

TEACH THE MACHINE

TRAINING DATA AND CLASSIFICATION



6-10 years



25-35 min



6 - 30 participants

Card sorting and role-play activity

OVERVIEW

Participants become teachers for a pretend AI system. Using picture cards, they provide examples that help the “machine” learn how to sort objects or animals into categories. They then test the machine with new examples and discover that its results depend on the quality, variety and accuracy of the information used to teach it.

LEARNING OBJECTIVES

Participants will:

- understand that many AI systems learn from examples
- recognize that examples used to teach an AI are called training data
- identify observable features that can help distinguish between categories
- understand that insufficient or unclear examples can lead to mistakes
- recognize the important role humans play in selecting and labelling training data



MATERIALS

- Approximately 20-30 picture cards showing familiar animals
- Two pieces of paper or signs for category labels
- Two hoops, boxes or marked areas for sorting the cards
- Adhesive tape or reusable adhesive
- Optional board or large sheet of paper
- Optional blank cards and coloured pencils

Suggested animal cards include:

- | | |
|-------------|-------------|
| • cat | • fish |
| • dog | • dolphin |
| • elephant | • whale |
| • monkey | • frog |
| • chicken | • crocodile |
| • eagle | • turtle |
| • bat | • penguin |
| • butterfly | • duck |

If animal cards are unavailable, pictures of food, vehicles, clothing or everyday objects can be used instead.

TEACH THE MACHINE

TRAINING DATA AND CLASSIFICATION



PREPARATION

Prepare two marked sorting areas.

Begin with two categories that younger children can understand, such as:

- Mostly lives on land
- Mostly lives in water

Select six clear examples for the first training round:

- three animals that mostly live on land
- three animals that mostly live in water

Keep the remaining cards hidden for the testing round. Include some examples that are easy to classify and some that are more difficult, such as a frog, crocodile, turtle, penguin or duck.

Choose one participant or facilitator to play the “machine.”

INTRODUCTION

Ask the participants:

- How do we know whether something is a cat, a dog or a bird?
- What features do we notice?
- How might a computer learn to recognize an animal in a picture?
- Can we explain everything using only one example?

Explain:

Some AI systems learn by examining many examples.

People provide pictures or other information and tell the system what each example represents. The AI looks for patterns that may help it classify new examples later.

Introduce the pretend machine:

Today, we are going to teach a machine how to sort animals.

The machine does not begin with all our knowledge.

It can only learn from the examples and labels we give it.



TEACH THE MACHINE

TRAINING DATA AND CLASSIFICATION



INSTRUCTIONS

1) INTRODUCE THE CATEGORIES

Show the two category areas:

- mostly lives on land
- mostly lives in water

Check that participants understand the categories. Explain that the categories refer to where the animal spends most of its life, not whether it can occasionally enter water or move on land.

2) PROVIDE THE FIRST TRAINING EXAMPLES

Show the first six animal cards one at a time.

Ask the group to identify each animal and decide where it belongs. Place each card in the corresponding category area.

For example:

- cat → mostly lives on land
- elephant → mostly lives on land
- monkey → mostly lives on land
- fish → mostly lives in water
- dolphin → mostly lives in water
- whale → mostly lives in water

Explain that these examples form the machine's training data. The category assigned to each card is its label.

3) LOOK FOR PATTERNS

Ask participants to describe what the training examples might teach the machine. Possible observations include:

- Water animals may have fins or flippers.
- Land animals may have legs.
- Water animals may have smooth bodies.
- Land animals may have fur.

Write or repeat the suggested features.

Do not correct every imperfect rule immediately. The testing round should reveal that some simple rules do not always work.

4) TEST THE MACHINE

Show a new card to the participant playing the machine. The machine must classify it using only the patterns identified during training.

Begin with relatively clear examples, such as a dog or shark. Then introduce more difficult animals:

- A penguin has legs but spends a great deal of time in water.
- A frog lives both on land and in water.
- A crocodile has legs but spends much of its time in water.
- A turtle may live on land, in freshwater or in the sea.
- A duck walks on land, swims and flies.

After each decision, ask the group:

- Does everyone agree?
- Which feature did the machine use?
- Did the training examples prepare it for this animal?

