Course 1: Generalities about AI

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

2025-10-02

Course 1: Generalities about AI

Course 1: Generalities about AI

Welcome to the first lesson of the AI course! I'll a AI do the introduction to this course. Demo on what modern AI system are able to achieve (generative AI based on large language models: chatGPT and Dall-E)

What is AI?

- **Intelligence**: ability to **extract knowledge** from observations
- This knowledge is used to **solve tasks in different contexts and environments** (automation)

Old way: Memorize

- Human experts code the machines
- Goods: we know what we are doing.
- Bads: requires **explicit** solutions (not available for some problems).

Modern way: Learning

- Let machines teach themselves how to solve a problem (**implicit**).
- Goods: universally applicable
- Bads: lack of understandability/robustness.
- Requires **training**.

Memorizing (explicit) **vs** Learning (implicit)

intelligence: ability to process information to inform future decisions. AI: focuses on building artificial algo that can do the same thing. Historically, this has been done by explicitly telling the machines (=coding algo) how to extract the required knowledge. This had the advantage of exactly knowing how the algo was working but the strong limitation of requiring explicit solution, not available for more complex tasks. The modern way to do AI, what we call ML is to teach machines how to do is without being explicitly programmed.

Global overview…

What is AI?

- **Intelligence**: ability to **extract knowledge** from observations
- This knowledge is used to **solve tasks in different contexts and environments** (automation)

Old way: Memorize

- Human experts code the machines
- Goods: we know what we are doing.
- Bads: requires **explicit** solutions (not available for some problems).

Modern way: Learning

- Let machines teach themselves how to solve a problem (**implicit**).
- Goods: universally applicable
- Bads: lack of understandability/robustness.
- Requires **training**.

Memorizing (explicit) **vs** Learning (implicit)

intelligence: ability to process information to inform future decisions. AI: focuses on building artificial algo that can do the same thing. Historically, this has been done by explicitly telling the machines (=coding algo) how to extract the required knowledge. This had the advantage of exactly knowing how the algo was working but the strong limitation of requiring explicit solution, not available for more complex tasks. The modern way to do AI, what we call ML is to teach machines how to do is without being explicitly programmed.

Global overview…

What is AI?

- **Intelligence**: ability to **extract knowledge** from observations
- This knowledge is used to **solve tasks in different contexts and environments** (automation)

Old way: Memorize

- Human experts code the machines
- Goods: we know what we are doing.
- Bads: requires **explicit** solutions (not available for some problems).

Modern way: Learning

- Let machines teach themselves how to solve a problem (**implicit**).
- Goods: universally applicable
- Bads: lack of understandability/robustness.
- Requires **training**.

Memorizing (explicit) **vs** Learning (implicit)

Global overview…
What is AI?
Intelligence: ability to extract knowledge from observations
This knowledge is used to solve tasks in different contexts and environments (automation)
Old way: Memorize
Human experts code the machines
Goods: we know what we are doing.
Bads: requires explicit solutions (not available for some problems).
Modern way: Learning
Let machines teach themselves how to solve a problem (implicit).
Goods: universally applicable
Bads: lack of understandability/robustness.
Requires training.
Memorizing (explicit) vs Learning (implicit)

intelligence: ability to process information to inform future decisions. AI: focuses on building artificial algo that can do the same thing. Historically, this has been done by explicitly telling the machines (=coding algo) how to extract the required knowledge. This had the advantage of exactly knowing how the algo was working but the strong limitation of requiring explicit solution, not available for more complex tasks. The modern way to do AI, what we call ML is to teach machines how to do is without being explicitly programmed.

Global overview…

What is AI?

- **Intelligence**: ability to **extract knowledge** from observations
- This knowledge is used to **solve tasks in different contexts and environments** (automation)

Old way: Memorize

- Human experts code the machines
- Goods: we know what we are doing.
- Bads: requires **explicit** solutions (not available for some problems).

