

Fostering Artificial Intelligence at Schools

Guideline for teachers in formal education



Table of Contents

Introduction	4
Why do we need to teach about artificial intelligence?	4
How to use this document	5
Background	6
What is AI?	6
What is data?	9
What is bias and algorithmic bias?	9
What is an algorithm?	11
What are the different subfields of AI?	11
Knowledge representation and reasoning (expert systems)	11
Automated planning	12
Machine learning	12
Reinforcement learning	12
Machine perception	13
Natural language processing	13
Robotics	13
Machine Learning	14
Supervised learning	15
Unsupervised learning	16
Reinforcement Learning	16
Classification	17
Prepare data	17
Train model	18
Test model	18
Deploy model	18
Important note on the choice of training data	19
Introduction to the LearningML tool	20
Demo Learning ML tool	20
Introduction to supervised machine learning	23
Lessons plans related to algorithmic bias	32
Activity one: Generic class about algorithmic bias	32
Activity two: Social sciences	39
Specific lessons for classroom subjects	43
Specific LearningML lessons	43
Mathematics	43
Religious buildings	52
Additional tools	55
Music	55
Glossary	56
Related bibliography	58

Imprint

2023: The FAIaS Project

Editors: Universidad Rey Juan Carlos (ES), Vrije Universiteit Brussel (BE), CollectiveUP (BE), Teatro Circo Braga (PT).

Authors: Gregorio Robles, Meritxell Díaz, Antonio J. Romero (Universidad Rey Juan Carlos), Marjon Blondeel (Vrije Universiteit Brussel), Liliana Carrillo, Chrysanthi Katrini, Loredana Bucseneanu (CollectiveUP), Joana Miranda and Sara Borges (Teatro Circo Braga). Revised by Brian Chen (Massachusetts Institute of Technology).

Layout by: Jurate Laugalyte (CollectiveUP), Carlos Nivia Otero (CollectiveUP).

We suggest citing this report as follows: Robles, Díaz, Romero, Blondel, et al. (2023). Fostering Artificial Intelligence at Schools: Guideline for teachers in formal education <http://www.fosteringai.net>

This publication has been written within the Erasmus + project Fostering AI at Schools, which aims to support educators to integrate and use Artificial Intelligence concepts in their lessons by creating and sharing resources about AI that can be used in the classroom.

CollectiveUP thanks the Digital Belgium Skills Fund (DBSF) for co-funding the AI4InclusiveEducation project (www.ai4inclusiveeducation.be) with a grant number DBSF2023-14. DBSF co-funded and partially supported the achievement of the Erasmus+ FAIaS project by creating an impact in Belgium.

This publication (in English, Spanish, Dutch, and Portuguese) and further publications and tools produced in the project can be downloaded free of charge at: www.fosteringai.net

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>



Disclaimer

Fostering AI at Schools is a project partnership, co-funded through the European Commission. This publication reflects the views only of the authors, and the European Commission cannot be held responsible for any use which may be made of the information contained therein.

Fostering AI at Schools - Erasmus+ Project, Action type "Cooperation partnerships in school Education", Grant number: 2020-1-ES01-KA201-083047

Image credits

Cover & back cover - Own design, using images created [by Freepik](#) and [by kjpargeter](#)
Pages 4, 6, 20, 43 - Header image created [by kjpargeter](#)



Introduction

Why do we need to teach about artificial intelligence?

We often think of Artificial Intelligence (AI) as some mysterious black box. Nevertheless, it is all around us and we use it everyday: spam filters, online chatbots, search engines, video games anticipating our next moves, navigation apps, rideshare apps, just to name a few. The underlying processes and concepts are rarely exposed to most people, making it difficult to understand what AI is and how it should be approached. Understanding the fundamental concepts however is becoming very important for people of all ages, including children who are growing up in an environment that integrates more AI products than ever before. Moreover, the children and students of today will enter a workforce increasingly powered by AI. Ignoring AI can be compared to ignoring the internet in the 1990s. By bringing AI into the classroom, students can understand the potential impact of AI in their daily life. It also allows them to understand the role of AI and the ethical considerations.

Knowing the aspects and impact of artificial intelligence is also of great importance for teachers. In the first place, because it is a reality that students have to get used to. This should also be seen as an opportunity, since they can include AI techniques in their classes, just as they are used in professional and academic environments. Secondly, because they have to be aware of the potential as well as the pernicious use of AI. As those responsible for developing the skills of their students, they need to know how AI can be used for that purpose. They must be the ones who get students to acquire capabilities that are innate to humans and that cannot be replaced with the use of AI. Among them are, for example, critical thinking or the identification of bias. That is why teachers have a great challenge before them and we hope that these materials and the set of activities that are made available to them will help them in their undertaking.

How to use this document

In this document, you will find a basic starting point to understand what AI (and in particular Machine Learning) is and how to apply it to your lessons, offering generic and specific lesson plans.

In particular, the outline of this document is as follows.

In the background section you will find:

- An introduction to AI, data, algorithms and (algorithmic) bias.
- An overview of the different subfields of AI.
- An introduction to machine learning, with examples and illustrations.
- A more in depth introduction to classification and how AI models are trained to perform classification.

In the next section we introduce the LearningML tool and explain how it can be used in the classroom. As an illustration, in the third section we give an overview of generic and specific lesson plans. The generic lesson plans can be adapted to different subjects. Guidelines on how to adapt the lessons are provided. We then present specific lesson plans that centre around a specific content, usually linked to a particular subject. The specific lesson plans follow the same structure, and have been thought to be adapted to your needs (duration, age, etc.). In particular, a set of activities about the concept of bias are proposed. We have put these activities apart from the lessons as we think that they allow to develop critical thinking of students, which beyond understanding what AI is and how it works or can be applied, is of major importance. You will also find some additional tools that can be used in the classroom.

Finally, the last section contains a glossary with the most relevant AI and machine learning concepts explained.

Background

This chapter will introduce and detail some concepts necessary to be able to understand the terminology used throughout the activities indicated in this document. The structure is as follows:

Firstly,

some key concepts will be explained, such as understanding what AI is, what algorithms are and what (algorithmic) bias is. An overview of the subfields of AI is also given.

Secondly,

a detailed explanation of one of the most important sub-branches of AI, machine learning (ML), is given.

Finally,

we look at classification and use it to explain the basic steps needed to create an ML model: data selection, training and testing.

What is AI?

Artificial intelligence (AI) is a discipline from computer science that tries to replicate and develop human intelligence and its implicit processes through computers. The term was coined in 1956 at the Dartmouth Conference by the renowned computer scientist John McCarthy, although some first mentions go back to shortly after World War II.

Although there is no agreement on a unique definition of AI, there is a consensus on four main approaches: two of them are based on humans (systems that think like them and systems that act like them) and two on rational aspects (systems that think rationally and systems that act rationally).

AI synthesises and automates tasks that are initially intellectual, and can therefore be potentially considered relevant to any field of human intellectual activity. We are talking of a genuinely universal field.

Therefore, AI is the combination of algorithms designed to create machines that have the same capabilities as human beings. Computer scientists Stuart Russell and Peter Norvig distinguish between four main types of AI:

1. Systems that think like humans

They automate activities such as decision-making, problem-solving and learning. One example is artificial neural networks illustrated in Figure 1.

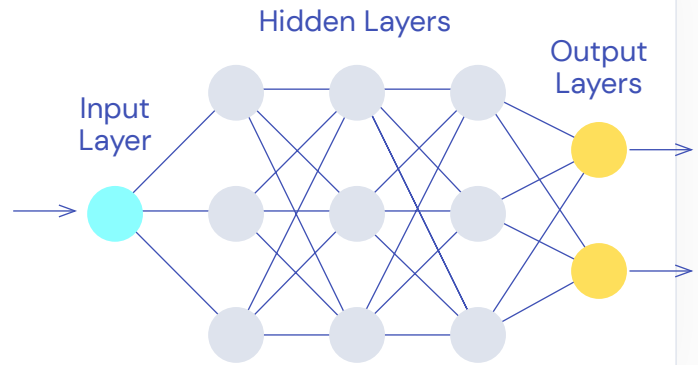


Figure 1. Neuronal Network workflow diagram

2. Systems that act like humans

These are computers that perform tasks in a similar way to people. This is the case of robots.

3. Systems that think rationally

They try to emulate the rational logical thinking of humans, i.e. research is carried out on how to make machines perceive, reason and act accordingly. Expert systems fall into this group, Figure 2 shows a schematic of its components.

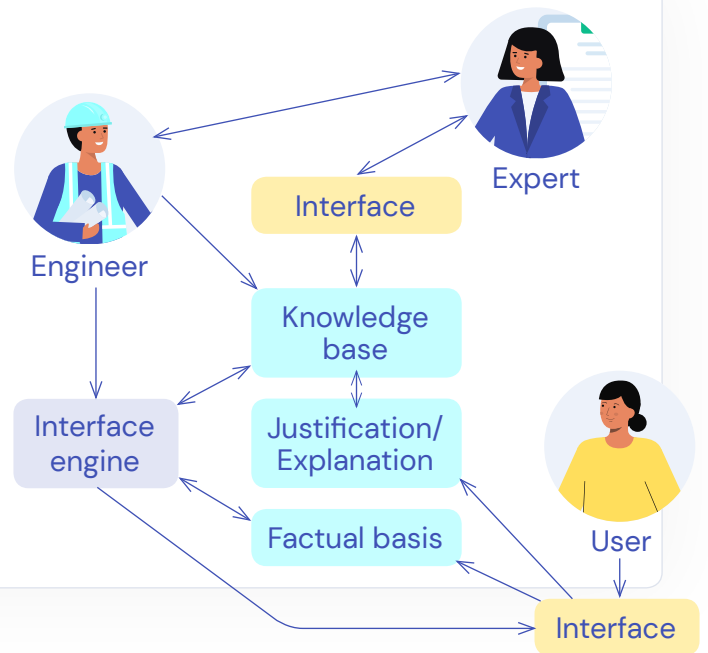
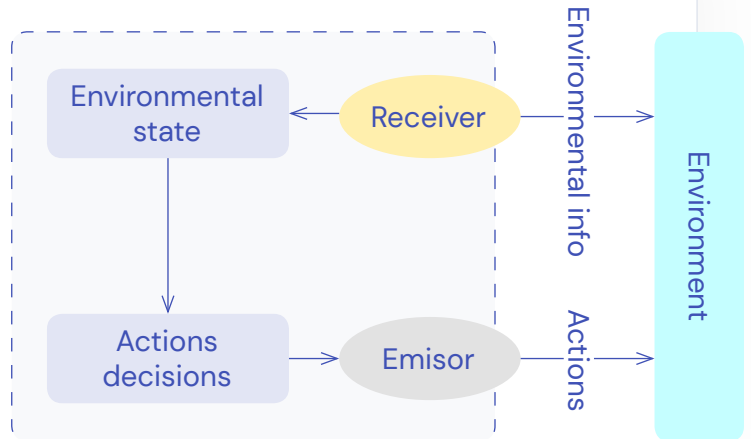


Figure 2. Expert systems workflow diagram

4. Systems that act rationally

Ideally, these are systems that try to rationally mimic human behaviour producing some type of change in the environment. Figure 3 shows a schematic of the components of an intelligent agent.

Figure 3. Diagram of the components of an intelligent agent

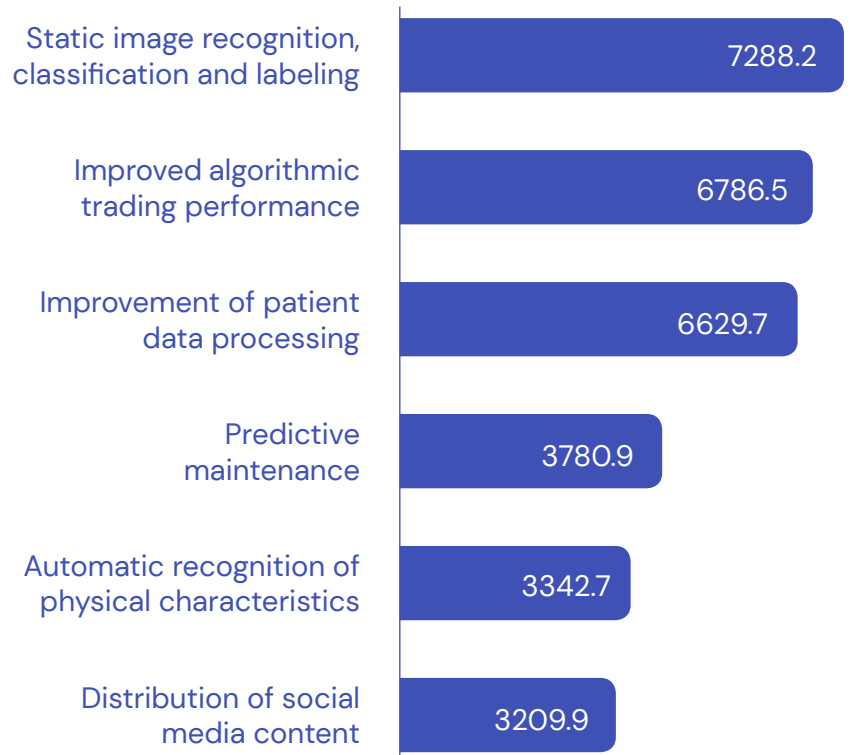


To show the importance of AI, Figure 4 shows some of the applications of AI and the estimated revenue generated.

As we can see, AI is a key element nowadays and it is necessary to be familiar with the concept and its possible applications.

