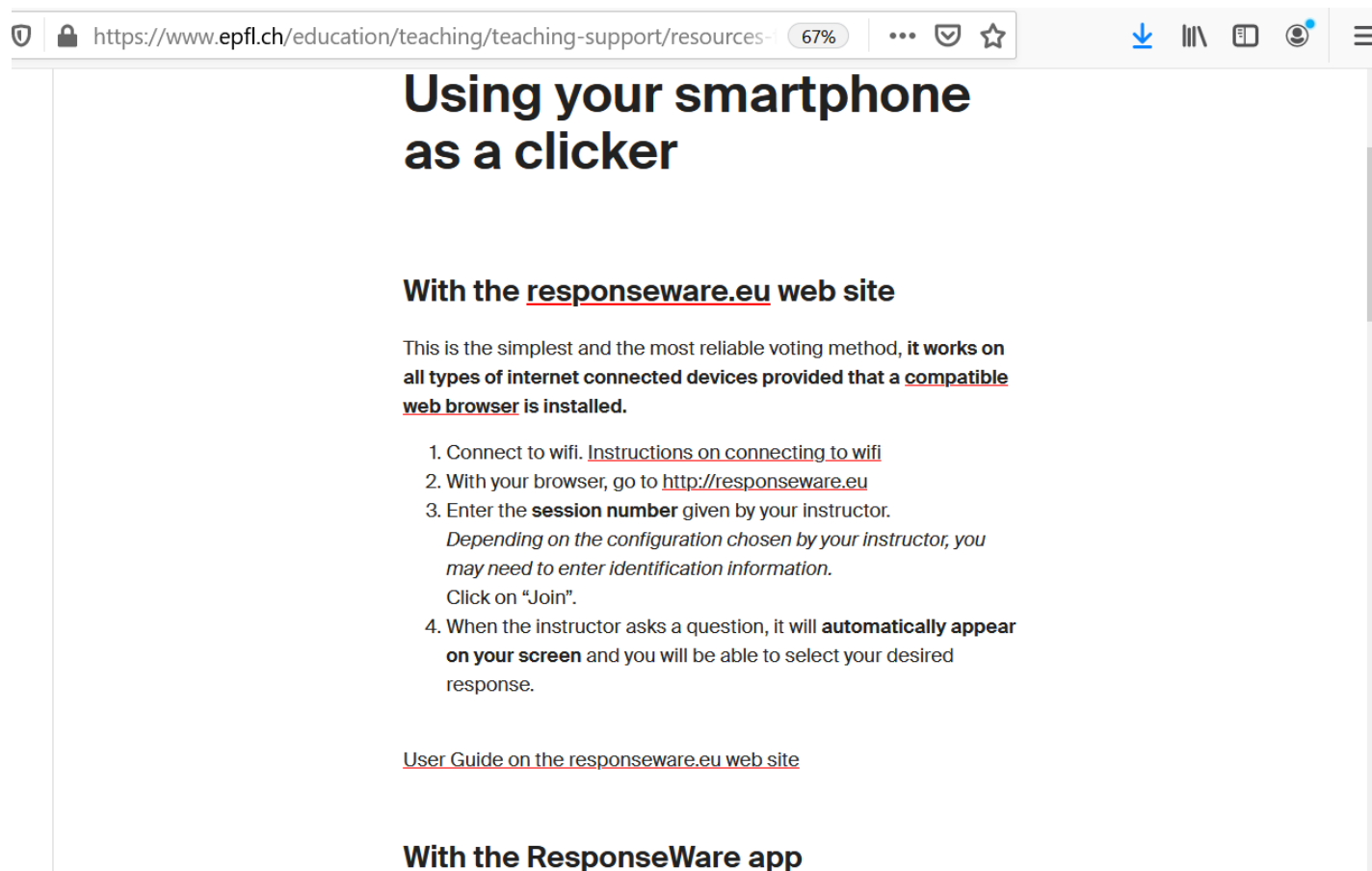
<https://www.turningtechnologies.com/downloads/>



You can also install either the mobile version (for android or iphone), or the desktop, but we encourage you to use only the web version.

Instructions

<https://www.epfl.ch/education/teaching/teaching-support/resources-for-students/student/using-your-smartphone/>



The screenshot shows a web browser window with the address bar displaying the URL: <https://www.epfl.ch/education/teaching/teaching-support/resources-for-students/student/using-your-smartphone/>. The page title is "Using your smartphone as a clicker". The main content area has a heading "Using your smartphone as a clicker" and a subheading "With the responseware.eu web site". Below this, a paragraph states: "This is the simplest and the most reliable voting method, **it works on all types of internet connected devices provided that a [compatible web browser](#) is installed.**". A numbered list follows: 1. Connect to wifi. [Instructions on connecting to wifi](#); 2. With your browser, go to <http://responseware.eu>; 3. Enter the **session number** given by your instructor. *Depending on the configuration chosen by your instructor, you may need to enter identification information.* Click on "Join". 4. When the instructor asks a question, it will **automatically appear on your screen** and you will be able to select your desired response. At the bottom, there is a link: [User Guide on the responseware.eu web site](#). The page is partially scrolled down, showing the beginning of the next section "With the ResponseWare app".

Using your smartphone as a clicker

With the responseware.eu web site

This is the simplest and the most reliable voting method, **it works on all types of internet connected devices provided that a [compatible web browser](#) is installed.**

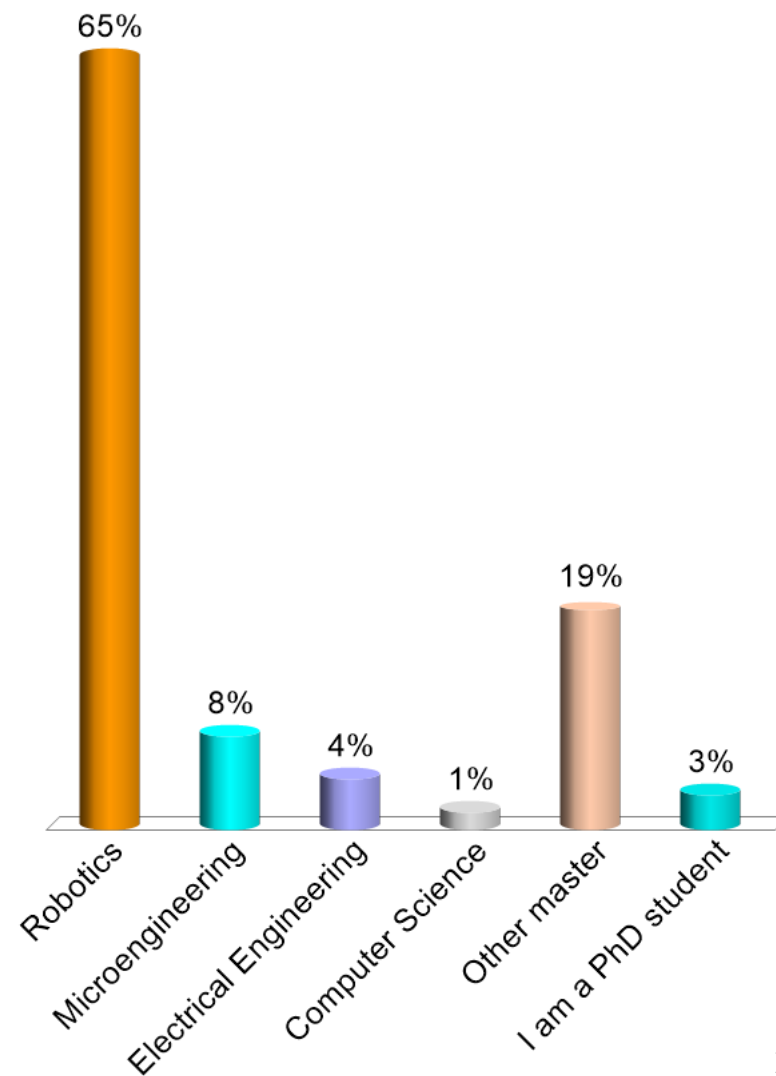
1. Connect to wifi. [Instructions on connecting to wifi](#)
2. With your browser, go to <http://responseware.eu>
3. Enter the **session number** given by your instructor.
Depending on the configuration chosen by your instructor, you may need to enter identification information.
Click on "Join".
4. When the instructor asks a question, it will **automatically appear on your screen** and you will be able to select your desired response.