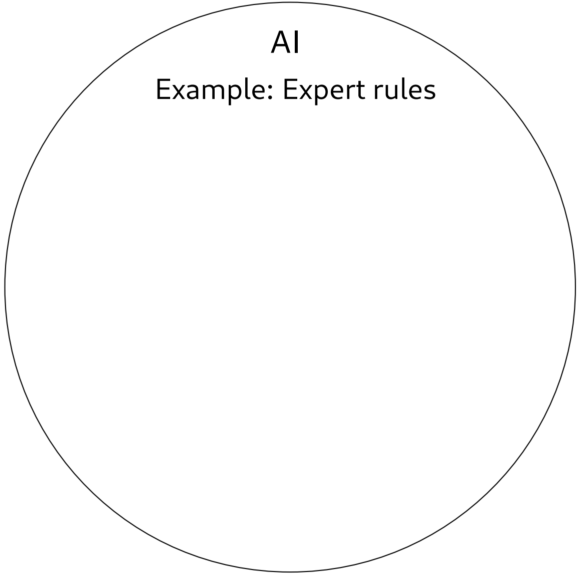
Modern way: Learning

- Let machines teach themselves how to solve a problem (**implicit**).
- Goods: universally applicable
- Bads: lack of understandability/robustness.
- Requires **training**.

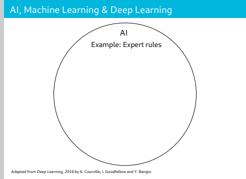
Memorizing (explicit) **vs** Learning (implicit)

Global overview…
What is AI?
Intelligence: ability to extract knowledge from observations
This knowledge is used to solve tasks in different contexts and environments (automation)
Old way: Memorize
Human experts code the machines
Goods: we know what we are doing.
Bads: requires explicit solutions (not available for some problems).
Modern way: Learning
Let machines teach themselves how to solve a problem (implicit).
Goods: universally applicable
Bads: lack of understandability/robustness.
Requires training.
Memorizing (explicit) vs Learning (implicit)

intelligence: ability to process information to inform future decisions. AI: focuses on building artificial algo that can do the same thing. Historically, this has been done by explicitly telling the machines (=coding algo) how to extract the required knowledge. This had the advantage of exactly knowing how the algo was working but the strong limitation of requiring explicit solution, not available for more complex tasks. The modern way to do AI, what we call ML is to teach machines how to do is without being explicitly programmed.

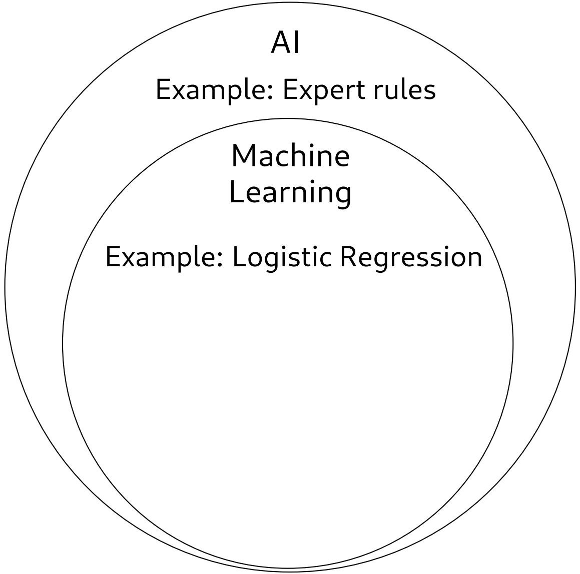


Adapted from *Deep Learning, 2016* by A. Courville, I. Goodfellow and Y. Bengio

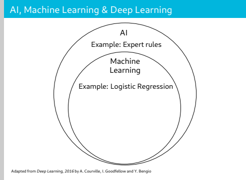


Deep Learning is a particular case of Machine Learning.