Figure 4. Applications of Artificial Intelligence (Source: Tractica)

Estimated global revenue of AI from 2016 to 2025 (in million U.S. dollars)



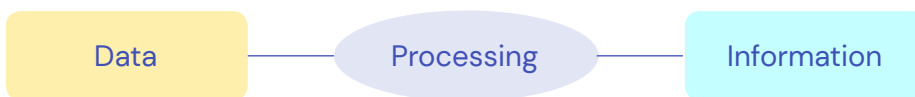
What is data?

AI, and in particular ML, makes heavy use of data. Using the atomic analogy, its smallest element of data is the datum. A datum is a symbolic representation of a quantitative or qualitative attribute or variable. The data describe empirical facts, events and entities. The nature of a datum can be numerical, alphabetic, algorithmic (we will talk about it later), or spatial, among others.

Although nowadays we hear very often about the value of data, by itself it may not contain humanly relevant information. Only when a set of data is examined together in the light of an approach, hypothesis or theory can the information contained in the data be appreciated. That is why properly grouped, structured and interpreted data is considered to be the basis of humanly relevant information that can be used in making decisions, reducing uncertainty or making calculations. This is the reason why many consider data as the petroleum of this century's most relevant machines, the computers.

In Figure 5 we can see, in a simplified way, the need to process data, the ultimate goal being to obtain information in order to gain knowledge about the data.

Figure 5. Diagram of the way to gain knowledge.



We will consider data as the input in our AI activities. Computers will learn from data and will reflect based on them. If we do not use the best data or our data is skewed, we might have bias, which might heavily influence our results, especially if humans are affected. We will talk more on bias in this document, and given its relevance, offer several activities to work on this concept with your students.

What is bias and algorithmic bias?

Every person has biases: these are tendencies, inclinations or prejudices toward or against someone or something. Some of them are positive and helpful. For instance, suppose you are approached by a stranger that has a knife in their hand. In this scenario it is good to be apprehensive. Another example is where you choose to only eat food that is considered healthy.

Most of the time however, biases result in prejudgments that lead to discrimination and rash decisions. A typical example is when a manager interviews people for an open position. The manager immediately likes the first person when he first meets him. During the interview however the candidate seems nervous, and the manager tells him to calm down and to start over again. After that, the interview went great. A couple of hours later, the manager interviews another person but immediately is turned off by some first impression. This person is also nervous but does not receive any words of encouragement or comfort. The interview went very bad. First impressions make a huge difference.

Automated systems are not inherently neutral. Computers must be coded by persons and as a consequence they reflect prejudices and preferences of those who code the AI algorithms behind these systems. Here we list some examples of possible causes of bias in AI.

One very common way bias is introduced in AI systems is because of biased data. For instance, suppose you want to automate the decision process of whether an applicant gets accepted into your university or college. Suppose that in the past you have admitted mostly men. If you use this historical data to train the new system, it is likely this gender bias is replicated.

When using real-world data, like news articles or social media posts, AI systems are likely to reinforce the existing prejudices. Automatic translation systems apps, for instance, tend to be biased against women when translating from languages that use gender neutral pronouns. Typical examples are “she takes care of the children” and “he invests”.

When looking at the task of classification, the way the data is collected is very important. Both undersampling and oversampling may lead to bias. An example of undersampling is when face detection systems are mostly trained on white male faces and as a consequence do not recognize darker female faces as good as lighter male faces. Oversampling on the other hand may lead to overrepresentation of certain groups. For instance, crimes committed in areas frequented by the police will be more present in the data set and the AI model trained on the data will probably reflect this disproportion and these areas will be labelled as more dangerous.

People can also reinforce bias in already deployed AI models. For example, an AI system might already be showing ads for a high-interest credit card to people with a lower level of education. They might click on this ad without realising this will reinforce the already existing bias and they will continue to receive such suggestions.

What is an algorithm?

An algorithm is a set of defined and non-ambiguous, ordered and finite instructions or rules that allow to solve a problem. In our case, computers will use algorithms to process data and perform other tasks such as classification. Not only computers use algorithms; for instance, in the kitchen humans also use them but in the form of recipes. In the case of computers, the ingredients are the data, and the recipe is the algorithm.

What are the different subfields of AI?

AI has evolved tremendously since 1956 when a small group of scientists gathered for the Dartmouth Summer Research Project on Artificial Intelligence which is now considered to be the birth of this field of research. Nowadays, when you hear about AI in the media, it is mostly focused on algorithms that learn from either lots of data or even from experiences. AI however is much more than that and includes several subfields. In this section we will attempt to give an overview of the different subfields. This is a very difficult task; AI is overlapping with other fields of research such as computer science, mathematics, psychology, biology, and many more. Some of the research tracks have also grown into fields of their own such as natural language processing. On top of that, most of the different subfields are also overlapping. Finally, in practice, most applications follow a hybrid approach and use ideas and techniques from several subfields.

Knowledge representation and reasoning (expert systems)

This subfield has been around since the early days of AI. It is concerned with representing information in a way that a computer can reason about it using rules, e.g. logic, without needing to write code that solves the problem. Applications include complex tasks like scheduling or solving sudoku puzzles. As an input, these systems take a description of the problem, and the internal machinery will make sense of the problem and output possible solutions. An important feature of these algorithms is their explainability: there is always a reasoning behind it. Of course, the world is very complex and cannot be entirely written down in rules. These algorithms are mostly suitable in very closed domains. You will never be able to build a good self-driving car with only these types of algorithms. In practice however these techniques are still used in self-driving cars, for instance to include domain knowledge about traffic signs.

Automated planning

Automated planning is another long-standing subfield of AI which is closely related to knowledge representation and reasoning. A typical application is planning (optimal) routes for robots in a warehouse. These algorithms compute routes using a representation of the warehouse as input. Again, these algorithms are transparent, but they can become very inefficient in situations where we have several robots in a relatively small warehouse. An advantage is that these algorithms will still work when you modify the layout of the warehouse, you will not have to change anything.

Machine learning

As mentioned before, the world is too complex to write down in rules. The idea of ML is not to program these rules ourselves but to let the computer learn (some of) these rules. In ML a computer – or algorithm – will learn from data such as numerical data or images. In essence these algorithms try to find patterns in data: by seeing a lot of images containing cats, it will learn to decide whether a new, unseen image contains a cat or not. Computers are really good at finding patterns in this way, but contrary to planning and knowledge presentation, these techniques are not transparent: a rule is being learned but there is no meaning attached to this rule.

Machine learning is often combined with other techniques. An example is where questions about pictures need to be answered. For instance, suppose you want to know how many cats are in a picture. ML will be able to recognize cats in the pictures, but it is very difficult to learn how many cats there are. Hence it is much easier to use ML to detect cats and then use logic and rules to count them.

In the media you often hear about deep learning. This is a part of ML. The main difference is that it needs less intervention from humans, it can work on raw data (e.g, pixels in an image) and it can learn from its own mistakes. The downside is that a huge amount of data is needed, it takes a very long time to learn anything significant and specialised infrastructure is needed.

Reinforcement learning

Reinforcement learning can be seen as a part of ML but it has grown so much that it has become a field of its own. The main difference with ML is that these algorithms learn from interaction and not from data. For instance, our self-driving car from earlier would not really work well using only ML. Suppose it would have to acquire data from other situations in order to learn how to act. This is an infeasible task: it is impossible to distil a representative list of situations. This is where reinforcement learning comes in. These algorithms learn which actions to take by interacting with the world. It is based on the idea of trial-and-error learning and is closely related to the field of psychology. During the learning phase the algorithms explore or exploit what they already know and they receive constant feedback. This is a very easy to understand framework that can be transferred to many applications: the algorithms receive an observation of the

world, take an action and receive feedback based on what the algorithm should learn, e.g. if you do not want to collide with other objects in the world, a negative reward signal should be send back when the car comes to close to other objects. These algorithms have proven to be very successful in environments that can change quickly: they have learned how to adapt quickly to new situations. An advantage is that you do not need data, you just need an environment to interact with. A disadvantage is that learning takes time, and these algorithms lack explainability as they are so complex in their inner workings.

Machine perception

Our self-driving car will not be able to learn anything useful if it cannot perceive the world: it needs ways to look around, to hear, feel and see what is happening. Machine perception is an umbrella term for techniques that perceive the world and try to make sense of it. Examples are computer vision, computer audition, speech recognition and computer touch. Most of the techniques use physics, mathematics, engineering, ML, ...

Natural language processing

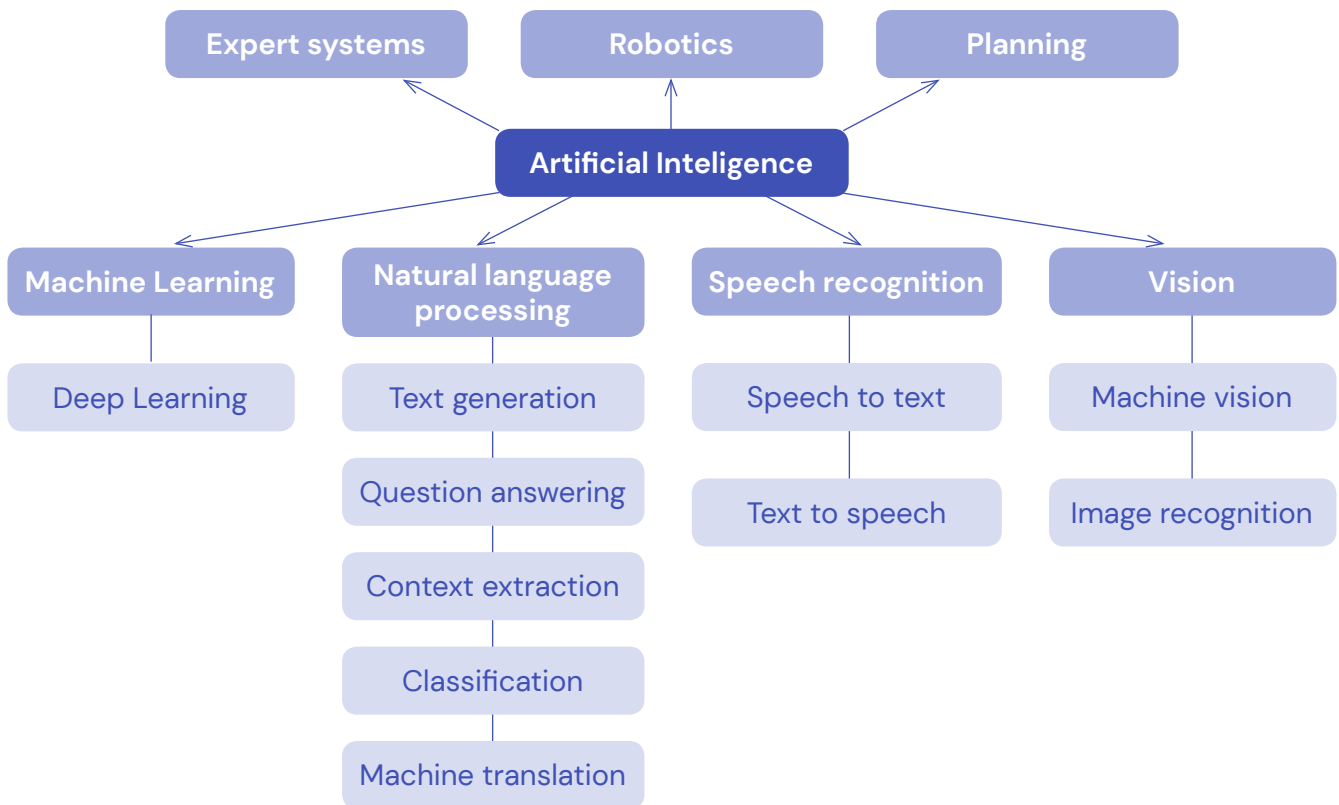
Natural language processing is the field in which researchers explore how computers can be used to understand and manipulate natural language. This subfield has a lot of influence from other fields such as linguistics, engineering, psychology, ML etc. Applications are machine translation, summarization of texts, speech recognition, text generation, question answering, and many more.

Robotics

A final interdisciplinary branch of AI involves the design, implementation, manufacturing and operation of robots. We encounter robots more and more in everyday life. Where they seemed to be used only in military settings and manufacturing units, we now regularly see self-driving vehicles, floor-cleaning robots, exoskeletons that help workers with the heavy lifting, robotic arms that take over some basic functionalities for amputees and robots that show you the way in airports, hotels, etc. They are widely used in assisted surgeries, autonomous warehouses and they are used to explore other planets like Mars. In order to function well, a lot of AI techniques but also engineering has to be combined. They need to observe the world, they need to interpret the observations and they need to know which actions to take in which situations. Moreover, now robots are becoming more and more part of our daily lives, researchers from social sciences are also getting involved: these robots will need to learn how to communicate and socially interact with us.

In the following page, Figure 6 shows a schematic of the different branches of AI

Figure 6. Hierarchical tree of the types of AI.



Machine Learning

In the rest of the background section, the focus will be on the subfield ML and more in particular the task classification, as these concepts are ubiquitous in AI applications.