[User Guide on the responseware.eu web site](#)

With the ResponseWare app

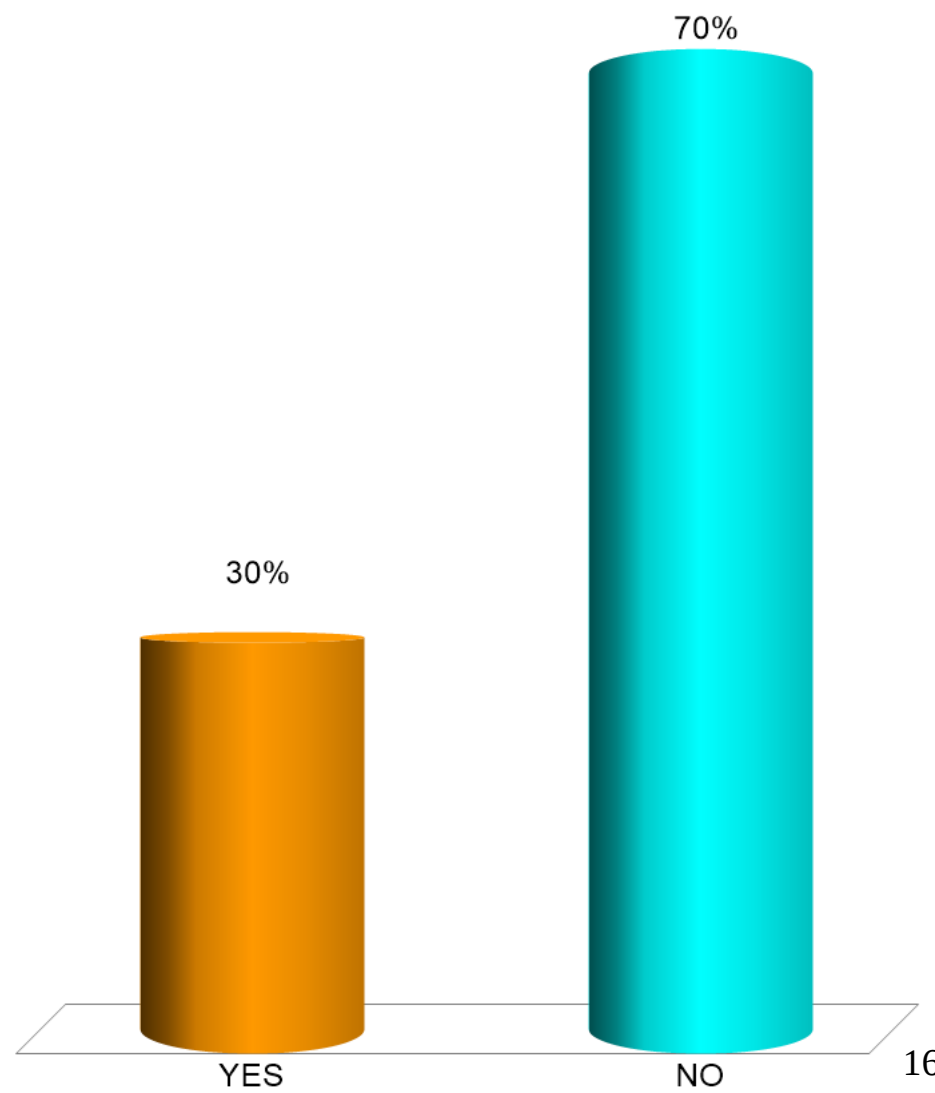
Which master do you follow?

- A. Robotics
- B. Microengineering
- C. Electrical Engineering
- D. Computer Science
- E. Other master
- F. I am a PhD student



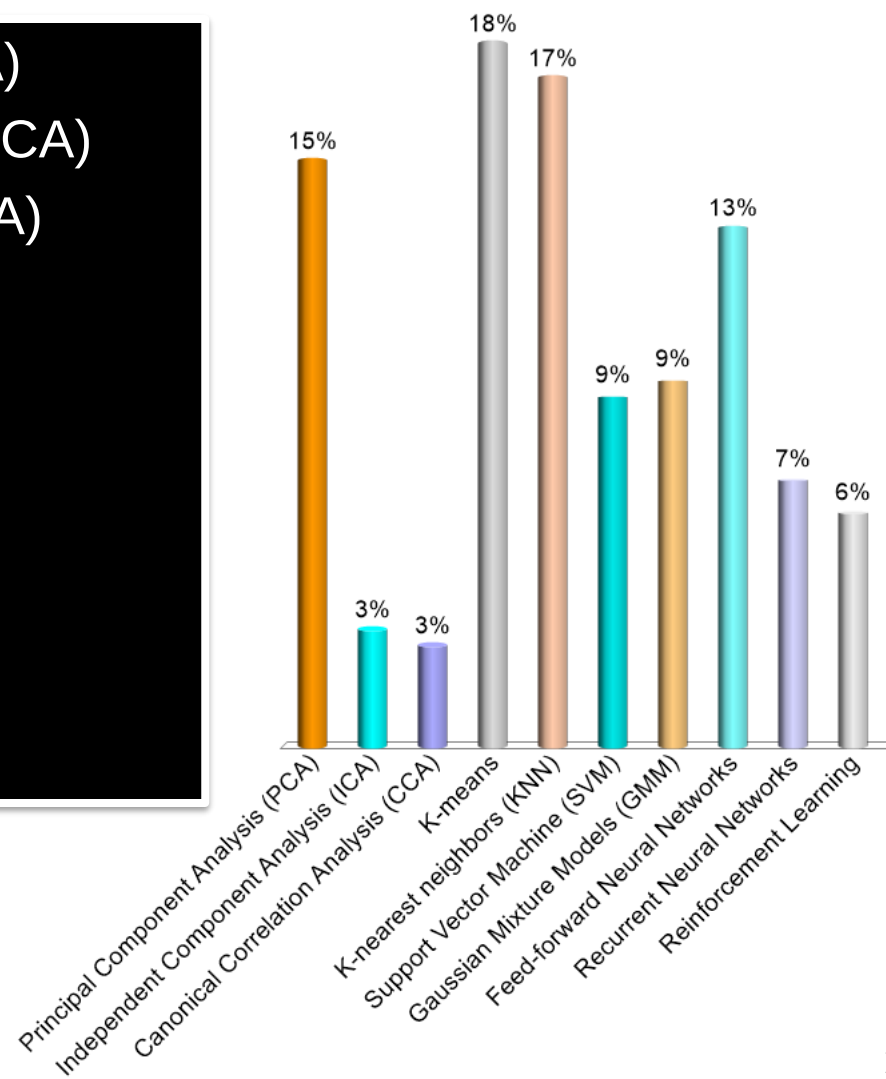
Have you taken a
Machine Learning course before?

- A. YES
- B. NO



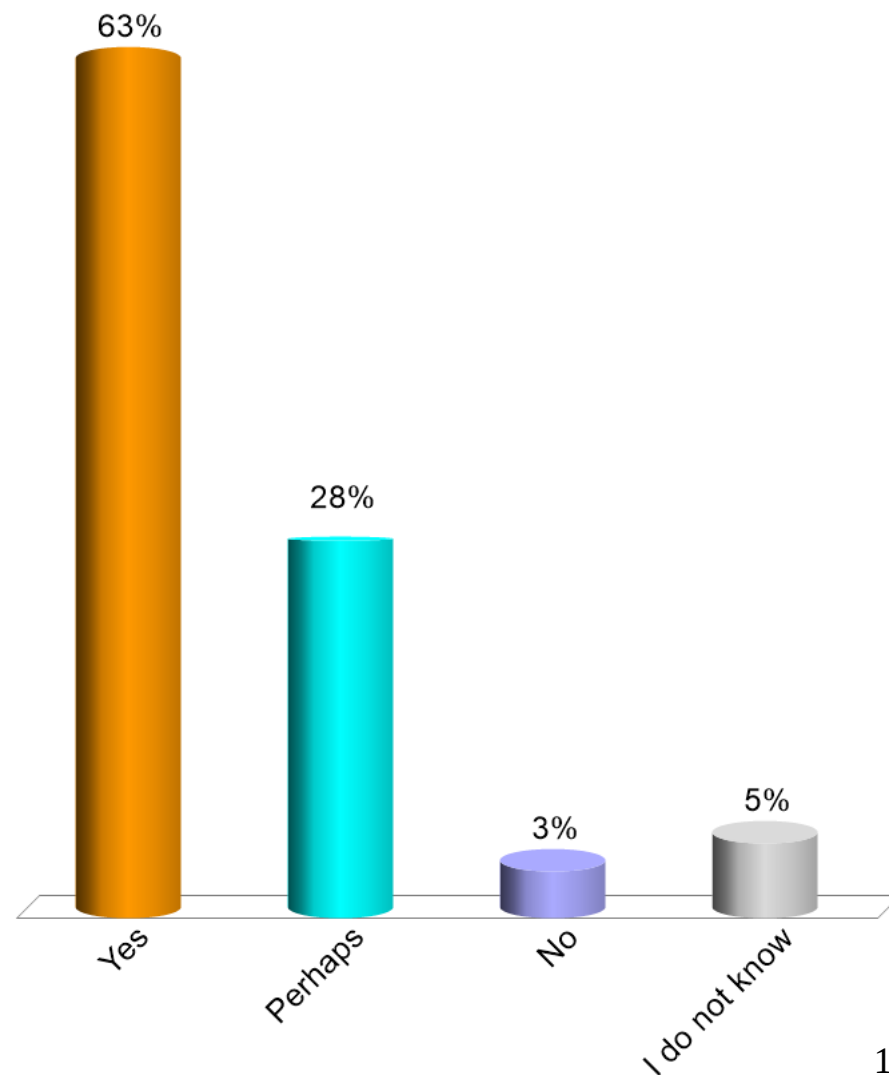
If you took a ML course before, which of these topics have you seen before

