



AI Detectives: How Do Machines Learn?

AI exploration phase

P2 - Building Hypothesis on AI

AI and Exploration Topics. What are we exploring?

Everyday presence of AI: where AI shows up in daily life.

Automation vs AI: why not every automatic system is AI.

Supervised learning: learning to classify from labelled examples.

Model: an internal representation built from data.

Algorithmic bias: how training data can skew results.

Materials Needed and to Create. What must be prepared before the activity?

1 set of cards with everyday AI / non-AI application images, per group (cards or printouts of everyday AI vs non-AI examples)

1 set of concave vs convex training cards (shapes), per group

2 posters or boxes labelled Pile 1 / Pile 2, 1 set per class

Recording sheets, 1 per group

Computer, 1 per group

Teachable Machine (online tool, used in Activity 3)



Exploration Steps and Description. What do participants actually do?

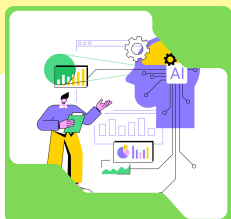
In this activity, young people investigate where AI appears in daily life, compare rule-based programming with machine learning, and train a real model, discovering that an AI learns patterns from data rather than understanding like a person.

Activity 1 - AI Detectives (30 min)

After an opening question (where do we find AI in daily life?), each group sorts everyday-situation cards (automatic door, social feed, traffic light, robot vacuum, face filter, voice assistant) into three columns: **surely AI, sometimes AI, surely not AI**. A whole-group discussion compares the answers and builds a first working definition: there is AI when a system uses data to recognise patterns and make predictions or decisions.

Activity 2 - Rules vs examples (90 min)

First the facilitator classifies shapes as a rigid program applying an explicit rule. Then each group plays an AI that learns the same task only from examples, with a training phase and a test phase where it predicts new shapes, including tricky weirdos it must force into a pile. The closing debate names **input, model, output, training data, test data and prediction**, and reveals the rule was **concave vs convex**.



AI Detectives: How Do Machines Learn?

Exploration Steps and Description. What do participants actually do?

Activity 3 - Teachable Machine (30 min)

Each group trains a model on a few example cards, then tests it on new cards and on changes of light, angle or distance, recording predictions in a table. They see the machine classifies without knowing the rule, and that **the result depends heavily on which examples were used**.

Activity 4 - What did we learn? (30 min)

Each group writes one key idea, one piece of evidence and one open question, and the pooling links the day to the year project: **a robot for an SDG challenge first needs to recognise, classify and decide**.

Expected Output. What do participants create, produce or gain from the exploration?

Young people can tell automation from AI, classify from examples versus fixed rules, train and break a simple model, see that results depend on training data, and use words like model, prediction, error and bias. These are their first tools for the final project.

The Hook and the Playfulness. What makes this fun/exciting for young people?

Playing detectives to crack a secret rule, training a machine and catching it making confident mistakes, and seeing that the problem is not only technical but also social and ethical.

Success indicators. How do we know it worked? What shows participants learned?

- They describe AI as a system trained on data, not just an automatic or intelligent technology.
- They state the difference between an explicit rule and a model trained on examples.
- They anticipate errors from the data and drop the idea of an infallible magic box.
- They raise questions about responsibility, data quality and fairness.



Appendix - AI Detectives: How Do Machines Learn?

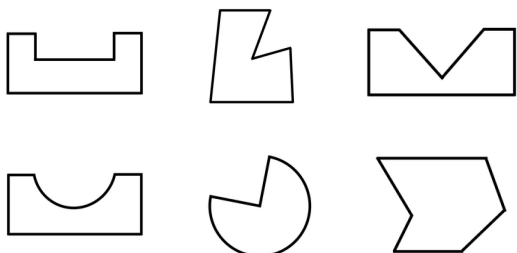
Appendix 1

Concave vs convex shape cards (and the "weirdos")

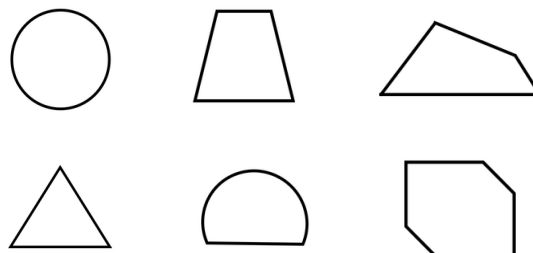
Content of the appendix

Sorting cards for Activity 2 (rules vs examples). Print one set per group

Pile 1: Concave
(interior angles greater than 180 degrees)



Pile 2: Convex

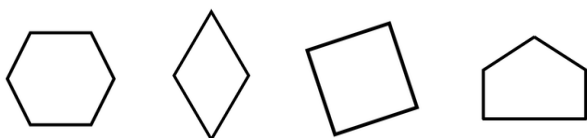


Test shapes (marked with a question mark): to be sorted by the group

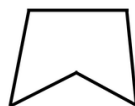
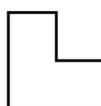
? Pile 1: Concave



Pile 2: Convex



? Special concave cases (the weirdos):



May look more regular or more closed, but is still concave (interior angles greater than 180 degrees).

Is rotated; the other concave shapes were shown the other way up.

How the cards are used

First decide which criterion to classify by and split the cards into two piles (Pile 1 becomes Class 1 in Teachable Machine, Pile 2 becomes Class 2). Then show the cards one by one and place them in the matching pile. Three criteria are possible: colour (yellow / blue), fill (pattern / solid) and shape (concave / convex). The example shown classifies by shape only; for colour and fill the classification is obvious. The shapes labelled with a question mark (test data) are the ones young people classify on their own after seeing how piles 1 and 2 were built.