Similar to how humans need a functioning brain even for survival, computers need instructions to do even the simplest things. Without code or instructions, a computer is basically useless. To describe what computers should do, humans write algorithms. These algorithms are then translated to code that can be executed by the computer. In ML, researchers investigate how computers or algorithms can learn from data without being explicitly programmed. The basic idea is to start with a base algorithm and let this algorithm improve itself until it can perform a given task in a manner that humans decide is good enough. Again, this idea can be compared to how humans learn. At birth you can only perform vital functions, during the first years of your life you learn a lot from examples, from interaction, etc. Typically, ML is classified into 3 types of algorithms:

- **Supervised learning**
- **Unsupervised learning**
- **Reinforcement learning**

In supervised and unsupervised learning, the algorithms learn from data. In reinforcement learning, the algorithms learn from interaction with an environment. In the following sections a brief overview will be given about the three types of learning algorithms.

Supervised learning

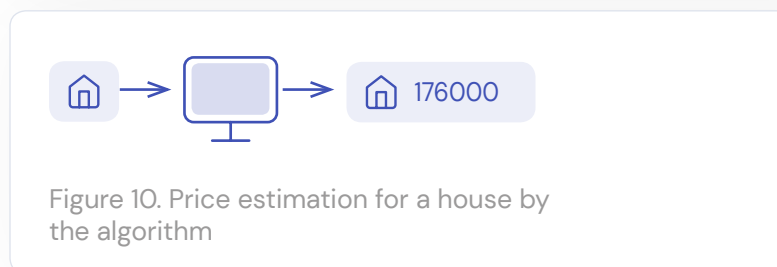
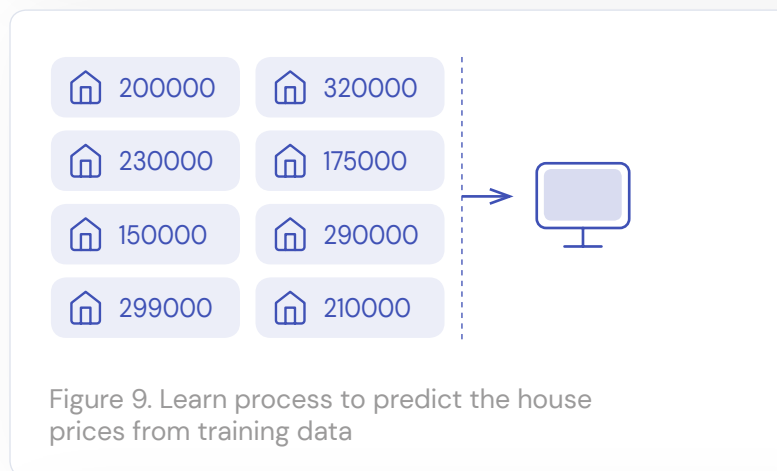
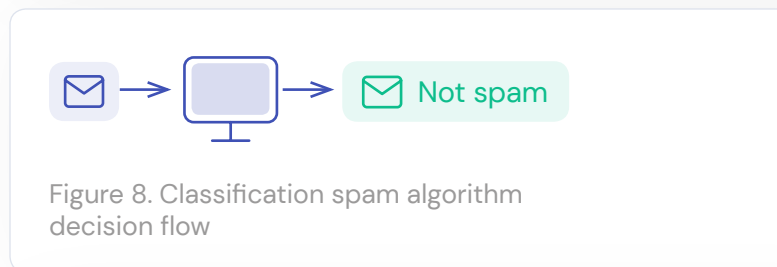
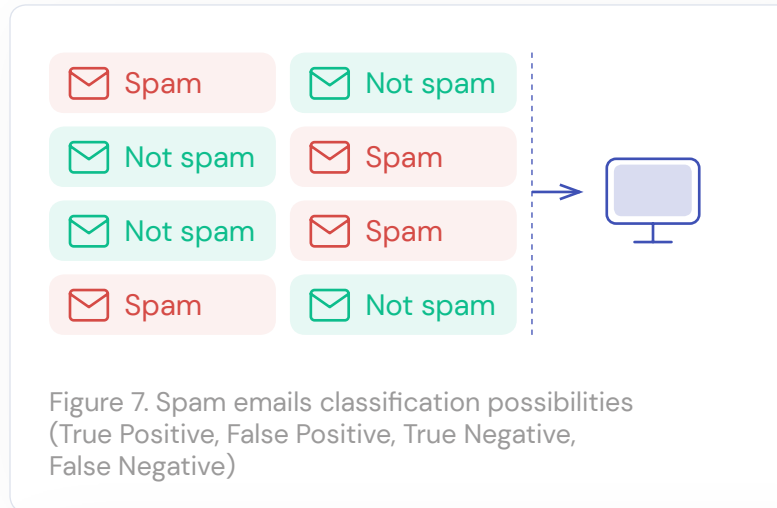
In supervised learning, an algorithm is presented with labelled data. Labelled data is a set of samples with informative tags. A typical example of such data is emails that are labelled as spam or not spam, represented in Figure 7.

The algorithm then learns to predict spam (or not spam) for emails it has not seen before. This is an example of a classification problem, where the output variable is a category like “spam” and “not spam”, as shown in the Figure 8.

Another example of a classification problem is when you want to predict “disease” or “no disease”.

The output variable can also be a real value, such as “euros” where the algorithm learns to predict the prices of a house based on other parameters such as area, total square metres, etc, as shown in Figure 9.

After the learning phase, it will predict the price for a new type of house. This kind of problem where you want to predict a real number instead of a category is called regression, as shown in Figure 10.

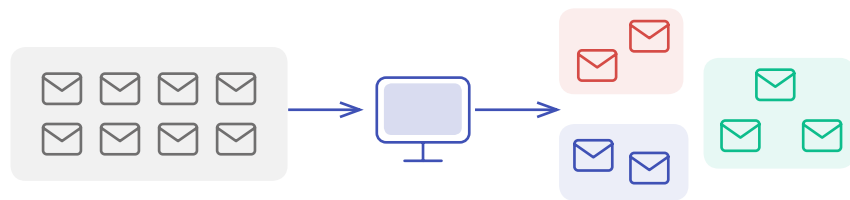


Unsupervised learning

In unsupervised learning, an algorithm is presented with unlabelled data. For instance, for the spam emails, the computer would be presented with emails but without any additional information about whether these are spam or not. Figure 11 shows how the computer receives emails and must categorise them according to their type.

The computer will then try to discover groups in the data. It will not label them however. This type of machine learning is called clustering. It is for instance used for customer segmentation where a company divides its customer base into groups that are similar. This way similar clients can be targeted in a similar way. Another application of unsupervised learning is where a company wants to discover associations between products. For instance: a customer that buys product A will likely also buy product B.

Figure 11. Classification of emails process



Reinforcement Learning

Finally, in a reinforcement learning problem, an algorithm learns not from data but from interaction. It uses a goal-based learning technique in which an algorithm learns by trial and error which actions it should take in different situations or states in order to optimise a reward. It is the main technology behind autonomous vehicles and it is used for numerous tasks such as real-time bidding and automated temperature control in data centres.

Many of the core algorithms in reinforcement learning were originally inspired by biological learning systems such as operant conditioning, a behaviourist technique in psychology, where behaviour is reinforced by positive or negative stimuli. Operant conditioning was extensively studied by Edward L. Thorndike (1874–1949) who observed the behaviour of hungry cats trying to escape puzzle boxes. Cats could escape the box and reach the food by means of simple actions such as pulling a cord. At first it took the cats a long time to get out of the box, but on successive trials it became easier to a point where the cats would be able to escape without hesitation.

Humans are really good at learning by trial and error. A typical example is how humans learn how to ride a bike: you cannot learn it from a book, you cannot learn it by watching others, the only way is by doing. At first you will make mistakes and probably fall off (negative reward) but eventually, after several failed attempts, you will figure out how to do it. A second example is parents who treat their children to chocolate when they finish their homework. Here we already see a potential pitfall: the child learns that they will receive a treat or a reward when they finish their homework, but they will not learn how to do it without mistakes. Indeed, when programming a reinforcement learning algorithm setting a correct reward is very important.

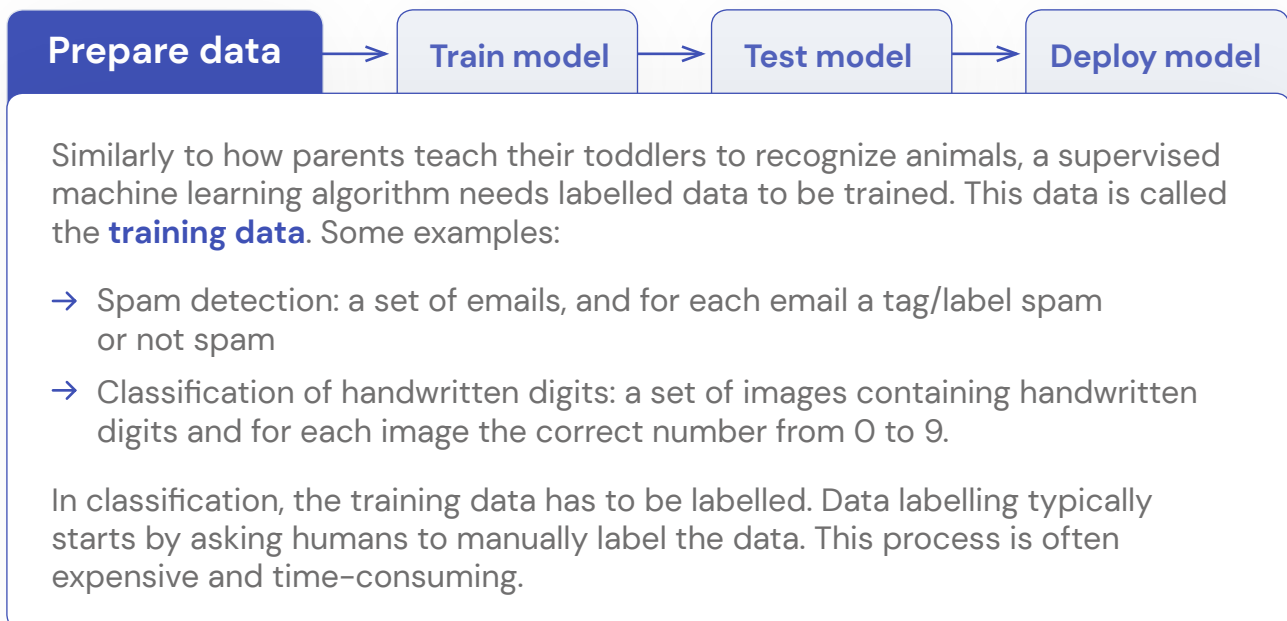
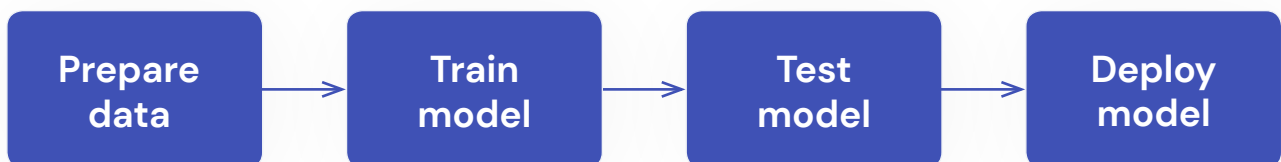
Classification