- A. Principal Component Analysis (PCA)
- B. Independent Component Analysis (ICA)
- C. Canonical Correlation Analysis (CCA)
- D. K-means
- E. K-nearest neighbors (KNN)
- F. Support Vector Machine (SVM)
- G. Gaussian Mixture Models (GMM)
- H. Feed-forward Neural Networks
- I. Recurrent Neural Networks
- J. Reinforcement Learning



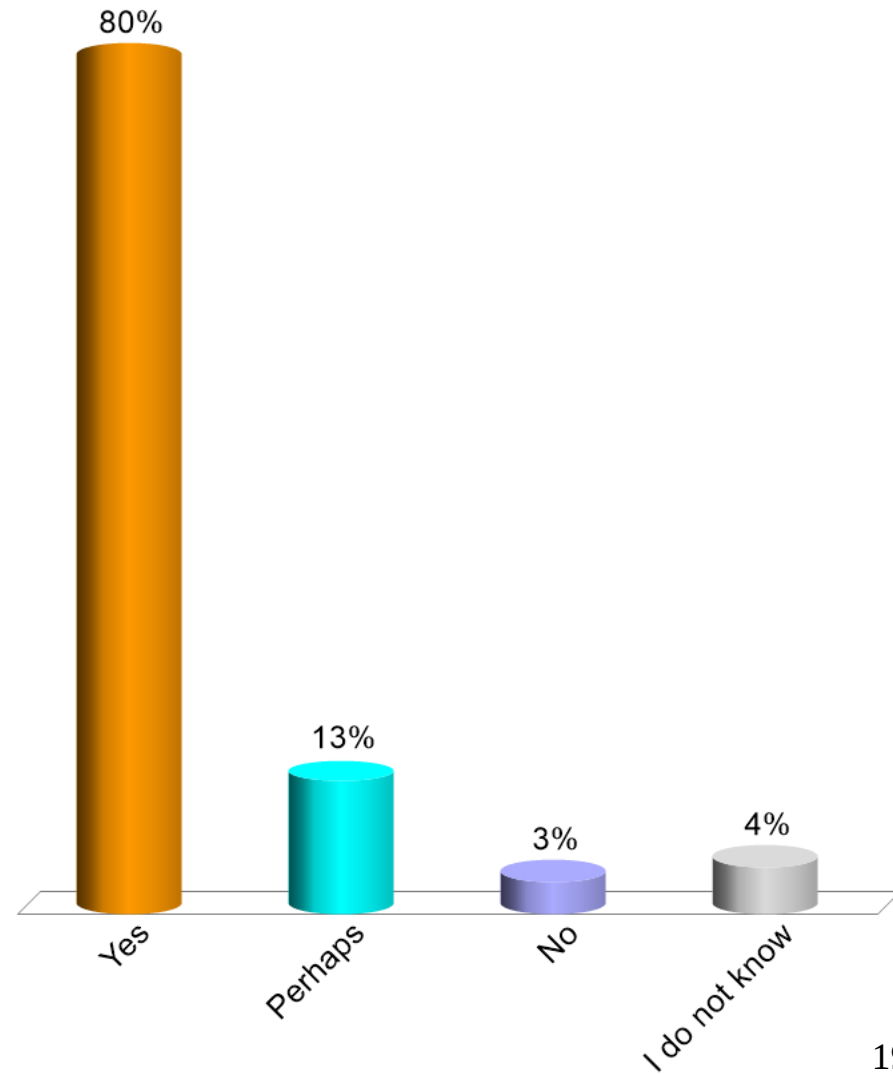
Do you expect to have to use machine learning
once you have graduated?

- A. Yes
- B. Perhaps
- C. No
- D. I do not know



Do you think that having taken a machine learning course will be a plus when you seek a job?

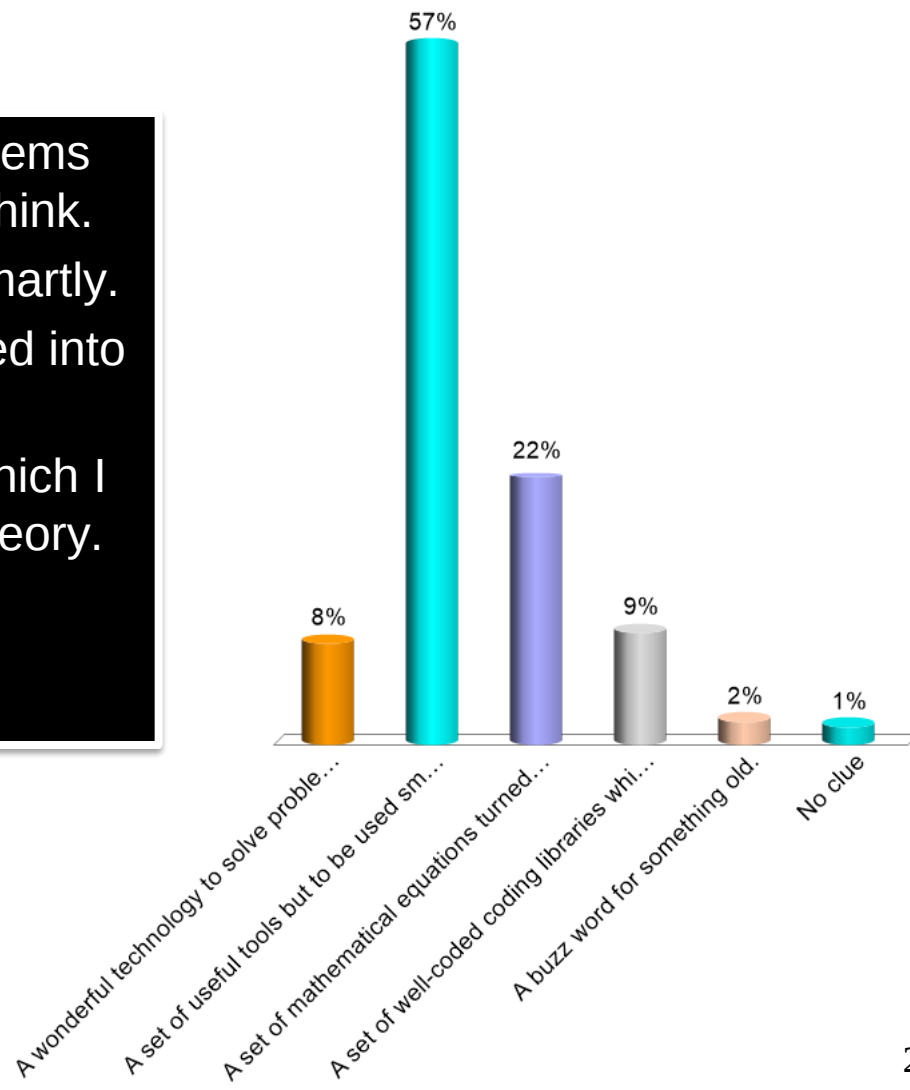
- A. Yes
- B. Perhaps
- C. No
- D. I do not know



What is machine learning to you?