AI, Machine Learning & Deep Learning

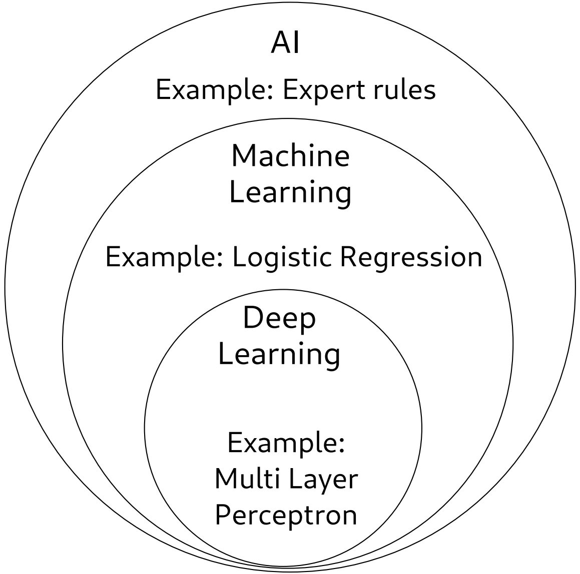


Adapted from *Deep Learning, 2016* by A. Courville, I. Goodfellow and Y. Bengio

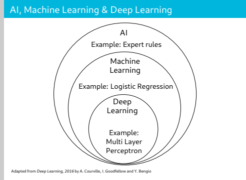


Deep Learning is a particular case of Machine Learning.

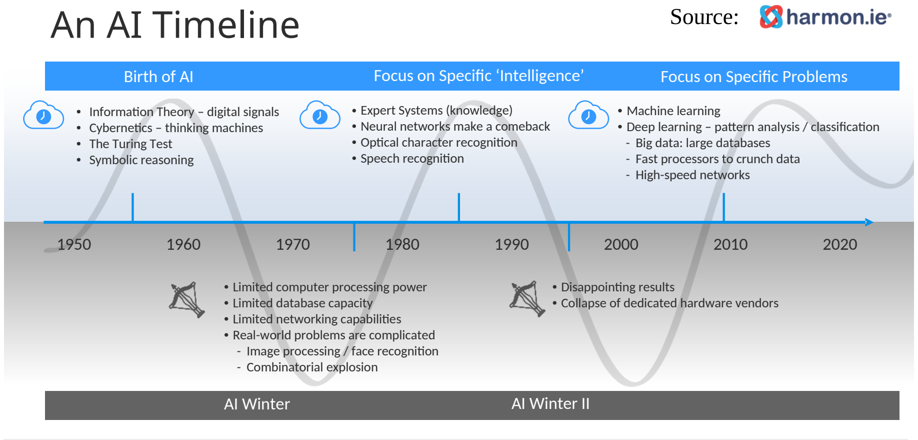
AI, Machine Learning & Deep Learning

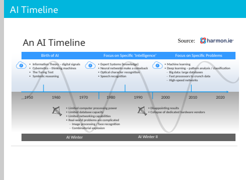


Adapted from *Deep Learning, 2016* by A. Courville, I. Goodfellow and Y. Bengio



Deep Learning is a particular case of Machine Learning.



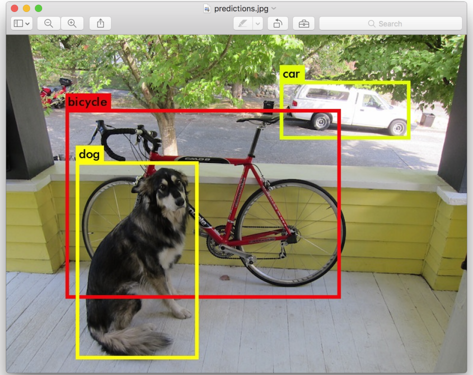


Here, we want to show that field of AI has gone through a number of "winters". This timeline shows that the basic building blocks of DL where there for decades, and algorithms to train them as well! For instance the SGD was proposed in 1952, BackPropagation in 1986. The reasons explaining the success of modern AI which is mostly based on Deep Learning are to be found in other aspects.

Traditional application domains of AI

Vision

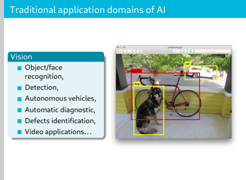
- Object/face recognition,
- Detection,
- Autonomous vehicles,
- Automatic diagnostic,
- Defects identification,
- Video applications...