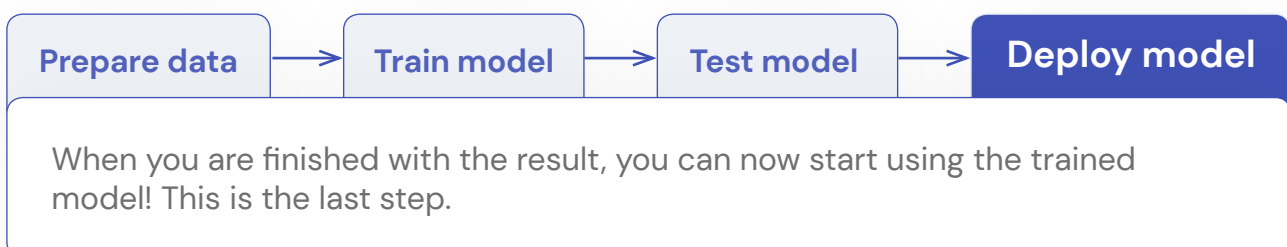
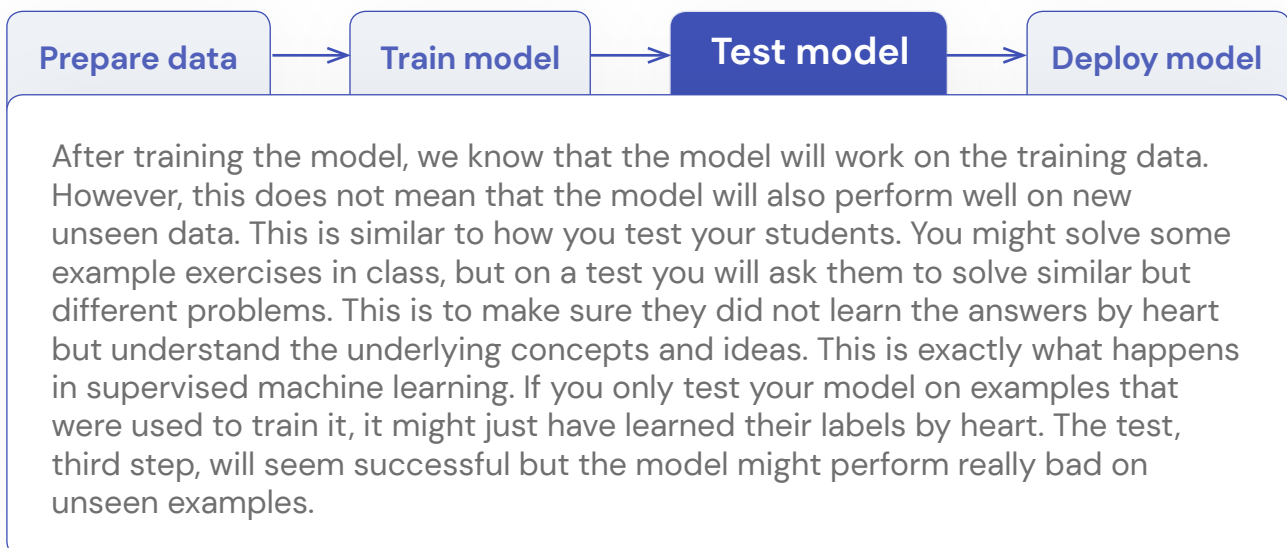
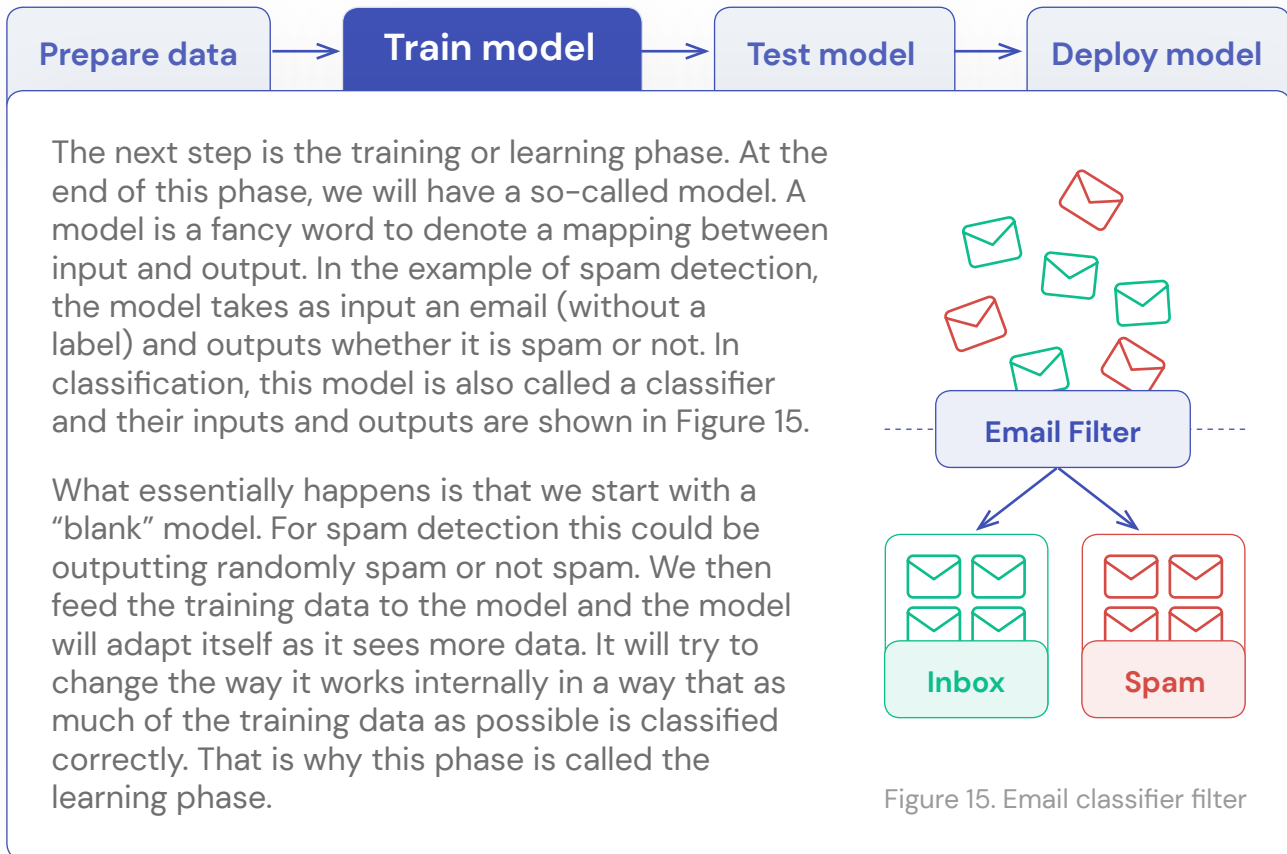
In this subsection we will take a deeper look at the classification problem. Classification basically means categorising a set of data into classes. Typical examples are spam detection, classification of handwritten digits, conversion prediction (buy or not buy), plant species classification, fraud detection..., just to name a few.

In AI, classification is a supervised machine learning technique that was already briefly discussed in the previous section. The way an algorithm is trained to classify data is similar to how parents teach their babies or toddlers to recognize animals: they point out an image and say “bear” or “cow”. The words “bear” and “cow” are the labels.

Roughly speaking, classification involves the steps illustrated in Figure 12.

Figure 12. Classification model preparation steps.





Important note on the choice of training data

Besides keeping truly unseen test data until the end, the choice of training data is also very important. Clear errors, mislabelled data and conflicting data should be avoided. Just as people get confused, training a model using such data could create anomalies. Insufficient data is also a reason why a model predicts with a low level of accuracy. It will not have seen enough data to understand the characteristics of what makes an email spam and what not.

It can get really dangerous when your training data is biased. More information and examples can be found in the algorithmic bias section.

Introduction to the LearningML tool

The previous section provided a brief introduction to the concepts that will be used throughout this handbook. In the activities that will be presented in the next section we will use a web tool called [LearningML](#), which is developed by [Juan David Rodriguez](#) a teacher and expert in software development.

This tool allows us to interact and experiment with AI models: it can be used to train models, to investigate how bias can be introduced in AI systems, etc.

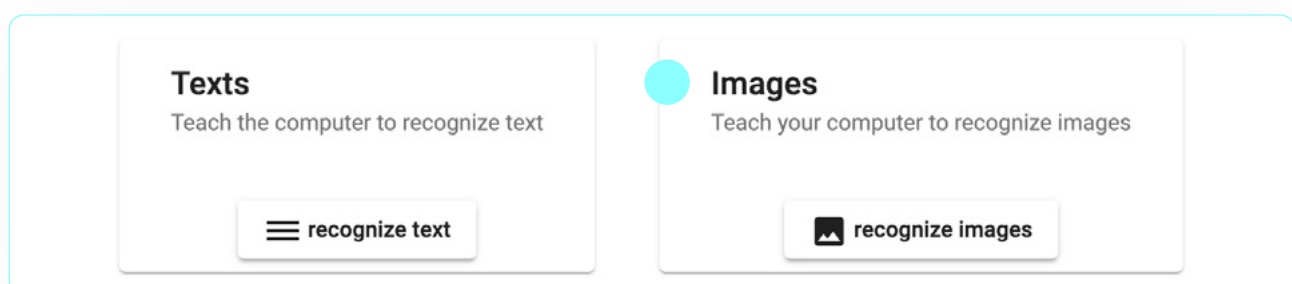
In this chapter we will introduce the tool and provide a small demo in which the basic functionality of the LearningML tool is shown. We will also revisit the steps necessary to create an AI model: prepare data, train model, test model. Finally, we will show some examples about algorithmic bias.

Demo Learning ML tool

Go to the tool: [LearningML](#).

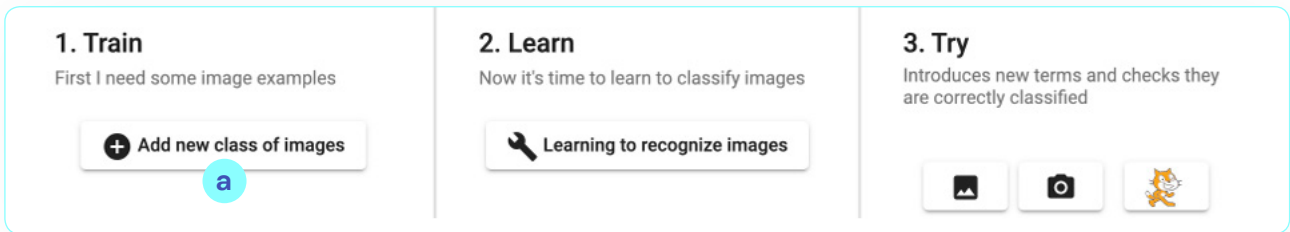
In this demo we are going to work with images: **select “recognize images”** in Figure 18. If you want to follow along you will need an image of a dog, and two images of a cat. You can also use the images used in the demo. They can be found in [this link](#).

Figure 18. LML V.1.2. classification options

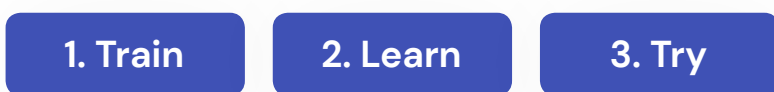


You will see the steps represented in Figure 19.

Figure 19. The three steps to use LML tool.



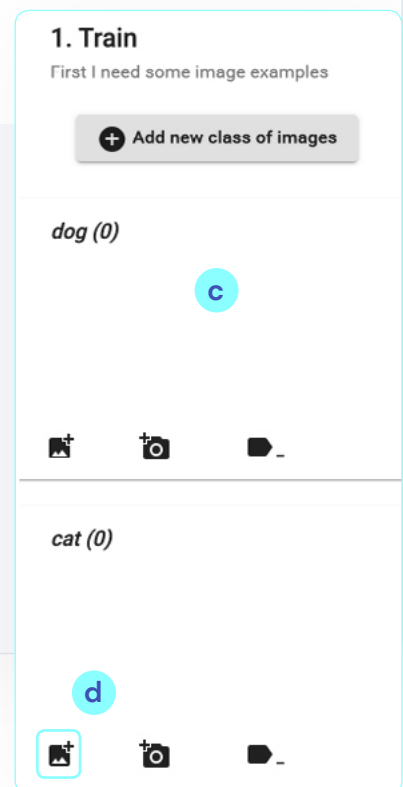
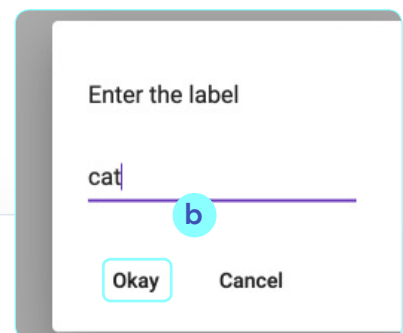
The process contains 3 steps:



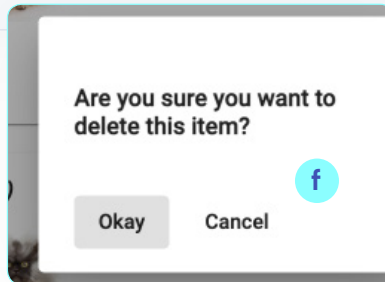
Step 1: Train

In this step we give the computer a number of images of dogs and cats. To help the computer we provide the correct answers, or labels.

- a Select "Add new class of images". This class corresponds to the label.
 - b Add a class "cat" and click "Okay".
 - c Repeat this step but now add a class "dog". You should now see something like image marked with c.
- Note** that it is not restricted to only two classes, you can add more classes. In this demo however we will use only dogs and cats.
- d You can now add images by clicking on the "add image" button for both classes.



- e After adding one image of a cat and one of a dog, you will see something similar to what is shown in figure marked with the letter e.
- f The uploaded images together with their label are called the training data. If you made a mistake and want to remove an image you can do so by clicking on that particular image and confirm you want to delete it.



1. Train
First I need some image examples

+ Add new class of images

dog (1)

cat (1)

Step 2: Learn

Now click on “Learning to recognize images”, as shown in figure marked with letter g. The model now goes through a learning phase in which it learns from the training data how it should classify dogs and cats.

2. Learn
Now it's time to learn to classify images

g Learning to recognize images

Step 3: Try

Finally we test our classifier. We do this by uploading images. This is called the test data.

- h Click on “upload a test image”, as shown and choose an image to use for testing. Choose an image you did not use as training data.
 - i The computer will now predict a label for your uploaded image. You will see something like the image shown in figure with the marked letter i.
- In this case, the prediction is “cat” since the computer’s score for the prediction is higher for cat (75.22%) than for dog (24.78%).
- These percentages quantify the probability that the image used belongs to one class or another.

3. Try
Introduces new terms and checks they are correctly classified

h

3. Try
Introduces new terms and checks they are correctly classified

I think it belongs to the class cat, isn't it?

i

- cat (75.22 %)
- dog (24.78 %)

Introduction to supervised machine learning

Subject

This class is a generic one, not related to any subject.

Age

Although it can be adapted to any age, it is recommended for use in groups of 10-12 years of age.

Duration

1 hour

Description and learning objectives

Students are introduced to the concept of classification. By exploring the LearningML tool, students learn about supervised machine learning. The students will learn by doing how important the training data plays in classification accuracy. In a separate follow up lesson students will work with a biased dataset and learn about algorithmic bias.

Situation before

This class does not require any particular knowledge besides being able to work with a computer.

Educational goals

This class is a generic one, not related to any subject. The goal of this activity is to provide a short and easy example to introduce AI to students.

Type of instruction

Indirect instruction where students work in groups and discuss and solve problems.

Materials, resources, equipment

- Each group of students will need a computer that has access to the internet.
- The link to the tool that is used in the lesson is: <https://learningml.org/editor/>
- An activity sheet for the students is provided at the end of this lesson plan.
- The images¹ used can be found in [this folder](#).

References and sources

Machine learning, classification

Teacher guide

Start with a class discussion.