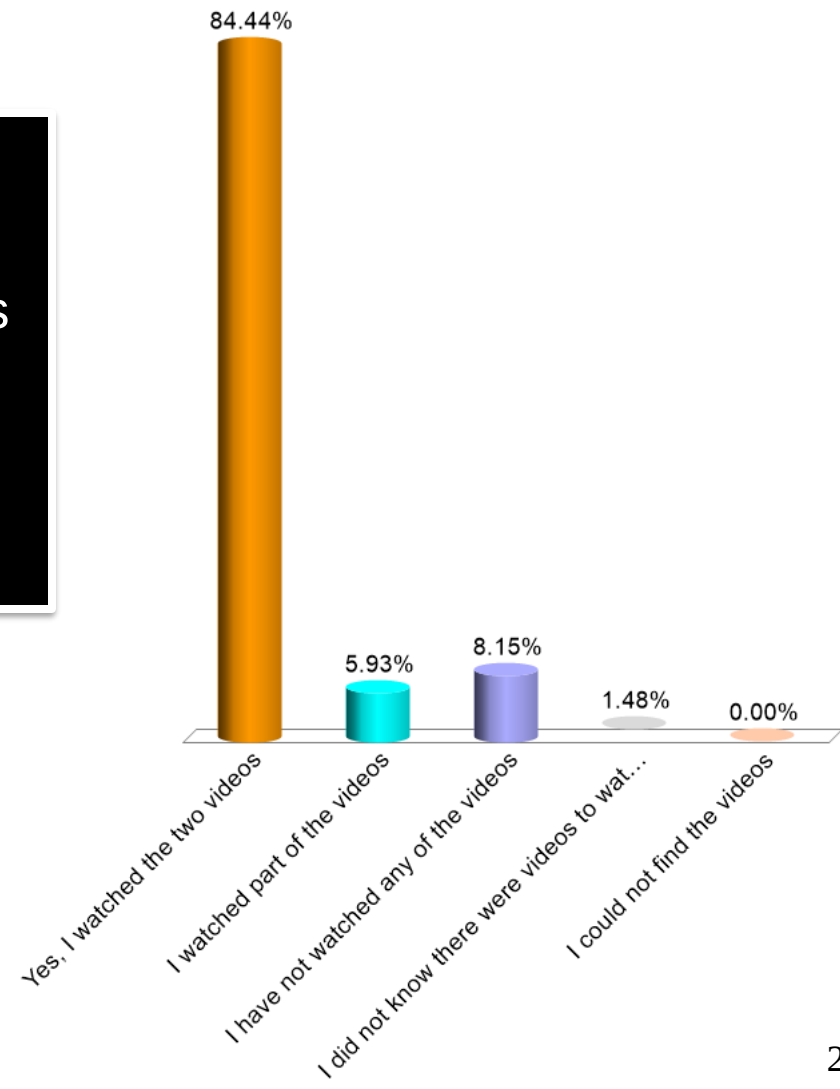
Multiple answers possible

- A. A wonderful technology to solve problems automatically without me needing to think.
- B. A set of useful tools but to be used smartly.
- C. A set of mathematical equations turned into computer code.
- D. A set of well-coded coding libraries which I can use with little knowledge of the theory.
- E. A buzz word for something old.
- F. No clue



Have you watched the videos for this class
prior to joining the class?

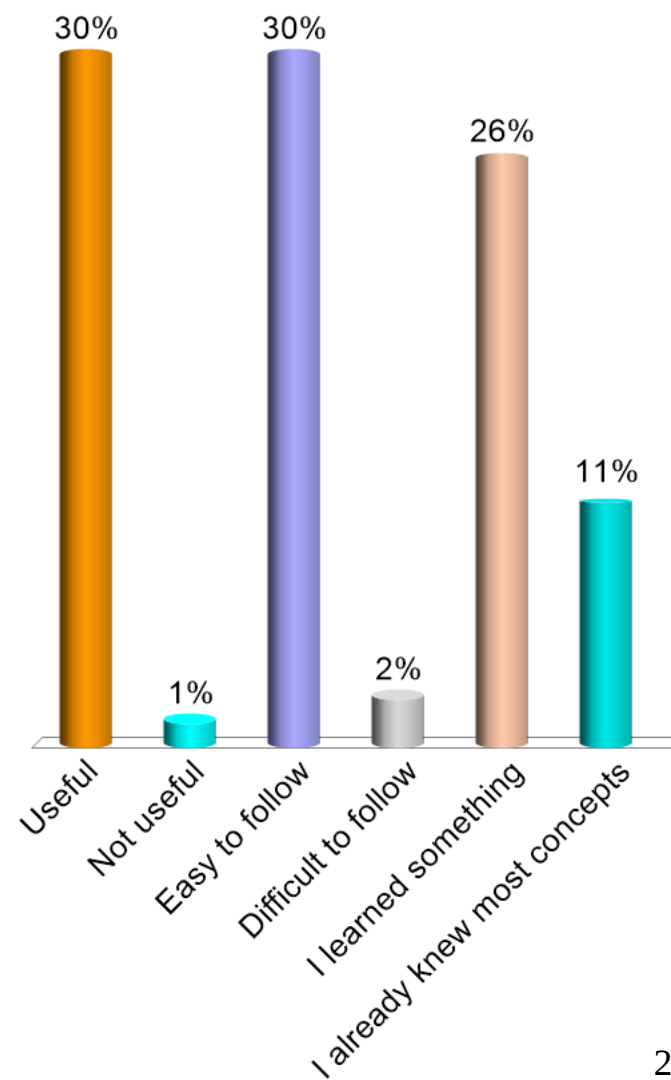
- A. Yes, I watched the two videos
- B. I watched part of the videos
- C. I have not watched any of the videos
- D. I did not know there were videos to watch
- E. I could not find the videos



Qualify the content of the videos

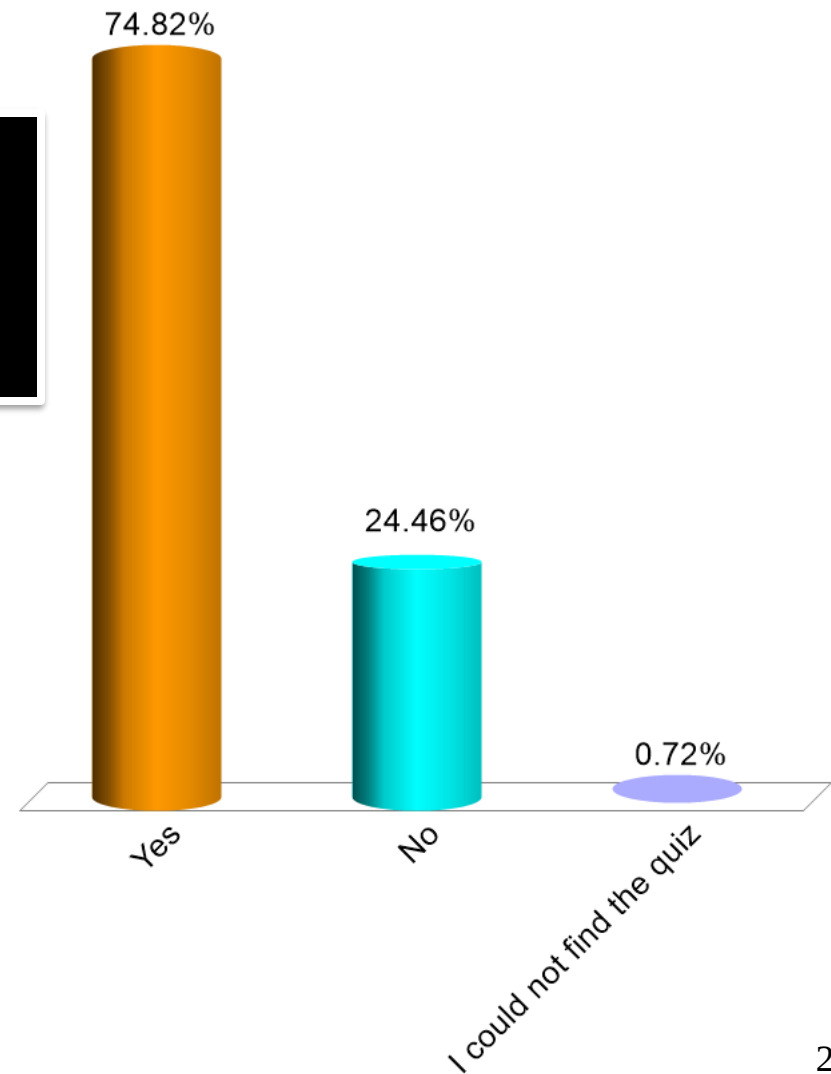
Multiple answers possible

- A. Useful
- B. Not useful
- C. Easy to follow
- D. Difficult to follow
- E. I learned something
- F. I already knew most concepts



Have you taken the quiz for this class
prior to joining the class?