2025-10-02

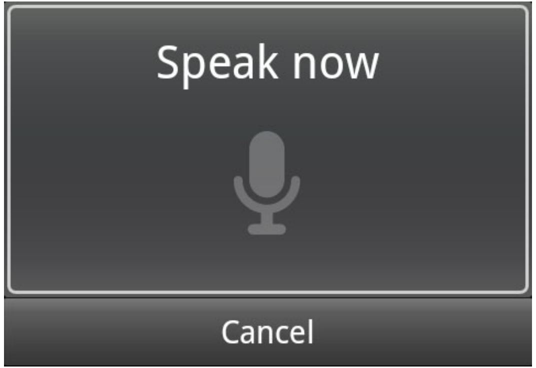
Course 1: Generalities about AI

└Traditional application domains of AI

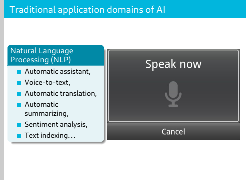


Natural Language Processing (NLP)

- Automatic assistant,
- Voice-to-text,
- Automatic translation,
- Automatic summarizing,
- Sentiment analysis,
- Text indexing…

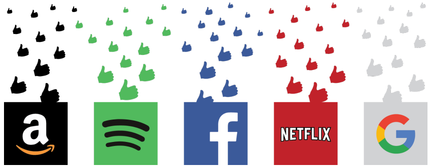


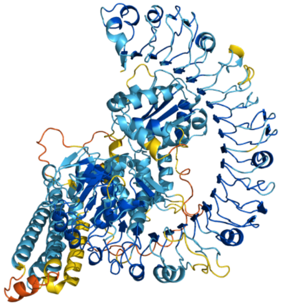
2025-10-02
Course 1: Generalities about AI
└Traditional application domains of AI



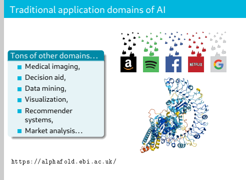
Tons of other domains...

- Medical imaging,
- Decision aid,
- Data mining,
- Visualization,
- Recommender systems,
- Market analysis...



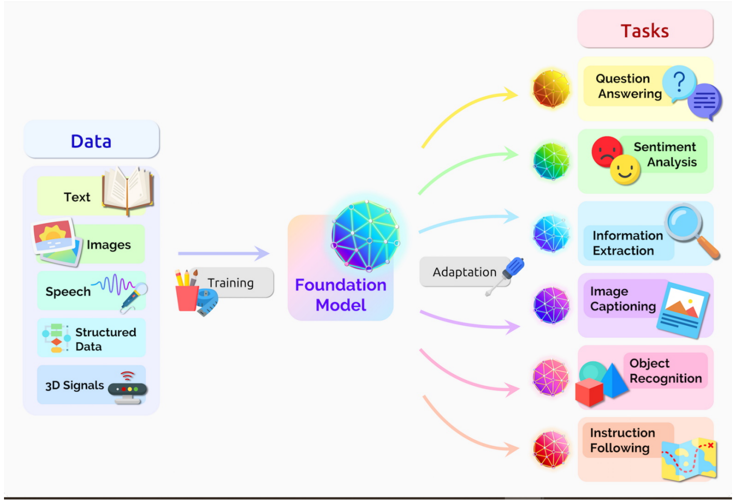


https://alphafold.ebi.ac.uk/



There is probably plenty more application domains. As an example of a breakthrough it is worth mentioning AlphaFold that has revolutioned the field of bioinformatics. The AlphaFold network directly predicts the 3D coordinates of all heavy atoms for a given protein using the primary amino acid sequence. It is an AI system that contains many "ingredients" of modern DL approaches: attention mechanisms, self-distillation,...).