Ask students about examples of classification that they use in their life or that they have seen in school. If needed, some examples to get the discussion started: spam detection, animal classification, the different sections in a library, handwriting detection used by postal services, movie genres, ...

Explain what we will do in this class.

In this class we will build our own classifier. We will teach a computer how to classify dogs and cats. We will do this by giving a lot of examples: we will show pictures of dogs and cats and we will also tell the computer which are dogs and which are cats. Remind the students of how parents teach their toddlers to recognize animals. They will sit with their child and point at pictures of animals in a book and say “cow” or “bear”.

Demo of LearningML

See **previous section** for a demo that can be shown in class. Make sure to introduce the concepts of training data and test data.

Let the students work with the tool.

An **activity sheet** is provided and links to folders containing images are provided. The students will be asked some questions.

Some pointers that can help you in guiding the activity and discuss these questions:

1. What happens if you test your classifier using the training data?

The prediction is correct and the probability score for the prediction is high. This is because the classifier was able to learn the label for this image by heart. Using training data as test data is not a good idea, it will give you the impression the classifier works well, but it does not tell you how well it will work on unseen examples.

2. What happens if you test the classifier for dogs/cats with an image of a lion?

It will output either dog or cat. In a way, these algorithms are really dumb. They figure out whether your image resembles most the dogs or the cats in your training data. It is not smart enough to tell you it is neither.

3. What happens when you train the classifier using more training data?

Generally speaking, the predictions will be better and the probabilities will be higher. The more examples the classifier can use to train on, the better it will understand the differences between the classes.

However, you need to be careful which data you add. By adding more images you can also add (more) bias in the data. For instance if you add only pictures of white cats, the classifier might understand that a white fur implies that it is a cat. Bias will be discussed in the follow up lesson.

4. What happens if you have only a limited number of images of dogs and many more of cats?

Most predicted labels will be cats. This is because there were not a lot of dogs in the training data and the classifier could not learn really well how to differentiate a dog from a cat. Most test data will have more in common with the cat training data than with the dog training data.

5. Is it ok if some cats/dogs are misclassified?

This is not a big problem. As mentioned before, there might not be enough training data, making it difficult for the classifier to distinguish them. There is only a problem if it systematically misclassified all the cats or all the dogs.

A filled out version of the activity sheet with some extra pointers and reflections, which is provided in the next section, is added at the end of this introduction to supervised learning chapter.

1. Go to <https://learningml.org/> and choose “**recognize images**”.

2. In this lesson we use the images provided in [this folder](#). This folder has several subfolders:

training_set

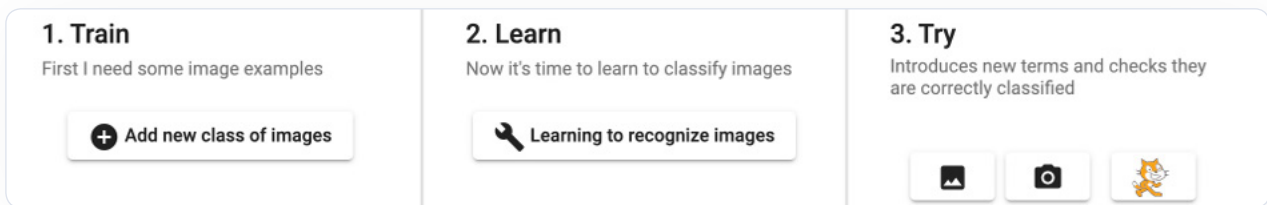
Test_set

lions

extra

3. Identify the three parts of a supervised machine learning system: **training data, learning phase and test data.**

Figure 27. Introduction – LML steps to train a classification model



4. Let's train the computer to classify dogs and cats.

- Add a new class of images: **dog**.
- Add a new class of images: **cat**.
- Now go to folder training set and upload the images of cats and dogs to the corresponding classes.
- Click on “**Learning to recognize images**”.
- Now test the classifier using the images in test_set and write down in **Table 1**. Classification results score table. predicted class, probability score and actual class.
- Answer the following questions:**
 - Are all predictions correct? If not, what do you think has happened?

Table 1. Introduction – Step 1: Classification results score table

Image	Predicted class	Score	Actual class
cat11			cat
cat12			cat
cat13			cat
cat14			cat
cat15			cat
dog11			dog
dog12			dog
dog13			dog
dog14			dog
dog15			dog

5. Let's keep the same training data but add more examples.

- Go to the extra folder and upload the images of cats and dogs to the corresponding classes.
- Click on "**Learning to recognize images**".
- Retest using the provided images in the test set folder and write down in **Table 2** predicted class, probability score and actual class.
- Answer the following questions:**
 - Compare these results with your previous results. Have predictions improved? Why?

Table 2. Introduction – Step 2: Classification results score table.

Image	Predicted class	Score	Actual class
cat11			cat
cat12			cat
cat13			cat
cat14			cat
cat15			cat
dog11			dog
dog12			dog
dog13			dog
dog14			dog
dog15			dog

6. Let's keep the same training data.

- Test the classifier on the following images from the training set and write down in **Table 3** predicted class, probability score and actual class.
- Answer the following questions:**
 - Why do you think the predictions are so good?
 - Is it a good idea to use training data for testing the classifier?

Table 3. Introduction – Step 3: Classification results score table.

Image	Predicted class	Score	Actual class
cat1			cat
cat2			cat
dog2			dog
dog3			dog

7. Let's keep the same training data

- a. Test the classifier on image lion1 in folder **lions** and write the predicted class and score in **Table 4**.
- b. **Answer the following questions:**
 - How did the model classify the lion?
 - Can you explain what has happened?

Table 4. Introduction – Step 4: Classification results score table.

Image	Predicted class	Score	Actual class
lion1			lion

8. Let's change the training data

- a. Remove all dogs. You can do this by clicking on the image you want to remove and confirm that you want to delete it.
- b. Go to folder **training_data** and add the 4 dog examples to the dog class.
- c. You should now have 4 training examples for dogs and 9 for cats.
- d. Click on **"Learning to recognize images"**,
- e. Retest using the provided images in the test set folder and write down in **Table 5** predicted class, probability score and actual class.
- f. **Answer the following questions:**
 - Why are the results so bad?

Table 5. Introduction – Step 5: Classification results score table.

Image	Predicted class	Score	Actual class
dog11			dog
dog12			dog
dog13			dog
dog14			dog
dog15			dog

The values given here were obtained with the version of the tool available on April 7, 2022. The values obtained by the students can be slightly different, but the overall trend should be the same.

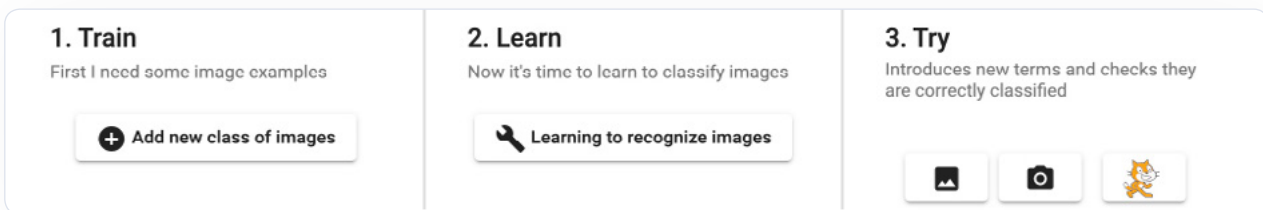
1. Go to <https://learningml.org/> and choose “recognize images”.

2. In this lesson we use the images provided in [this folder](#). This folder has several subfolders:



3. Identify the three parts of a supervised machine learning system: **training data, learning phase and test data.**

Figure 28. Introduction - LML steps to train a classification model



4. Let’s train the computer to classify dogs and cats.

- a. Add a new class of images: **dog**.
- b. Add a new class of images: **cat**.
- c. Now go to folder training set and upload the images of cats and dogs to the corresponding classes.
- d. Click on “**Learning to recognize images**”.
- e. Now test the classifier using the images in test_set and write down the predicted class, probability score and actual class. **Table 6** shows the results that should be obtained.
- f. **Answer the following questions:**
 - Are all predictions correct? If not, what do you think has happened?

Table 6. Introduction - Step 1: Classification results score table

Image	Predicted class	Score	Actual class
cat11	cat	87.04%	cat
cat12	cat	66.98%	cat
cat13	cat	94.90%	cat
cat14	cat	61.28%	cat
cat15	cat	78.72%	cat
dog11	dog	90.31%	dog
dog12	dog	58.75%	dog
dog13	dog	70.17%	dog
dog14	dog	71.13%	dog
dog15	dog	68.82%	dog

In this example we have one dog that was incorrectly classified. The reason this dog was classified as a cat is because the dog has more in common (colour of fur, placing of the paws) with the cats in the dataset than with the dogs.

5. Let's keep the same training data but add more examples.
 - a. Go to the extra folder and upload the images of cats and dogs to the corresponding classes.
 - b. Click on "Learning to recognize images".
 - c. Retest using the provided images in the test_set folder and write down the predicted class, probability score and actual class. **Table 7** shows the results that should be obtained.
 - d. **Answer the following questions:**
 - Compare these results with your previous results. Have predictions improved? Why?

Table 7. Introduction - Step 2: Classification results score table.

Image	Predicted class	Score	Actual class
cat11	cat	90.35%	cat
cat12	cat	60.63%	cat
cat13	cat	88.17%	cat
cat14	cat	74.53%	cat
cat15	cat	69.78%	cat
dog11	dog	94.46%	dog
dog12	dog	85.31%	dog
dog13	dog	90.93%	dog
dog14	dog	90.01%	dog
dog15	dog	89.16%	dog

Not all prediction scores are higher, but overall the predictions for dogs have gotten better. The classifier is getting better in distinguishing cats from dogs.

6. Let's keep the same training data.
 - a. Test the classifier on the following images from training_set and write down the predicted class, probability score and actual class. **Table 8** shows the results that should be obtained.
 - b. **Answer the following questions:**
 - Why do you think the predictions are so good?
 - Is it a good idea to use training data for testing the classifier?

Table 8. Introduction - Step 3: Classification results score table.

Image	Predicted class	Score	Actual class
cat1	cat	98.53%	cat
cat2	cat	97.39%	cat
dog2	dog	99.31%	dog
dog3	dog	98.82%	dog

The classifier works really well on all examples in the training data. Probability scores for the predictions are high. This is because the classifier was able to learn the label for this image by heart. Using training data as test data is not a good idea, it will give you the impression the classifier works well, but it does not tell you how well it will work on unseen examples.

7. Let's keep the same training data

- a. Test the classifier on image lion1 in the **lions** folder. **Table 9** shows the results that should be obtained.
- b. **Answer the following questions:**
 - How did the model classify the lion?
 - Can you explain what has happened?

Table 9. Introduction – Step 4: Classification results score table.

Image	Predicted class	Score	Actual class
lion1	dog	80.10%	lion

The classifier classifies the lion as a dog. The classifier will always output either cat or dog. Based on the training data, the lion resembles a dog.

8. Let's change the training data

- a. Remove all dogs. You can do this by clicking on the image you want to remove and confirm that you want to delete it.
- b. Go to folder **training_data** and add the 4 dog examples to the dog class.
- c. You should now have 4 training examples for dogs and 9 for cats.
- d. Click on "**Learning to recognize images**",
- e. Retest using the provided images in folder test set folder and write down the predicted class, probability score and actual class. **Table 10** shows the results that should be obtained.
- f. **Answer the following questions:**
 - Why are the results so bad?

Table 10. Introduction – Step 5: Classification results score table.

Image	Predicted class	Score	Actual class
dog11	dog	63.64%	dog
dog12	cat	51.06%	dog
dog13	cat	64.04%	dog
dog14	cat	69.29%	dog
dog15	cat	68.90%	dog

Almost all dogs are classified as cats. This is because there weren't a lot of dogs in the training data and the classifier could not learn really well how to differentiate a dog from a cat.

Lessons plans related to algorithmic bias

Activity one: Generic class about algorithmic bias

Subject

This class is a generic one, not related to any subject.

Age

This class is not related to any age, but it can be adapted.

Duration

1 hour

Description and learning objectives

This class can stand on its own or could be organised as a follow up class for the class on supervised machine learning. In this lesson, students will unknowingly use a data set that is biased. They will learn what algorithmic bias is and how it is caused. They will reflect on how to prevent and deal with algorithmic bias.

Situation before

This class does not require any particular knowledge besides being able to work with a computer. If not given as a follow-up class, the teacher should give a demo of the LearningML tool.

Educational goals

This class is a generic one, not related to a specific subject. Nevertheless, this activity introduces critical thinking and bias to students, allowing them to distinguish when it exists and, above all, how to avoid it.

Type of instruction

Indirect instruction where students work in groups and discuss and solve problems.

Materials, resources, equipment

- Each group of students will need a computer that has access to the internet.
- The link to the tool that is used in the lesson is: <https://learningml.org/editor/>
- An activity sheet for the students is provided at the end of this lesson plan.
- The images¹ used can be found in [this folder](#).

References and sources

Machine learning, classification, algorithmic bias

Teacher guide

Start with a class discussion

Ask students if they have encountered situations where technology was unfair towards them or someone they know. If needed prompting, ask them how technology could be biased.

Some typical examples if needed prompting:

- Some facial recognition systems do not recognize darker female faces as lighter male faces.
- A soap dispenser that does not work for black people: https://www.youtube.com/watch?v=YJjv_OeiHmo
- The system of the bank not approving your loan because you live in a certain underdeveloped area.
- An online platform that shows jobs ads in nursing or secretarial work primarily to women.