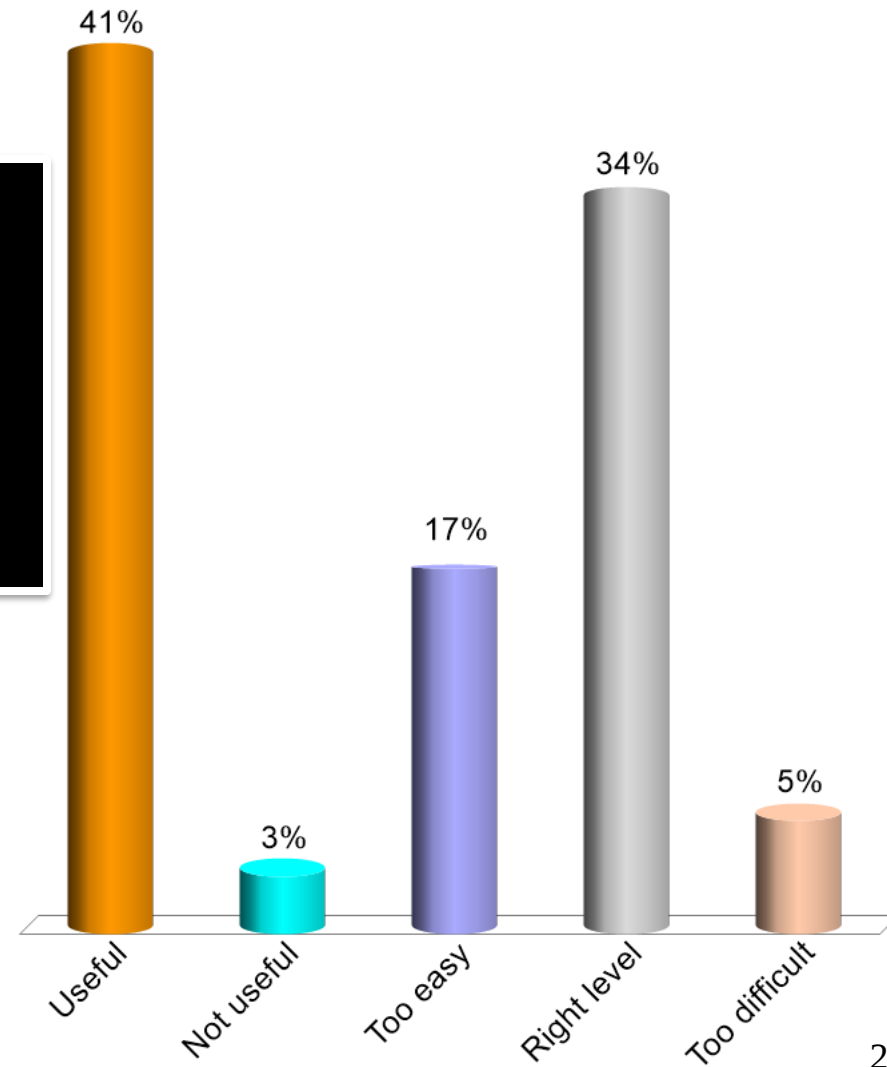
- A. Yes
- B. No
- C. I could not find the quiz



Qualify the content of the quiz

Multiple answers possible

- A. Useful
- B. Not useful
- C. Too easy
- D. Right level
- E. Too difficult

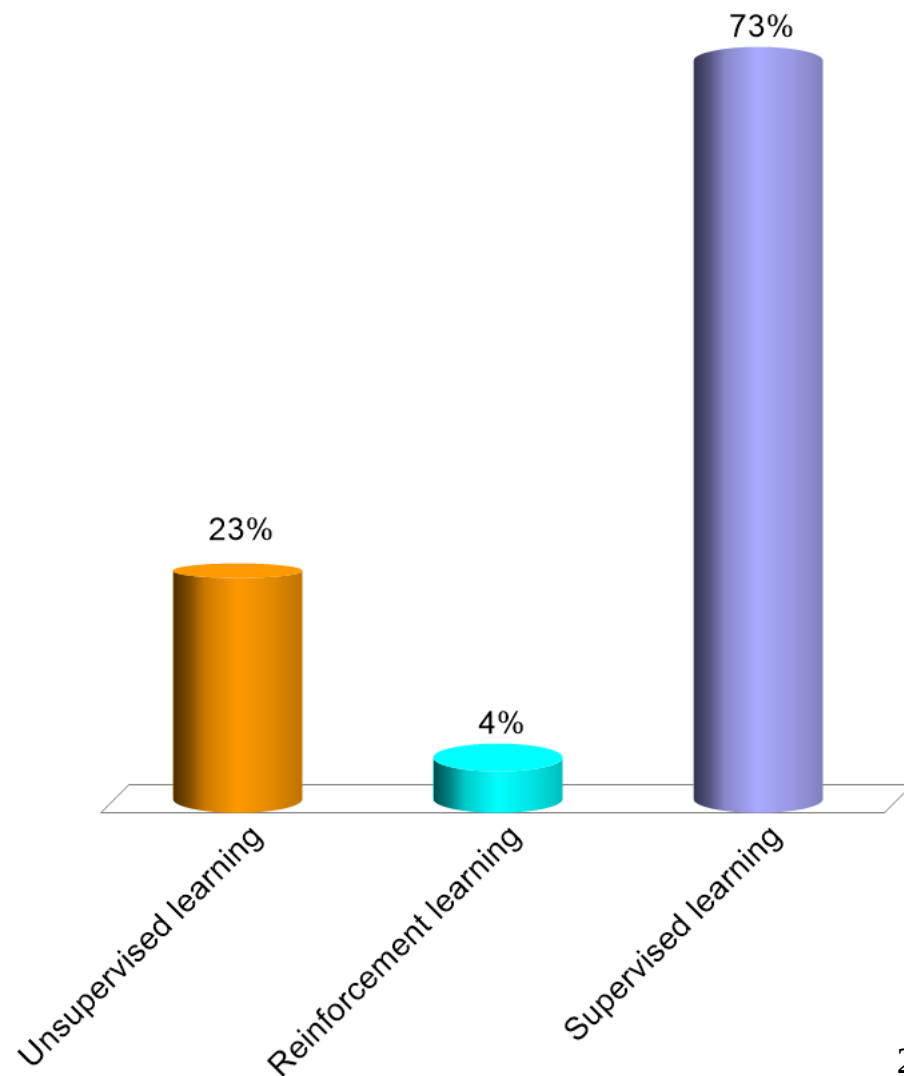


Applied Machine Learning Course

Recap of concepts seen in the videos

What type of learning is classification?

- A. Unsupervised learning
- B. Reinforcement learning
- C. Supervised learning



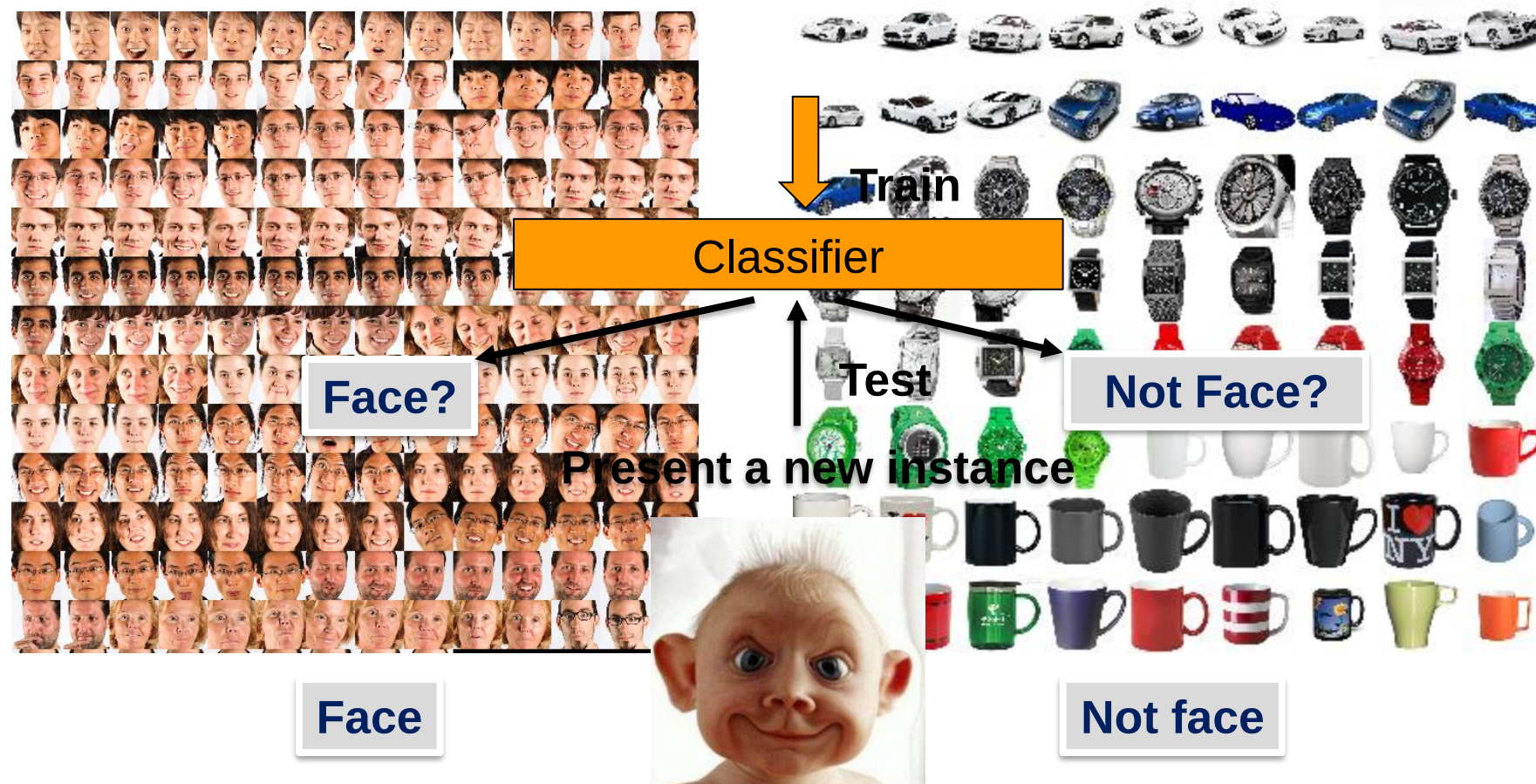
Classification is a type of *supervised learning*