Foundation models: a game changer

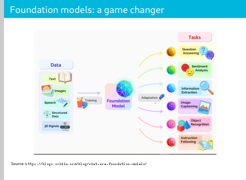


Source: `https://blogs.nvidia.com/blog/what-are-foundation-models/`

2025-10-02

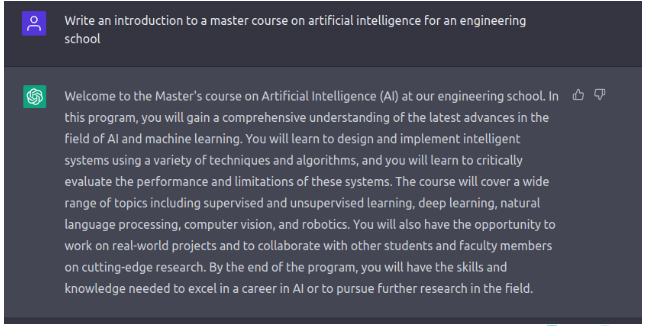
Course 1: Generalities about AI

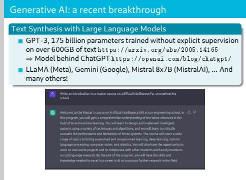
└ Foundation models: a game changer



Text Synthesis with Large Language Models

- GPT-3, 175 billion parameters trained without explicit supervision on over 600GB of text `https://arxiv.org/abs/2005.14165`
  ⇒ Model behind ChatGPT `https://openai.com/blog/chatgpt/`
- LLaMA (Meta), Gemini (Google), Mistral 8x7B (MistralAI), ... And many others!





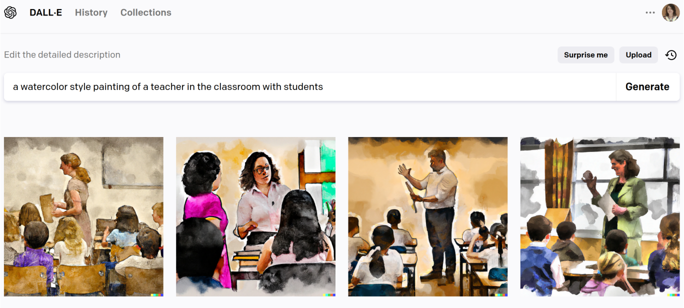
trained on CommonCrawl, Webtext, Books, Wikipedia

Generative AI: a recent breakthrough

Image Synthesis

- Stable Diffusion Models (`https://arxiv.org/pdf/2403.03206`)
- DALL-E (`https://labs.openai.com/`)
- Midjourney (`https://docs.midjourney.com/`)

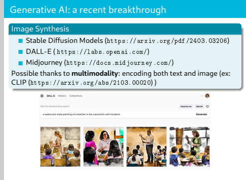
Possible thanks to **multimodality**: encoding both text and image (ex: CLIP (`https://arxiv.org/abs/2103.00020`))



2025-10-02

Course 1: Generalities about AI

└─Generative AI: a recent breakthrough



Both use diffusion models to encode and Vision Transformer

Generative AI: a recent breakthrough

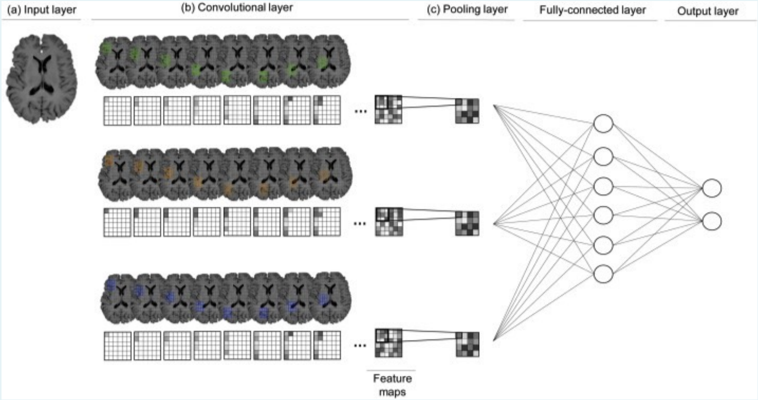
Many other modalities...