When algorithms have outcomes that are unfair in a systematic way, we call this algorithmic bias.

Recall from last lesson if used as a follow-up class

Ask them what happened in the last lessons when we trained the classifier using a small set of dogs and a large set of cats in the training data.

Most predicted labels were cats. This is because there were not a lot of dogs in the training data and the classifier could not learn really well how to differentiate a dog from a cat. Most test data will have more in common with the cat training data than with the dog training data.

We say that our algorithm is biased towards cats since it works really well for them and biased against dogs since it does not work as well for them.

If not used as a follow-up class, give a demo on the tool

See previous section for a demo that can be shown in class. Make sure to introduce the concepts of training data and test data.

Let the students work with the tool

An activity sheet is provided.

The images the students will be working with can be found in [this folder](#). In the activity sheet the students will be asked some questions. A filled out version of the activity sheet with some extra pointers and reflections is added after the activity sheet.

Reflect

The classifier classified all ginger dogs as cats and except for the ginger cat all cats were misclassified as dogs. The classifier had the following bias: it assumed that in order to be a cat, the fur must be ginger. This bias comes directly from the training data: all cats in the training data are ginger.

Ask the students if they can think of a real-world situation where training data is biased.

A typical example is the one where facial detection systems work well on lighter males, but bad on dark females. This is because standard training databases are predominantly white and male.

Why is this potentially harmful?

For instance, facial recognition used on the street can lead to false arrests because people were incorrectly identified.

This is something we need to be careful of when using data to train a classifier, our classifier should not exclude anyone.

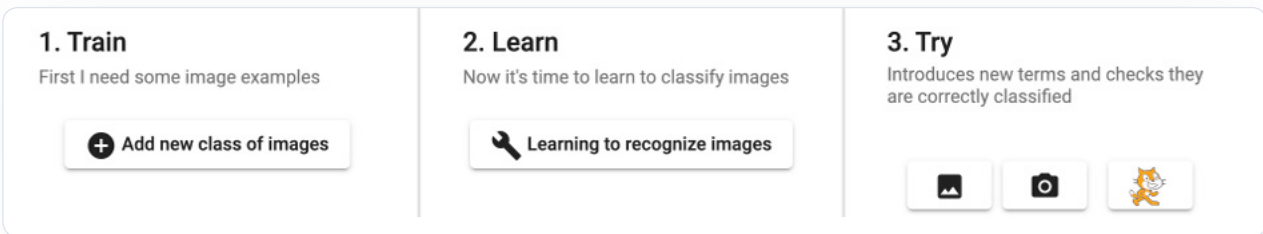
1. Go to <https://learningml.org/> and choose “**recognize images**”.

2. In this lesson we use the images provided in [this folder](#). This folder has several subfolders:



3. Identify the three parts of a supervised machine learning system: **training data, learning phase and test data.**

Figure 29. Introduction – LML steps to train a classification model



4. Let’s train the computer to classify dogs and cats.

- a. Add a new class of images: **dog**.
- b. Add a new class of images: **cat**.
- c. Now go to the **training set** folder and upload the images of cats and dogs to the corresponding classes.
- d. For the dog training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?
- e. For the cat training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?
- f. Click on “**Learning to recognize images**”.
- g. Now test the classifier using the images in test set and write down the predicted class, probability score and actual class.
- h. For what class did your classifier perform best? Why?

Table 11. Algorithmic Bias – Step 1: Classification results score table.

Image	Predicted class	Score	Actual class
cat11			cat
cat12			cat
cat13			cat
cat14			cat
cat15			cat
dog11			dog
dog12			dog
dog13			dog
dog14			dog
dog15			dog

5. Try to fix the training data set. Go to the extra folder and augment your data set by choosing images of cats and dogs and uploading them to the corresponding classes. If you want, you can also browse the internet and upload images.

- a. For the dog training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?
- b. For the cat training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?
- c. Click on **“Learning to recognize images”**.
- d. Test the classifier on the test dataset and write down the predicted class, probability score and actual class.
- e. For what class did your classifier perform best? Or did it work the same for both classes? How did the changes you made have an impact? Why?

Table 12. Algorithmic Bias – Step 2: Classification results score table.

Image	Predicted class	Score	Actual class
cat11			cat
cat12			cat
cat13			cat
cat14			cat
cat15			cat
dog11			dog
dog12			dog
dog13			dog
dog14			dog
dog15			dog

The values given here were obtained with the version of the tool available on April 7, 2022. The values obtained by the students can be slightly different, but the overall trend should be the same.

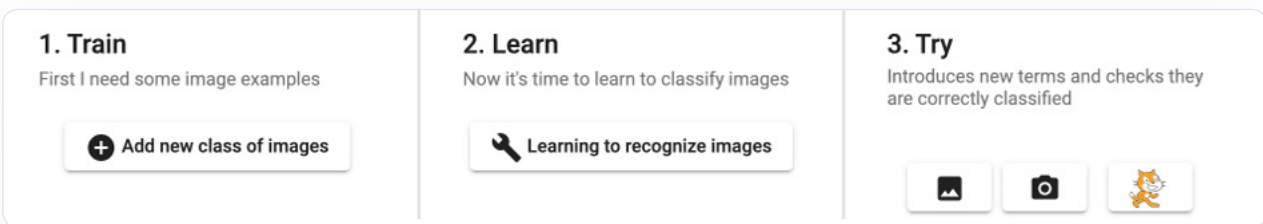
1. Go to <https://learningml.org/> and choose “recognize images”.

2. In this lesson we use the images provided in [this folder](#). This folder has several subfolders:



3. Identify the three parts of a supervised machine learning system: **training data, learning phase and test data.**

Figure 30. Introduction - LML steps to train a classification model



4. Let’s train the computer to classify dogs and cats.

- a. Add a new class of images: **dog**.
- b. Add a new class of images: **cat**.
- c. Now go to the **training set** folder and upload the images of cats and dogs to the corresponding classes.
- d. For the dog training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?

- There are 6 images.
 - They have different colours of fur and different breeds. Only two dogs have the same colour of fur.
 - There is a variety of sitting positions; some are walking.

- e. For the cat training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?

- There are 6 images.
 - All are ginger cats with stripes.
 - There is a variety of sitting positions; some are walking.

f. Click on “**Learning to recognize images**”.

- g. Now test the classifier using the images in test set and write down the predicted class, probability score and actual class.
- h. For what class did your classifier perform best? Why?

All cats are misclassified, except for the ginger cat. The ginger dog was classified as a cat. The classifier has learned that if the colour of the fur is ginger, it must be a cat.

Table 13. Algorithmic Bias - Classification results score table answers.

Image	Predicted class	Score	Actual class
cat11	cat	87.02%	cat
cat12	dog	61.87%	cat
cat13	dog	70.33%	cat
cat14	dog	55.93%	cat
cat15	dog	55.64%	cat
dog11	dog	84.57%	dog
dog12	dog	93.76%	dog
dog13	dog	70.94%	dog
dog14	cat	52.77%	dog
dog15	dog	91.61%	dog

Important note for question 5. There is no unique solution, this question should be seen as an experiment. The students try to add images in order to make the predictions better. The objective is to make sure more cats are classified as cats and not as dogs while making sure the dogs are still classified as dogs. If the students are not happy with the results, they can add even more images or remove some of the images they uploaded. Remind them they have to retrain the computer after changing the training data.

5. Try to fix the training data set. Go to the extra folder and augment your data set by choosing images of cats and dogs and uploading them to the corresponding classes. If you want, you can also browse the internet and upload images.

- a. For the dog training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?

As an example we have added the following dog images from the extra folder: dog22, dog23, dog25. This way we have added even more variety: we now also have a white dog for example.

- b. For the cat training dataset:
 - How many images are included?
 - Are the images similar?
 - Are the images different?

As an example we have added the following cat images from the extra folder: cat23, cat24, cat25. We now have a more balanced dataset: there is more variety in the cat training data, not all cats are ginger and we also have cats without stripes.

- c. Click on **“Learning to recognize images”**.
- d. Test the classifier on the test dataset and write down the predicted class, probability score and actual class.

- e. For what class did your classifier perform best? Or did it work the same for both classes? How did the changes you made have an impact? Why?

The accuracy of the classifier has increased significantly. There are still two animals that have been incorrectly classified: cat13 and dog13. If you take a closer look at the training data, you can see why this could be happening:

- cat13 resembles dog1 from the data training, much more than it resembles any of the cats in the training data.
- dog13 resembles the newly added cat25 from the training data, much more than it resembles any of the dogs in the training data.




Table 14. Algorithmic Bias – Step 2: Classification results score table.

Image	Predicted class	Score	Actual class
cat11	cat	96.43%	cat
cat12	cat	84.88%	cat
cat13	dog	70.71%	cat
cat14	cat	74.66%	cat
cat15	cat	70.64%	cat
dog11	dog	89.86%	dog
dog12	dog	87.87%	dog
dog13	cat	55.41%	dog
dog14	dog	59.77%	dog
dog15	dog	80.78%	dog

Activity two: Social sciences

In this activity about bias, students will experience the effects of bias. We have created two different activities that are based on the same model created by Thomas Schelling. Links to interactive content have been provided.

Note that these activities do not require the LearningML tool.

 Subject	 Age	 Duration
This class is related to social sciences subject.	There are two versions of the experiment. The first one is more suitable for older students (from 14 onwards). The second one can even be played by younger children in primary school.	1 hour

Description and learning objectives

In this section students will learn how a small bias can have big consequences on the results. Specifically, we focus on the discipline of social sciences where bias is best identified.

Previous knowledge

- Students understand what algorithmic bias is.
- Students understand that biased data can lead to algorithmic bias.
- Students understand that personal biases of the programmer can lead to algorithmic bias.

Resources

- Example 1: [Bias game](#)
- Example 2: [Segregation](#)
- Video on YouTube: [Explanation about Schelling's model](#)

Useful background

Algorithmic bias

Type

Teach about the impact of AI

Background

When using AI based tools it is important to know that these are susceptible to bias. This is because data can be biased and every programmer has personal preferences. What people often do not realise are the effects a small bias can have. In this section we list two experiments that can be done by students to show how a small bias can have a potentially massive impact.

These experiments are based on a simple model created by Thomas Schelling in 1960. In this model a mixed group of people are spontaneously segregated by race. This happens even though no one in the population wants that outcome. Initially, black and white families are distributed randomly. At each fixed time step every family decides whether they want to move or not, depending on the racial composition of their immediate neighbourhood. This procedure is repeated until everybody finds a satisfactory home.

What happens, of course, depends on the preferences of the families. If everybody is indifferent about race, no one will move. On the other end, if nobody wants to live next to someone of another race, total segregation is the only solution. The interesting cases are the ones in between these two extremes. Schelling discovered that segregation could emerge when families do not want to be in the minority. A relatively mild preference for neighbours of similar race can lead to high levels of segregation.

Example 1: Bias game

Create 21 cards, using 3 different colours: of each colour you have 7 cards. Each colour represents a type of person, e.g. a race. Start by placing three different cards next to each other, as shown in Figure 30.

The remaining cards are shuffled and placed in the middle. Students go in turns. When a student takes a card, he must place it on the table. However, no-one wants to be the only one that looks like us in a neighbourhood. So, you can only put the card next to a person of the same type. For instance, if you need to place a red person, you can only put it in the positions shown in Figure 31.

When the students are done with adding all the cards, they see that they've created a segregated group of people.

Figure 30. Social Sciences Activity: People symbolic representation.

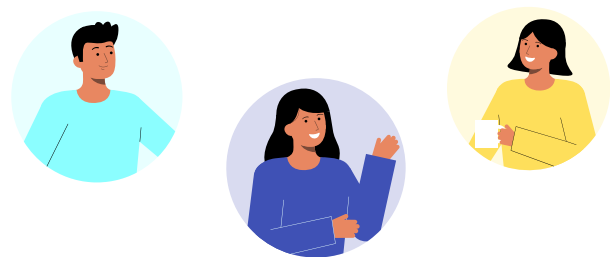
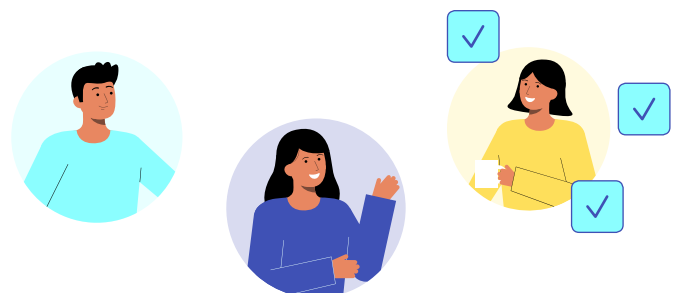


Figure 31. Social Sciences Activity: People cards representation.



Example 2: Segregation

This experiment is directly based on Schelling's model and is also explained in the [video](#). If you want, you can even use eggs like in the video. Otherwise, let them work in groups of 4-5. For each group:

- Draw a 6 by 6 grid on a large piece of paper.
- You will need 30 pieces of coloured paper that fit in the tiles you drew on the paper. You will need pieces of paper of two different colours. The experiment will work with different distributions, but it should be more or less 50-50% of each colour since we want to start in a mixed neighbourhood.
- Pick two colours that are very distinct, e.g. blue and red.
- Instead of paper, you can also use chess pieces or anything of two different colours.
- Place the pieces of paper on 30 of the tiles in the grid and mix them up in order to have a mixed neighbourhood.