TEACH THE MACHINE

TRAINING DATA AND CLASSIFICATION



INSTRUCTIONS

5) IMPROVE THE TRAINING DATA

If the machine struggles, ask participants what additional examples would help it.

Add new training cards representing animals with different characteristics. Participants may decide that the original categories also need clearer definitions or that a third category is necessary:

- mostly lives on land
- mostly lives in water
- lives both on land and in water

Explain that improving the examples and labels is part of improving the system.

6) TEST THE MACHINE AGAIN

Repeat the testing round using the expanded training set. Discuss whether the additional examples make classification easier or more accurate.

Invite a new participant to play the machine if time allows.

DISCUSSION AND REFLECTION

Ask the participants:

- How did the machine learn what to do?
- What information did we give it?
- Which animals were easy to classify?
- Which animals were difficult to classify, and why?
- Did one example teach the machine everything it needed to know?
- What happened when the training examples were too similar?
- Who decided the categories and labels?
- Could two people classify the same animal differently?
- What could happen if a training card were given the wrong label?
- How could we help the machine improve?

Conclude with the key idea: An AI system learns from the information and examples people provide. If its training data are limited, unclear or incorrect, the system may also make mistakes.

TEACH THE MACHINE

TRAINING DATA AND CLASSIFICATION



CONNECTION TO ARTIFICIAL INTELLIGENCE

Many AI systems are trained to recognize patterns in data. For example, an image-recognition system may examine thousands or millions of labelled images to learn which visual features are commonly associated with particular animals, plants or objects.

When the system receives a new image, it compares its features with patterns found in the training data and makes a prediction. It does not recognize the animal through life experience in the same way a person does.

The quality of the result depends partly on:

- how many examples the system received
- how varied those examples were
- whether the examples were labelled correctly
- whether the categories were clearly defined
- whether the new example is similar to anything in the training data

AI can therefore be very useful for tasks such as identifying wildlife, analysing scientific images or sorting information. However, its predictions still need to be checked, particularly when an example is unusual or the consequences of a mistake are important.

EASIER VARIATION

For participants aged six or seven, use two visually distinct categories, such as:

- animals with wings
- animals without wings

Alternatively, sort objects by colour, shape or size.

Use only eight to ten cards and avoid ambiguous examples during the first round. After the children understand the activity, introduce one surprising example, such as a bat in a set of animals with wings. The facilitator can play the machine throughout the activity.

TEACH THE MACHINE

TRAINING DATA AND CLASSIFICATION



MORE CHALLENGING VARIATION

For older participants, divide the group into two teams and give each team a different training set for the same categories.

For example:

- Team A receives several pictures of large dogs.
- Team B receives pictures of dogs of different sizes, colours and breeds.

Test both machines using unfamiliar dog pictures.

Compare their performance and ask why one training set may produce better results.

Participants can then design their own classification system by:

1. selecting two or three categories
2. choosing training examples
3. labelling each example
4. listing the features the machine might learn
5. testing another group's "machine" with new cards

FACILITATOR NOTES

- Keep the robot's misunderstandings humorous but reasonable.
- The aim is to demonstrate ambiguity, not frustrate the group.
- Do not blindfold participants. The robot needs to see the activity area to move safely.
- Use walking only; running should not be permitted.
- Make sure every participant has an opportunity to suggest an instruction.
- Accept different successful solutions. There may be several ways to complete the same task.
- For multilingual groups, demonstrate important commands physically before beginning.
- Emphasize that unexpected results are useful: they help us identify what needs to be clarified or corrected.

TEACH THE MACHINE

TRAINING DATA AND CLASSIFICATION



KEY WORDS

ARTIFICIAL INTELLIGENCE

Technology that can perform tasks involving abilities such as recognizing patterns, making predictions or generating content.

TRAINING DATA

Examples or information used to teach an AI system.

LABEL

The name or category assigned to a training example.

FEATURE

An observable characteristic, such as shape, colour, texture or number of legs.

CLASSIFICATION

Placing something into a particular category.

PREDICTION

A result produced by an AI system based on patterns it has learned.