- Speech (Many actors, ex: Kyutai)
- Music (StableAudio, Udio, Suno, ...)
- Video (StableVideo, Sora (OpenAI), ...)
- ...
- **Multimodal** models are arising

Generative AI: a recent breakthrough
Many other modalities...
Speech (Many actors, ex: Kyutai)
Music (StableAudio, Udio, Suno, ...)
Video (StableVideo, Sora (OpenAI), ...)
...
Multimodal models are arising

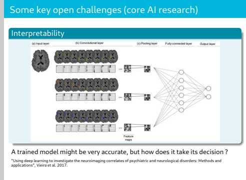
Some key open challenges (core AI research)

Interpretability



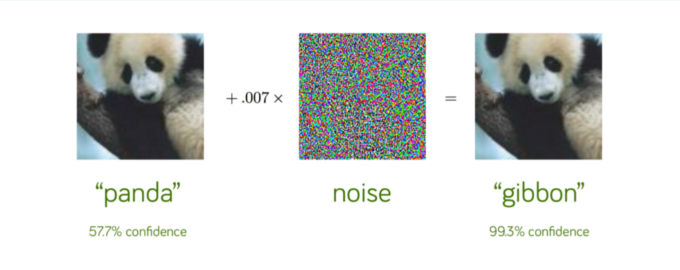
A trained model might be very accurate, but how does it take its decision ?

"Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications", Vieira et al. 2017.



Some key open challenges (core AI research)

Learning what should be learned (robustness / adversarial attacks)



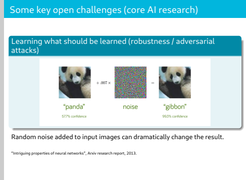
Random noise added to input images can dramatically change the result.

"Intriguing properties of neural networks", Arxiv research report, 2013.

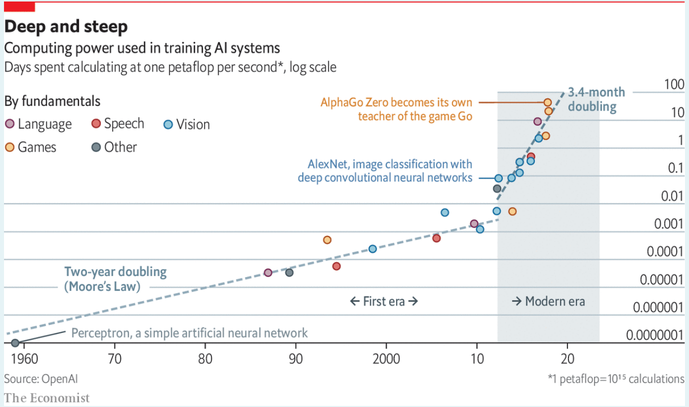
2025-10-02

Course 1: Generalities about AI

└─Some key open challenges (core AI research)



Computational and memory footprints

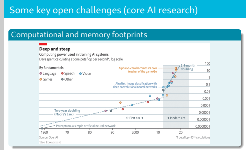


Training a large algorithm: thousands to millions of parameters using Gigabytes of data.

2025-10-02

Course 1: Generalities about AI

└─Some key open challenges (core AI research)



Training a large algorithm: thousands to millions of parameters using Gigabytes of data.

Let's dive into details!

Machine Learning

Examples

- Learning to play chess through playing games,
- Learning to recognize dogs and cats in images from annotated examples . . .

Machine learning

- Supervised: Learning from **inputs** and **human annotations**
- Unsupervised: Learning from **inputs** only (patterns)
- Self-supervised: Learning by reconstructing the **inputs** from distorted versions

Generalization

- Generalization refers to the ability to infer **good decisions or representations** from **examples**.
- Goal: transfer knowledge from a dataset to another

Machine Learning
Examples
Learning to play chess through playing games,
Learning to recognize dogs and cats in images from annotated examples . . .
Machine learning
Supervised: Learning from inputs and human annotations
Unsupervised: Learning from inputs only (patterns)
Self-supervised: Learning by reconstructing the inputs from distorted versions
Generalization
Generalization refers to the ability to infer good decisions or representations from examples.
Goal: transfer knowledge from a dataset to another

The most common framework in ML is Supervised Learning: we learn a function from inputs/examples/data and their outputs/labels. In this setting, labels are given by humans (ex: cat pictures, with the annotation "cat"). A second framework is Unsupervised Learning, where the function is learned from the data only. In general, it consists of finding repeated patterns or structure in data. A third framework, very popular nowadays, consist of leveraging the data itself to compute the annotations, and remove the human in the loop. It can be very powerful if a relevant way to compute annotations is found. A particular example, used for most Large Language Models: remove one word in a sentence, and train the network to find this word.