The classifier learns a mapping between

- a set of inputs (e.g. images) and
- a set of desired outputs (class labels).

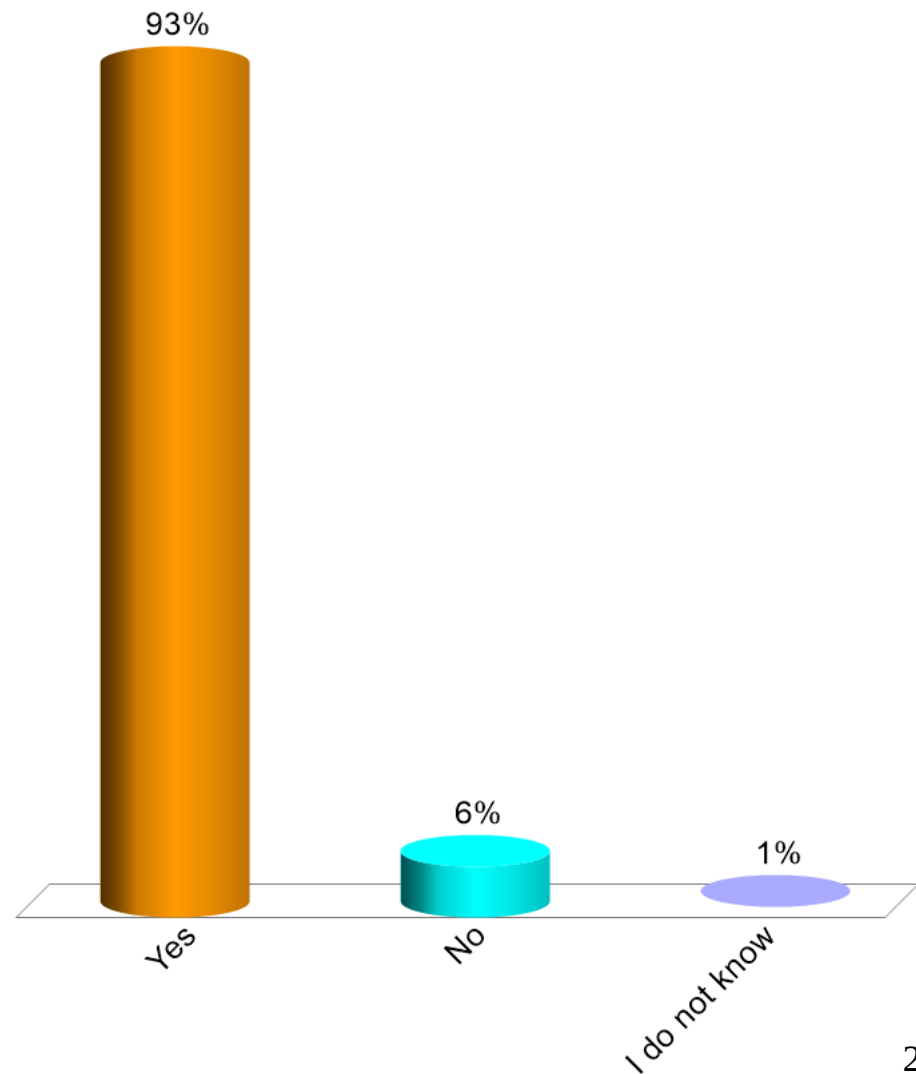
Classification: example

Gather instances of the two classes



Can classification of face vs. no face
suffer from unbalanced datasets?

- A. Yes
- B. No
- C. I do not know

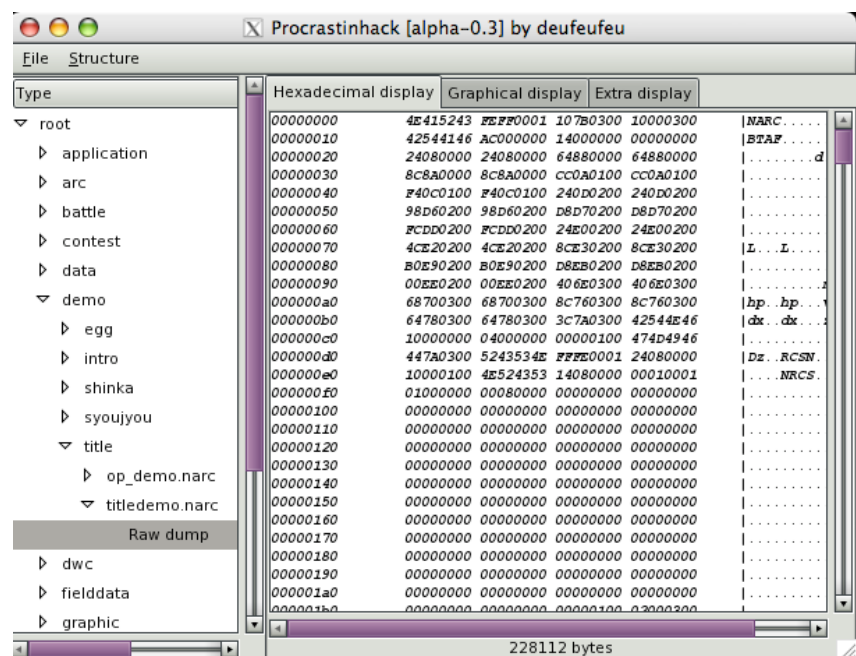
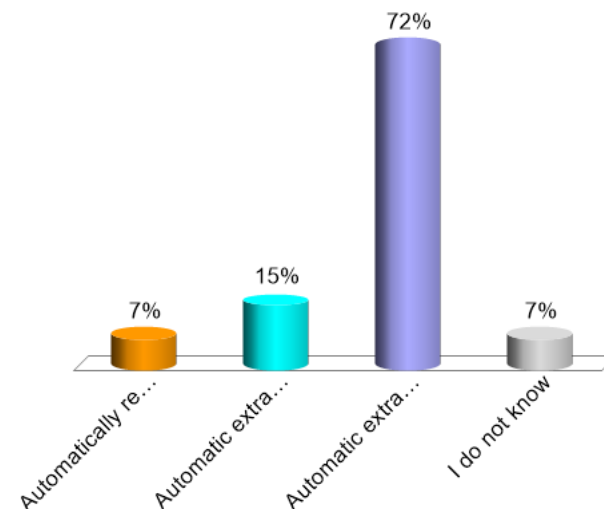


Datasets composed of samples of
faces versus non faces classes
can be severely unbalanced

There are more examples of images of objects
that are not faces than of images of faces.

What does *structure discovery* refer to in machine learning?