Then let the experiment begin. In what follows a piece of paper will be called a house, and we have houses of red and blue colours. For every house do the following:

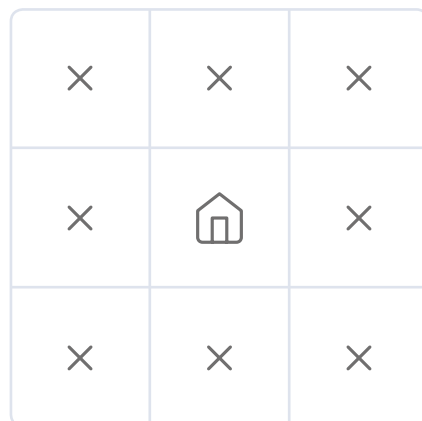
- Count the number of houses of each colour in its local neighbourhood. All x's are tiles that can contain a direct neighbour, as shown in Figure 32.
- Do not forget to also count the house itself.
- Should you move the red house in the middle? Examples:

Situation one

Red = 2 Blue = 2

Since there are at least as many red houses as blue houses we do not move the house.

Figure 32. Social Sciences Activity – Step 1: Houses example.



Situation two**Red = 3** **Blue = 2**

Since there are at least as many red houses as blue houses we do not move the house.

**Situation three****Red = 2** **Blue = 3**

Since there are more blue houses than red houses, the house should be moved.



- If the red house should be moved, move the house to an empty tile where the neighbours are such that if you move the house, there will not be more blue houses than red houses.
- The students keep on moving houses until every home is in a neighbourhood that agrees with their preferences or until it becomes clear that the neighbourhood is segregating themselves over time.

During the activity you can see how we unconsciously introduce bias when creating the neighbourhood. This can also occur when we choose the data with which to train an ML model, leading to situations that could be avoided.

Specific lessons for classroom subjects

This chapter is divided into two subsections. The first one contains ready-made lessons for specific subjects, using the LearningML tool. The second one provides an overview of tools that can be used to teach about AI in some specific subjects.

Our goal is to provide an overview of useful AI powered tools that can enhance your lessons.

Specific LearningML lessons

Mathematics

Geometry

 Subject

Geometry/Mathematics

 Age

10-12 years

 Duration

1 hour

Description and learning objectives

This example lesson is an activity that can be used in a geometry class in order to distinguish different shapes like the parallelograms, squares, trapezoids and rectangles. In this activity, the students will learn how to distinguish these shapes based on their definitions and properties. The activity uses machine learning to strengthen the learning of the learners, as well as other soft skills such as critical thinking, communication and argumentation skills.

Topic

Geometric shapes

Educational goals

- Distinguish parallelograms, squares, trapezoids and rectangles based on their characteristics.
- Be able to identify their main characteristics.
- Apply critical thinking when using AI for the classification of the shapes.

Situation before

What students need to know.

- Definitions of parallelograms, squares, trapezoids and rectangles
- Terminology: angle, sides, diagonals and triangles

Type of instruction

Indirect instruction where students work in groups and discuss and solve problems guided by the teacher.

Materials, resources, equipment

- [LearningML tool](#), in particular its image functionality
- Computer with good Internet connection
- [Link to the material](#)

Background knowledge

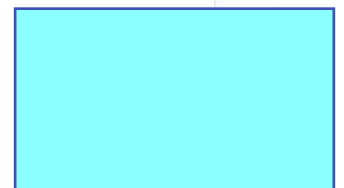
Quadrilateral: a four-sided polygon. A quadrilateral has four sides, four vertices and the interior angles add to 360 degrees. There are some special types of quadrilaterals:

Quadrilateral

Rectangle: a quadrilateral in which every angle measures 90 degrees.

Properties:

- Opposite sides are equal and parallel.
- A rectangle has two diagonals of equal length which bisect each other and divide the rectangle in two equal parts.

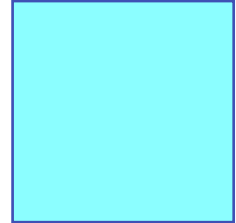


Quadrilateral

Square: a quadrilateral in which every angle measures 90 degrees and all sides are of equal length.

Properties:

- Every square is a rectangle.
- Opposite sides are equal and parallel.
- A square has two diagonals of equal length which bisect each other and divide the rectangle in two equal parts.
- The diagonals bisect each other at 90 degrees.
- The diagonals divide the square into two congruent isosceles right-angled triangles.
- A square has four lines of symmetry.
- A square has a rotational symmetry of order 4 about its centre of rotation.

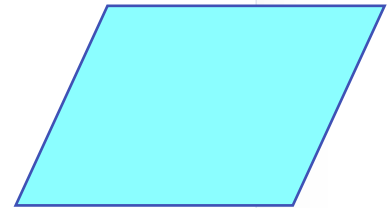


Quadrilateral

Parallelogram: a quadrilateral with opposite sides parallel

Properties:

- Every parallelogram is a trapezoid.
- Same-side interior angles supplement each other.
- The opposite sides are equal.
- The opposite angles are equal.
- The diagonals bisect each other.
- The diagonals divide the parallelogram into two congruent triangles.

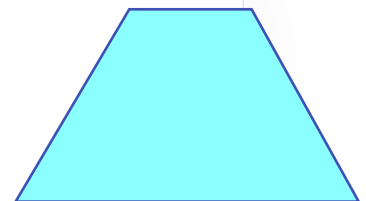


Quadrilateral

Trapezoid: a quadrilateral with one pair of parallel sides. The parallel sides are known as the bases and the non-parallel sides are known as the legs of the trapezoid.

Properties:

- Same-side interior angles supplement each other.



Description of the lesson

Step 1

5–30
min

This step includes two options (a and b) depending on the source of the images. These options are not subsequent, the teacher must select either one of them. Once the teacher has selected one of them, the teacher will ask the students to do the last part of this step, section c.

- a. Upload at least 10 pictures from each of the shapes using the provided link (5 minutes)
- b. Ask students to find a minimum of 10 images from each shape on the internet. The Google Image Search engine can be of great help for this task (30 minutes)
- c. Ask the learners to upload only 2 different pictures for each shape. By this way they will see how well the ML model is doing it by reducing training data.

All images should be uploaded to LearningML, having at the end of this step four groups of shapes (rectangles, squares, parallelograms and trapezoids).

Step 2

15 min

Let the model learn. In this step, LearningML learns to classify images, based on the data provided in Step 1. At this point it is a good opportunity to present the most important concepts of the different shapes. Learners should already be aware of the most significant differences, based on the pictures from Step 1. The most important concepts and keywords should be presented at this stage.

Step 3

30–60
min

Test & discuss. Ask the students to look for other images of shapes. Let them upload the image to LearningML to see the predicted shape. As an alternative the provided link can be used to find images. For each uploaded picture, students should answer following questions:

- Are the results correct? (i.e., is the most probable shape the one you think it is?)
- If the result is not correct, can you think of a reason why the model predicted that shape?
- What properties did you (as a human) use to identify the shapes?
- What properties do you think the algorithm uses to identify the shapes?

Ask learners to upload a picture of a table and to use it as a test image.

- What is the prediction?
- Are the results correct?
- What do you think is the most significant property or detail used by the model to identify its type?
- If the result is not correct, can you think of a reason why the model predicted that shape?

Angles of triangles

 Subject	 Age	 Duration
Geometry/ Mathematics	10-12 years	1h or 2h (depending on whether they use the existing dataset or create it on their own)

Description and learning objectives

This example lesson is an activity that can be used in a geometry class about triangles. In this activity the students will learn how to distinguish triangles based on properties of their angles. The activity uses machine learning to strengthen the learning of the learners, as well as other soft skills such as critical thinking, communication and argumentation skills.

Situation before

What students need to know:

- Acute, right and obtuse triangles
- Properties of the triangles based on their angles, such as the sum of all the angles of a triangle is 180 degrees

Educational goals

- Learn the three types of triangles, based on their angles
- Know the types of triangles
- Be able to identify their main characteristics
- Promote critical thinking when using AI in the classification of the respective triangles

Type of instruction

Indirect instruction where students work in groups and discuss and solve problems guided by the teacher.

Material, resources, equipment

- [LearningML tool](#), in particular its image functionality
- Computer with good internet connection
- [Link to the material](#)

Background knowledge

In geometry, a triangle is a polygon with three edges and three vertices. Based on the length of the sides, the measurement of the angles, there are six kinds of triangles: scalene triangle, right triangle, acute triangle, obtuse triangle, isosceles triangle, and equilateral triangle. However, in this activity we will only look at following three types:

- If one of the interior angles of the triangle measures more than 90 degrees, then the triangle is called an obtuse triangle.

Background knowledge (continued)

- An acute triangle is a triangle in which all the three interior angles measure less than 90 degrees.
- A right triangle is a triangle in which one angle measures 90 degrees. A 90-degree angle is called a right angle, and hence a triangle with a right angle is called a right triangle. In a right triangle, the Pythagorean theorem holds.

The sum of all the angles in a triangle (of all types) is equal to 180 deg.

Description of the lesson

Step 1

5–30 min

This step includes two options (a and b) depending on the source of the images. These options are not subsequent, the teacher must select either one of them. Once the teacher has selected one of them, the teacher will ask the students to do the last part of this step, section c.

- a. Upload 10 images for each type using the provided link (5 minutes)
- b. Ask the students to look up a minimum of 10 images for each shape on the internet. The Google Image Search engine can be of great help for this task (30 minutes)
- c. Ask the learners to upload only 2 different pictures for each shape. By this way they will see how well the ML model is doing it by reducing training data.

All images should be uploaded to LearningML, having at the end of this step three groups of shapes (obtuse, acute and right triangles).

Step 2

15 min

Let the model learn. In this step, LearningML learns to classify images, based on the data provided in Step 1. At this point it is a good opportunity to present the most important concepts of the different triangles. Learners should already be aware of the most significant differences, based on the pictures from Step 1. The most important concepts and keywords should be presented at this stage.

Step 3

30–60 min

Test & discuss. Ask learners to download pictures of triangles from the Internet and upload them to learningML to see the results. As an alternative, the provided link can be used to find images. For each uploaded image, learners should answer the following questions:

- Are the results correct? (i.e., is the prediction made by the AI model correct?)
- If the result is not correct, can you think of a reason why the model predicted that shape?
- What properties did you (as a human) use to identify the shapes?
- What properties do you think the algorithm uses to identify the shapes?

Ask the students to upload a square and use it as a test image.

- What is the prediction made by the model?
- Why did the model predict that shape?

Sides of triangles

 Subject	 Age	 Duration
Geometry/ Mathematics	10-12 years	1h or 2h (depending on whether they use the existing dataset or create it on their own)

Description and learning objectives

This example lesson is an activity that can be used in a geometry class about triangles. In this activity the students will learn how to distinguish triangles based on properties of their sides. The activity uses machine learning to strengthen the learning of the learners, as well as other soft skills such as critical thinking, communication and argumentation skills.

Situation before

What students need to know:

- Isosceles, scalene and equilateral triangles
- Properties of the triangles based on their angles, such as the sum of all the angles of a triangle is 180 degrees

Educational goals

- Learn the three types of triangles, based on their sides.
- Know the types of triangles
- Be able to identify their main characteristics
- Promote critical thinking when using AI in the classification of the respective triangles

Type of instruction

Indirect instruction where students work in groups and discuss and solve problems guided by the teacher.

Material, resources, equipment

- [LearningML tool](#), in particular its image functionality
- Computer with good internet connection
- [Link to the material](#)

Background knowledge

In geometry, a triangle is a polygon with three edges and three vertices. Based on the length of the sides, the measurement of the angles, there are six kinds of triangles: scalene triangle, right triangle, acute triangle, obtuse triangle, isosceles triangle, and equilateral triangle. However, in this activity we will only look at following three types:

Background knowledge (continued)

- An **equilateral triangle** is a triangle in which all three sides have the same length.
- An **isosceles triangle** is a triangle in which two sides are equal.
- A **scalene triangle** is a triangle in which none of the sides are equal.

Properties of equilateral triangles:

- The angles are congruent and are equal to 60 degrees.
- The perpendicular drawn from any vertex to its opposite side bisects the side in equal lengths.

Properties of isosceles triangles:

- In an isosceles triangle, the two equal sides are called the legs and the angle between them the apex angle. The side opposite the apex angle is called the base.
- The perpendicular from the apex angle bisects the base and the apex angle.
- The base angles are equal.
- The perpendicular drawn from the apex angle divides the triangle into two congruent triangles. This is also the line of symmetry of the triangle.

Properties of scalene triangles:

- The angles are of different measurements.
- It has no parallel or equal sides. Hence there is no line of symmetry.
- A scalene triangle can be an obtuse, acute or right-angled triangle.

Recall that an obtuse triangle is a triangle in which one angle measures more than 90 degrees. An obtuse triangle can be a scalene or isosceles triangle, but will never be equilateral.

Description of the lesson

Step 1

5–30
min

This step includes two options (a and b) depending on the source of the images. These options are not subsequent, the teacher must select either one of them. Once the teacher has selected one of them, the teacher will ask the students to do the last part of this step, section c.

- a. Upload 10 images for each type using the provided link (5 minutes)
- b. Ask the students to look up for a minimum of 10 images for each shape on the internet. The Google Image Search engine can be of great help for this task (30 minutes)
- c. Ask the learners to upload only 2 different pictures for each shape. By this way they will see how well the ML model is doing it by reducing training data.