One key challenge in ML is Generalization. It can be seen as the ability to generalize on unseen data (different from data (inputs/labels) the model has been trained on). In this course we will only focus on different techniques that are used to build systems that can learn on data, and generalize on unseen data.

Machine Learning

Examples

- Learning to play chess through playing games,
- Learning to recognize dogs and cats in images from annotated examples . . .

Machine learning

- Supervised: Learning from **inputs** and **human annotations**
- Unsupervised: Learning from **inputs** only (patterns)
- Self-supervised: Learning by reconstructing the **inputs** from distorted versions

Generalization

- Generalization refers to the ability to infer **good decisions or representations** from **examples**.
- Goal: transfer knowledge from a dataset to another

Machine Learning
Examples
Learning to play chess through playing games,
Learning to recognize dogs and cats in images from annotated examples . . .
Machine learning
Supervised: Learning from inputs and human annotations
Unsupervised: Learning from inputs only (patterns)
Self-supervised: Learning by reconstructing the inputs from distorted versions
Generalization
Generalization refers to the ability to infer good decisions or representations from examples.
Goal: transfer knowledge from a dataset to another

The most common framework in ML is Supervised Learning: we learn a function from inputs/examples/data and their outputs/labels. In this setting, labels are given by humans (ex: cat pictures, with the annotation "cat"). A second framework is Unsupervised Learning, where the function is learned from the data only. In general, it consists of finding repeated patterns or structure in data. A third framework, very popular nowadays, consist of leveraging the data itself to compute the annotations, and remove the human in the loop. It can be very powerful if a relevant way to compute annotations is found. A particular example, used for most Large Language Models: remove one word in a sentence, and train the network to find this word.

One key challenge in ML is Generalization. It can be seen as the ability to generalize on unseen data (different from data (inputs/labels) the model has been trained on). In this course we will only focus on different techniques that are used to build systems that can learn on data, and generalize on unseen data.

Machine Learning

Examples

- Learning to play chess through playing games,
- Learning to recognize dogs and cats in images from annotated examples ...

Machine learning

- Supervised: Learning from **inputs** and **human annotations**
- Unsupervised: Learning from **inputs** only (patterns)
- Self-supervised: Learning by reconstructing the **inputs** from distorted versions

Generalization

- Generalization refers to the ability to infer **good decisions or representations** from **examples**.
- Goal: transfer knowledge from a dataset to another

Machine Learning
Examples
- Learning to play chess through playing games,
- Learning to recognize dogs and cats in images from annotated examples ...
Machine learning
- Supervised: Learning from **inputs** and **human annotations**
- Unsupervised: Learning from **inputs** only (patterns)
- Self-supervised: Learning by reconstructing the **inputs** from distorted versions
Generalization
- Generalization refers to the ability to infer **good decisions or representations** from **examples**.
- Goal: transfer knowledge from a dataset to another

The most common framework in ML is Supervised Learning: we learn a function from inputs/examples/data and their outputs/labels. In this setting, labels are given by humans (ex: cat pictures, with the annotation "cat"). A second framework is Unsupervised Learning, where the function is learned from the data only. In general, it consists of finding repeated patterns or structure in data. A third framework, very popular nowadays, consist of leveraging the data itself to compute the annotations, and remove the human in the loop. It can be very powerful if a relevant way to compute annotations is found. A particular example, used for most Large Language Models: remove one word in a sentence, and train the network to find this word.

One key challenge in ML is Generalization. It can be seen as the ability to generalize on unseen data (different from data (inputs/labels) the model has been trained on). In this course we will only focus on different techniques that are used to build systems that can learn on data, and generalize on unseen data.

Input/output

- Goal: infer a function of parameters $\mathbf{W}$ from an input (often tensor) space to an output (often tensor) space, $\mathbf{y} = f(\mathbf{x}, \mathbf{W})$.
- Example: input can be an image, output a vector where the largest value indicate the category the image belongs to.