- A. Automatically reordering of data in a spreadsheet
- B. Automatic extraction of outliers in a large set of datapoints
- C. Automatic extractions of regularities in across datapoints?
- D. I do not know



Structure discovery: Clustering

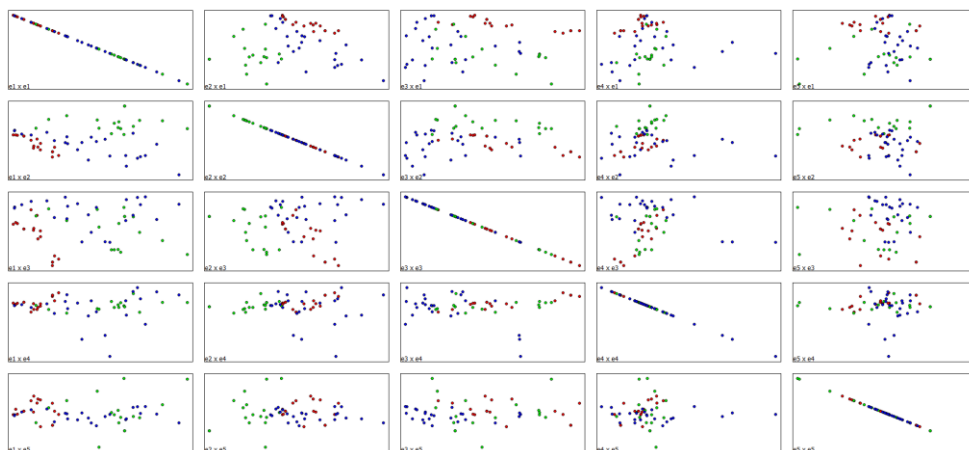


Raw data

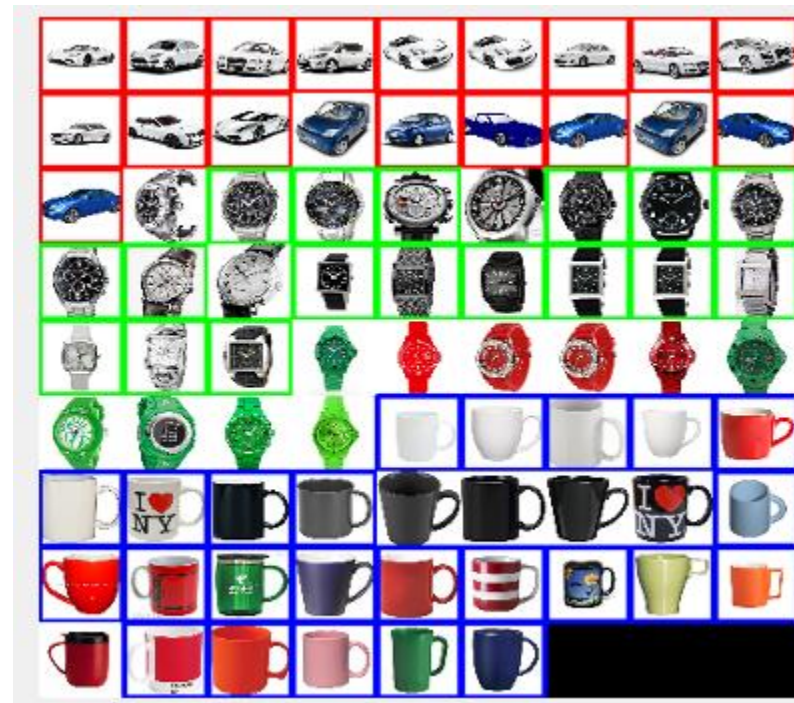


True labels

Structure discovery: Clustering

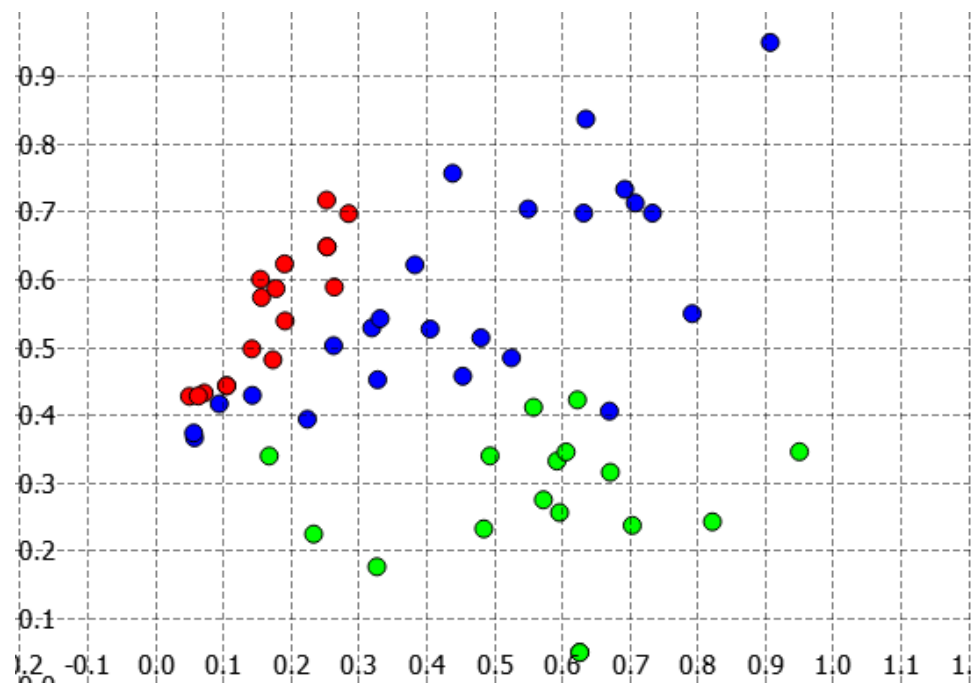


Distribution of data in pixel space

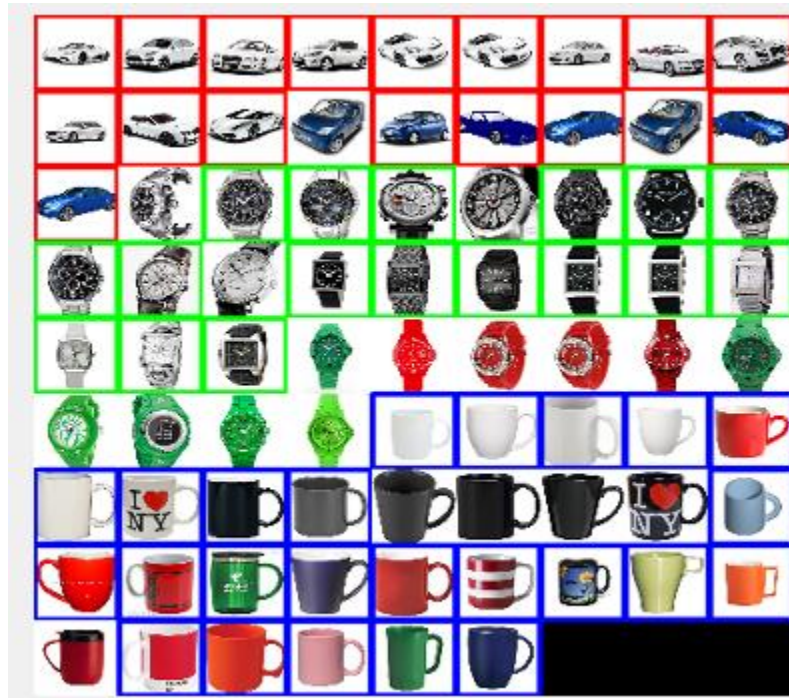


True labels

Structure discovery: Clustering

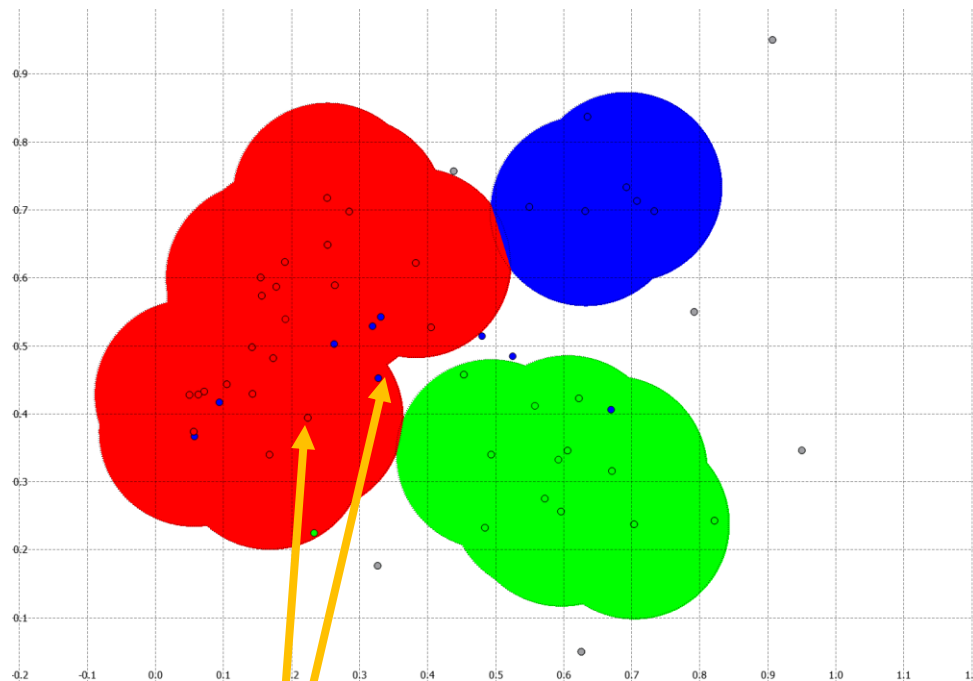


Distribution in low dimensional space



True labels

Structure discovery: Clustering



Automatic grouping

Lack of ground truth and lack of clear separation can lead to erroneous grouping

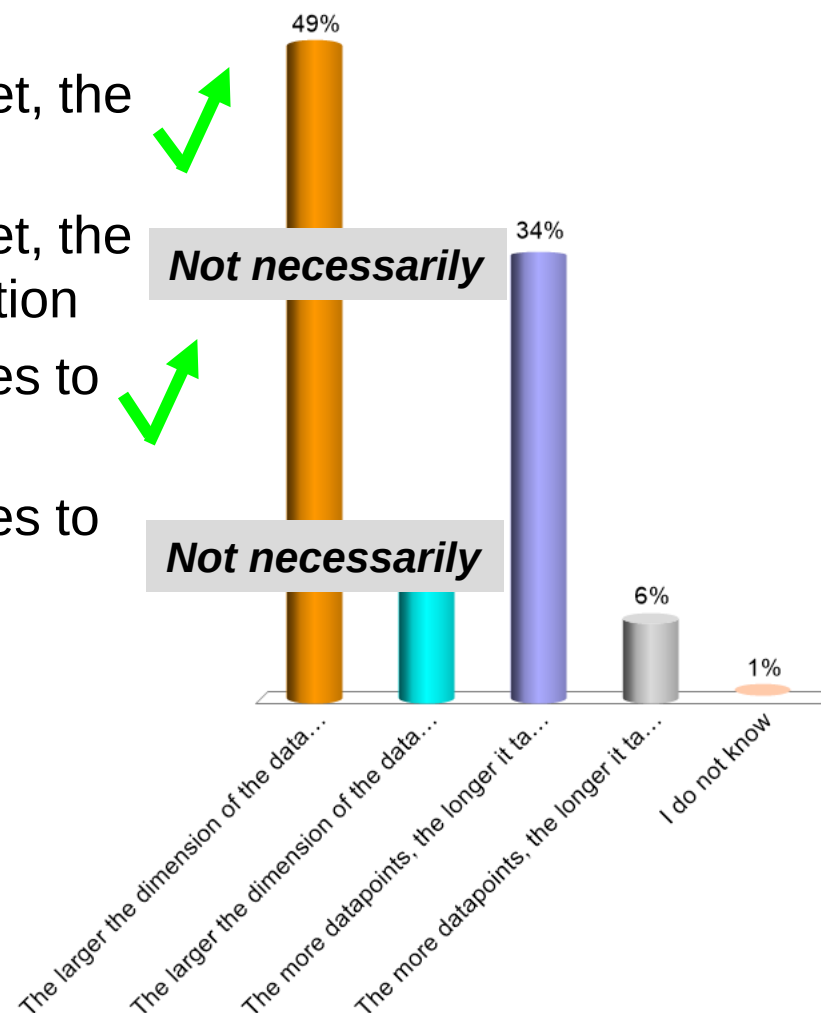


True labels

What is the curse of dimensionality?

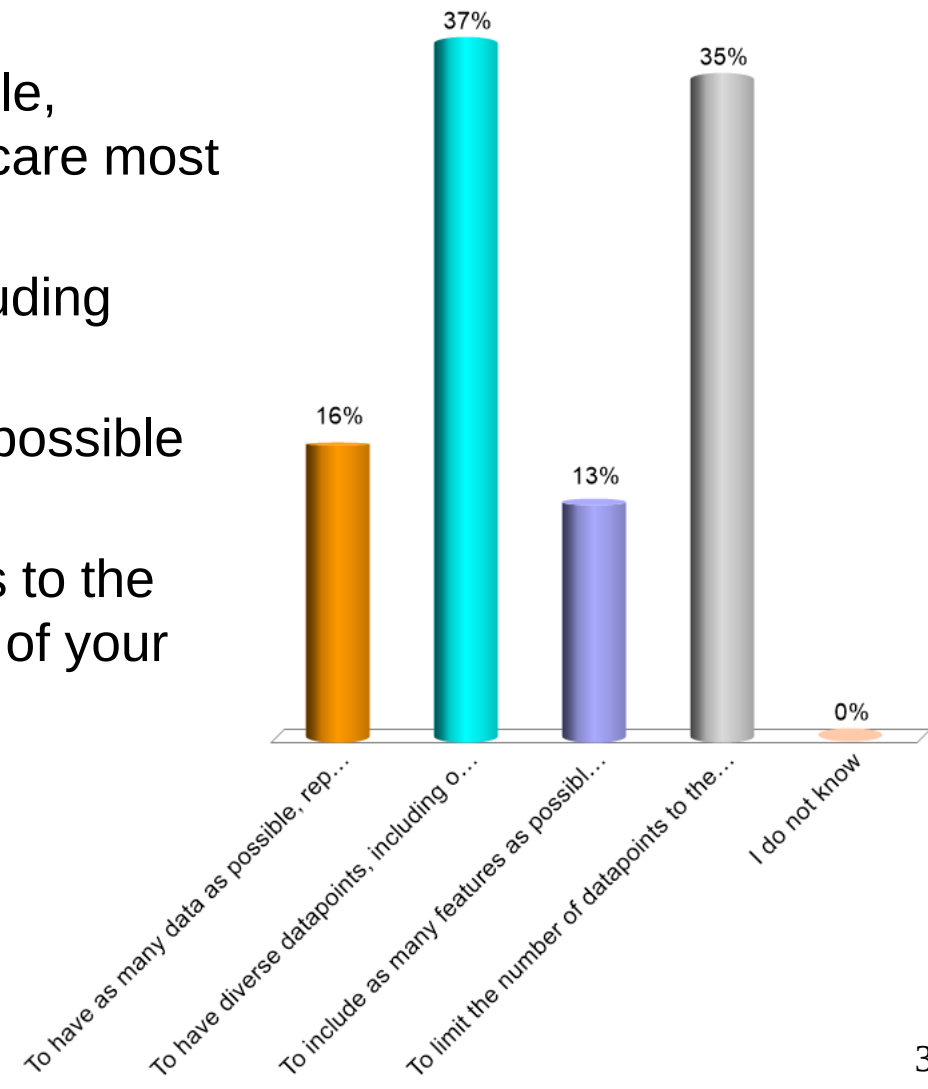
- A. The larger the dimension of the dataset, the longer it takes to compute the solution
- B. The larger the dimension of the dataset, the more difficult it is to find a correct solution
- C. The more datapoints, the longer it takes to compute the solution.
- D. The more datapoints, the longer it takes to gather the dataset.
- E. I do not know

2 correct answers



When you select your dataset,
what should you pay attention to?

- A. To have as many data as possible, representative of the group you care most about.
- B. To have diverse datapoints, including outliers.
- C. To include as many features as possible about your dataset.
- D. To limit the number of datapoints to the memory and maximal resources of your system.
- E. I do not know



Multiple answers possible

When you select your dataset, what should you pay attention to?

- A. To have as many data as possible, representative of the group you care most about.
- B. To have diverse datapoints, including outliers.
- C. To include as many features as possible about your dataset.
- D. To limit the number of datapoints to the memory and maximal resources of your system.
- E. I do not know

A. Yes, have balanced classes!
To distinguish this group from the rest (face vs. no face).

B. Yes, but take the proportion of outliers into account when computing performance

C. Not if some features are really known to be irrelevant. The more features, the more data required

D. Yes, if you care about real-world applications on embedded platforms. No, if you can increase your computer resources.

TODO LIST BEFORE NEXT CLASS

❖ Go to moodle:

- 1: Watch the videos of this lecture (*40 minutes*)
- 2: Do the quiz (*2-3 min*)
- 3: Watch the video of the next lecture on PCA (*40 minutes*)
- 4: Do the quiz (*2-3 min*)

❖ Install software:

- Turning point
- Discord
- Mldemos