All images should be uploaded to LearningML, having at the end of this step three groups of shapes (isosceles, scalene and equilateral triangles).

Step 2

15 min

Let the model learn. In this step, LearningML learns to classify images, based on the data provided in Step 1. At this point it is a good opportunity to present the most important concepts of the different triangles. Learners should already be aware of the most significant differences, based on the pictures from Step 1. The most important concepts and keywords should be presented at this stage.

Step 3

30–60
min

Test & discuss. Ask learners to download pictures of triangles from the Internet and upload them to learningML to see the results. As an alternative the provided link can be used to find images. For each uploaded image, learners should answer the following questions:

- Are the results correct? (i.e., is the prediction made by the AI model correct?)
- If the result is not correct, can you think of a reason why the model predicted that shape?
- What properties did you (as a human) use to identify the shapes?
- What properties do you think the algorithm uses to identify the shapes?

Ask the students to upload a square and use it as a test image.

- What is the prediction made by the model?
- Why did the model predict that shape?

Religious buildings

 Subject	 Age	 Duration
History	10-12 years	1h or 2h (depending on whether they use the existing dataset or create it on their own)

Description and learning objectives

This example lesson is an activity that can be used in a history class. In this activity the students will study the characteristics of different places of worship. There are an estimated 10,000 distinct religions worldwide but we will cover the architecture of the most popular ones, Christianity (Orthodox differences in Greece and Russia) – Buddhism – Islam. The activity uses machine learning to strengthen the learning of the learners, as well as other soft skills such as critical thinking, communication and argumentation skills.

Situation before

This class is a generic one and does not require any particular knowledge besides being able to work with a computer.

Educational goals

- Learn the differences in architecture in the different places of worship.
- Be able to identify their main characteristics.
- Promote critical thinking when using AI in the classification the buildings.

Type of instruction

Indirect instruction where students work in groups and discuss and solve problems guided by the teacher.

Material, resources, equipment

- [LearningML tool](#), in particular its image functionality
- Computer with good internet connection
- [Link to the material](#)

Background knowledge

(Adapted from Wikipedia). The word **church** derives from the Greek ekklesia, meaning the called-out ones. Its original meaning is to refer to the body of believers, or the body of Christ. The word church is used to refer to a Christian place of worship by some Christian denominations, including Anglicans and Catholics. Other Christian denominations object to the use of the word “church” to refer to a building, as they argue that this word should be reserved

Background knowledge (continued)

for the body of believers who worship there. Instead, these groups use words such as “Hall” to identify their places of worship or any building in use by them for the purpose of assembly.

A **mosque**, literally meaning “place of prostration”, is a place of worship for followers of Islam. There are strict and detailed requirements in Sunni jurisprudence (fiqh) for a place of worship to be considered a masjid, with places that do not meet these requirements regarded as musallas. There are stringent restrictions on the uses of the area formally demarcated as the mosque (which is often a small portion of the larger complex), and, in the Islamic Sharia law, after an area is formally designated as a mosque, it remains so until the Last Day. Many mosques have elaborate domes, minarets, and prayer halls, in varying styles of architecture. Mosques originated on the Arabian Peninsula, but are now found in all inhabited continents. The mosque serves as a place where Muslims can come together for salat (meaning “prayer”) as well as a centre for information, education, social welfare, and dispute settlement. The imam leads the congregation in prayer.

Temples in Buddhism represent the pure land or pure environment of a Buddha. Traditional Buddhist temples are designed to inspire inner and outer peace. Its structure and architecture varies from region to region. Usually, the temple consists not only of its buildings, but also the surrounding environment. The architectural elements of a Buddhist temple are meant to embody themes and teachings of Buddhism.

Description of the lesson

Step 1

5–30
min

This step includes two options (a and b) depending on the source of the images. These options are not subsequent, the teacher must select either one of them. Once the teacher has selected one of them, the teacher will ask the students to do the last part of this step, section c.

- a. Upload at least 10 images for each type using the provided link (5 minutes)
- b. Ask the students to look up at least 10 images for each type of religious building on the internet. The Google Image Search engine can be of great help for this task (30 minutes)
- c. Ask the learners to upload only 2 different pictures for each shape. By this way they will see how well the ML model is doing it by reducing training data.

All images should be uploaded to LearningML, having at the end of this step three groups of religious buildings (Christian churches, mosques and Buddhist temples).

Step 2

15 min

Let the model learn. In this step, LearningML learns to classify images, based on the data provided in Step 1. At this point it is a good opportunity to present the most important concepts of the different religious buildings. Compare the different examples together with the students.

Step 3

30–60
min

Test & discuss. Ask learners to download pictures of religious buildings from the Internet and upload them to learningML to see the results. As an alternative the provided link can be used to find images. For each uploaded image, learners should answer the following questions:

- Are the results correct? (i.e., is the prediction made by the AI model correct?)
- If the result is not correct, can you think of a reason why the model predicted that shape?
- What properties did you (as a human) use to identify the buildings?
- What properties do you think the algorithm uses to identify the buildings?

Ask the students to upload an image of Saint Basil's church in Moscow and use it as a test image.

- What is the prediction made by the model?
- If the result is not correct, can you think of a reason why the model predicted that shape?
- What kind of building is it? Which reasoning did you use?
- What properties do you think the algorithm uses to identify this building?

Additional tools

In this section we present some tools that can be used in the classroom. These are not lessons on their own.

Music

[Magenta](#) is an open source research project which explores how machine learning can be used in the creative process. A primary goal is to show that machine learning can be used to enable and enhance the creative potential in people. It provides a variety of open-source demos that involve interacting with music.

Using machine learning to generate music and working with computer-generated music can open up a new world of opportunities in teaching music. For instance, changing or improving music that has been composed using machine learning can give students an exciting challenge.

Ready to use tools can be found here: <https://magenta.tensorflow.org/demos>. This project is still very alive and more tools are in development.

Glossary

Algorithm	A set of defined and non-ambiguous, ordered and finite instructions or rules that allow to solve a problem. In our case, computers will use algorithms to process data and perform other tasks such as classification.
Artificial Intelligence (AI)	Discipline from computer science that tries to replicate and develop human intelligence and its implicit processes through computers.
Artificial neural networks	Also called neural networks, are computing systems inspired by the biological neural networks that constitute animal brains. ANNs consist of a collection of connected nodes or neurons, typically organised in layers. Each node receives information, processes it and sends it to the nodes connected to it. The connections are called edges and typically have a weight that is adjusted as learning proceeds.
Bias	Tendencies, inclinations or prejudices toward or against someone or something. AI systems can be biased because of several reasons such as quality of the data, under and oversampling, etc.
Black box	In science, computing, and engineering, a black box is a device, system, or object which produces information without revealing any information about its internal working.
Chatbot	Computer program that simulates and processes human conversation (either written or spoken), allowing humans to interact with digital devices as if they were communicating with a real person.
Classification	Categorising a set of data into classes, examples are spam detection, classification of handwritten digits, fraud detection, ... Classification models predict a category for an unseen datapoint.
Clustering	Organising data into subgroups with similar attributes or elements
Data	Group of datum. Data describe empirical facts, events and entities. Data may represent abstract ideas or concrete measurements.
Datum	Minimum unit of data. Strictly speaking datum is the singular form.

Labelled data	Set of samples that have been tagged with one or more labels. Labelled data is used in supervised learning.
Learning phase	This is the phase in which the model learns how to map input to output based on the training data. The output of this phase is an AI model that can be used to predict an output for new unseen input.
LearningML	Open-source web-based tool designed to teach AI to kids. It simplifies the process and it provides an user-friendly interface to train and test a classification model. It allows image, text or datasets to be used as data.
Machine Learning	Subfield of AI in which algorithms learn from data. Abbreviation – ML
Model	The function that maps input to output. In machine learning this function is learned using data, also called the training data. In machine learning, the learned model is a black box.
Neural network	See Artificial neural networks
Pattern	A sequence of repeating objects, shapes or numbers
Pattern recognition	Process of recognizing patterns and regularities in data
Qualitative attribute or variable	Refers to variables that cannot be easily measured, such as smells, tastes, textures, and attractiveness
Quantitative attribute or variable	Refers to variables that can be described by numbers, or a set of them, i.e. variables that can be measured objectively.
Reinforcement learning	Subfield of AI in which algorithms learn from interaction
Regression	Modelling the relationship between independent variables and a dependent (target) variable where the target variable is a continuous variable such as temperature, age, salary, price
Supervised learning	Subfield of machine learning. Supervised machine learning algorithms are trained using labelled data. It is used in classification and regression.
Training phase	Synonymous with the learning phase
Unsupervised learning	Subfield of machine learning. Unsupervised machine learning algorithms are trained using unlabelled data. It is used in clustering.

Related bibliography

LearningML tool

- Rodríguez, Juan & Moreno-León, Jesús & Román-González, Marcos & Robles, Gregorio. (2021). Evaluation of an Online Intervention to Teach Artificial Intelligence With LearningML to 10–16–Year–Old Students.
 - LearningML frontpage: <https://web.learningml.org/>
-

Bias in AI

- Rosenfeld, Hannah, Danni Yu, and Abhishek Gupta. 2023. “Ethical AI requires future-ready, inclusive education system | World Economic Forum.” The World Economic Forum, June 5, 2023. <https://www.weforum.org/agenda/2023/06/ethical-ai-future-ready-inclusive-education-system/>
 - Blackman, Reid. 2020. “A Practical Guide to Building Ethical AI.” Harvard Business Review, October 15, 2020. <https://hbr.org/2020/10/a-practical-guide-to-building-ethical-ai>
 - Burton, Emanuelle, Judy Goldsmith, Sven Koenig, Benjamin Kuipers, Nicholas Mattei, and Toby Walsh. 2017. “Ethical Considerations in Artificial Intelligence Courses.” AI Magazine 38(2) (July): pp22–34. <https://doi.org/10.1609/aimag.v38i2.2731>
 - Warikoo, Natasha, Stacey Sinclair, Jessica Fei, and Drew Jacoby-Senghor. n.d. “Examining Racial Bias in Education: A New Approach.” Educational Researcher 45 (9). <https://doi.org/10.3102/0013189X16683408>
 - Zittleman, Karen, and David Sadker. 2022. “Gender Bias in Teacher Education Texts: New (and Old) Lessons.” Journal of Teacher Education 53, no. 2 (March): 168–180. <https://doi.org/10.1177/0022487102053002008>
 - Hao, Karen. 2019. “Facebook’s ad-serving algorithm discriminates by gender and race.” MIT Technology Review. <https://www.technologyreview.com/2019/04/05/1175/facebook-algorithm-discriminates-ai-bias/>
-

AI in non-formal education

- Norqvist, Lars, and Eva Leffler. 2017. “Learning in non-formal education: Is it “youthful” for youth in action?” Springer link 63 (March): pp235–256. <https://doi.org/10.1007/s11159-017-9631-8>
 - “Visa SVP Carolina Barcenás: “Educating the general public on AI is key to ethical use and smart” 2019. Medium. <https://medium.com/authority-magazine/visa-svp-carolina-barcenas-educating-the-general-public-on-ai-is-key-to-ethical-use-and-smart-ea080f37966c>
-

AI in formal education

- Willis, Robert J., and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87, no. 5 (October). <https://doi.org/10.1086/260821>
- Glewwe, Paul. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *Journal of Economic Literature* 40, no. 2 (June): pp436-482. 10.1257/002205102320161258
- Gururaj, Tejasri. 2023. "10 examples of how artificial intelligence is improving education." *Interesting Engineering*. <https://interestingengineering.com/lists/examples-how-artificial-intelligence-improving-education>
- Morron, Mirriam. 2023. "The Impact of AI on Learning and Creativity: A Critical Look at the Future." LinkedIn. <https://www.linkedin.com/pulse/impact-ai-learning-creativity-critical-look-future-miriam-morron/>
- "Role Models in AI: Stephanie Lampkin | by AI4ALL Team | AI4ALL." 2019. Medium. <https://medium.com/ai4allorg/role-models-in-ai-stephanie-lampkin-f50e6885afa4>
- "Female role models changing the field of Artificial Intelligence in Norway - NORA." n.d. Norwegian Artificial Intelligence Research Consortium. Accessed June 16, 2023. <https://www.nora.ai/news-and-events/news/30-women-in-norway-changing-the-field-of-artificia.html>

Best practises for AI

- Zowghi, Didar, and Francesca da Rimini. 2023. *Responsible AI: Best Practices for Creating Trustworthy AI Systems*. N.p.: Pearson Addison Wesley. <https://doi.org/10.48550/arXiv.2305.12728>
- Xu, Feiyu, Hans Uszkoreit, Yangzhou Du, Wei Fan, Dongyan Zhao, and Jun Zhu. 2019. "Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges." In *Lecture Notes in Computer Science*, pp563-574. Vol. 11839. N.p.: Springer. https://doi.org/10.1007/978-3-030-32236-6_51

Games with AI

- "The tag game | Activities." n.d. Scouts. Accessed June 16, 2023. <https://www.scouts.org.uk/activities/the-tag-game/>
- "Describe and draw | Activities." n.d. Scouts. Accessed June 16, 2023. <https://www.scouts.org.uk/activities/draw-the-description/>

Glossaries

- UNESCO Institute for Statistics, [Glossary](#)
 - College of the Environment, University of Washington, [Diversity, Equity and Inclusion Glossary](#)
 - Harvard Human Resources, [Glossary of Diversity, Inclusion and Belonging Terms](#)
-

Fostering Artificial Intelligence at Schools

Guideline for teachers in formal education