Error/Loss

- Loss $\mathcal{L}$: nonnegative measure of the discrepancy between expected output $\hat{\mathbf{y}}$ and obtained output $\mathbf{y}$.
- Example: output should be $[0, 1]$ but is $[0.2, 0.8]$.

Parameters

- $f(\cdot, \mathbf{W})$ contains **parameters W** to be trained,
- In most cases, an ideal $f(\cdot, \mathbf{W})$ exists but is **hard to find in practice**,
- Learning is a **regression ill-posed** problem.

Global formalism
Input/output
Goal: infer a function of parameters W from an input (often tensor) space to an output (often tensor) space, y = f(x, W).
Example: input can be an image, output a vector where the largest value indicate the category the image belongs to.
Error/Loss
Loss L: nonnegative measure of the discrepancy between expected output ŷ and obtained output y.
Example: output should be [0, 1] but is [0.2, 0.8].
Parameters
f(·, W) contains parameters W to be trained,
In most cases, an ideal f(·, W) exists but is hard to find in practice,
Learning is a regression ill-posed problem.

Loss: it's a way to tell the model when it is wrong and train the model accordingly. The model contains parameters (model weights and bias) and usually, given a task, an optimal set of parameters exist but again finding it is ill posed problem (many solutions exist)

Global formalism

Input/output

- Goal: infer a function of parameters $\mathbf{W}$ from an input (often tensor) space to an output (often tensor) space, $\mathbf{y} = f(\mathbf{x}, \mathbf{W})$.
- Example: input can be an image, output a vector where the largest value indicate the category the image belongs to.

Error/Loss

- Loss $\mathcal{L}$: nonnegative measure of the discrepancy between expected output $\hat{\mathbf{y}}$ and obtained output $\mathbf{y}$.
- Example: output should be $[0, 1]$ but is $[0.2, 0.8]$.

Parameters

- $f(\cdot, \mathbf{W})$ contains **parameters $\mathbf{W}$** to be trained,
- In most cases, an ideal $f(\cdot, \mathbf{W})$ exists but is **hard to find in practice**,
- Learning is a **regression ill-posed** problem.

Global formalism
Input/output
Goal: infer a function of parameters W from an input (often tensor) space to an output (often tensor) space, $\mathbf{y} = f(\mathbf{x}, \mathbf{W})$.
Example: input can be an image, output a vector where the largest value indicate the category the image belongs to.
Error/Loss
Loss $\mathcal{L}$: nonnegative measure of the discrepancy between expected output $\hat{\mathbf{y}}$ and obtained output $\mathbf{y}$.
Example: output should be [0, 1] but is [0.2, 0.8].
Parameters
$f(\cdot, \mathbf{W})$ contains parameters W to be trained,
In most cases, an ideal $f(\cdot, \mathbf{W})$ exists but is hard to find in practice,
Learning is a regression ill-posed problem.

Loss: it's a way to tell the model when it is wrong and train the model accordingly. The model contains parameters (model weights and bias) and usually, given a task, an optimal set of parameters exist but again finding it is ill posed problem (many solutions exist)

Global formalism

Input/output

- Goal: infer a function of parameters $\mathbf{W}$ from an input (often tensor) space to an output (often tensor) space, $\mathbf{y} = f(\mathbf{x}, \mathbf{W})$.
- Example: input can be an image, output a vector where the largest value indicate the category the image belongs to.

Error/Loss

- Loss $\mathcal{L}$: nonnegative measure of the discrepancy between expected output $\hat{\mathbf{y}}$ and obtained output $\mathbf{y}$.
- Example: output should be $[0, 1]$ but is $[0.2, 0.8]$.

Parameters

- $f(\cdot, \mathbf{W})$ contains **parameters W** to be trained,
- In most cases, an ideal $f(\cdot, \mathbf{W})$ exists but is **hard to find in practice**,
- Learning is a **regression ill-posed** problem.

Global formalism
Input/output
Goal: infer a function of parameters W from an input (often tensor) space to an output (often tensor) space, y = f(x, W).
Example: input can be an image, output a vector where the largest value indicate the category the image belongs to.
Error/Loss
Loss ℒ: nonnegative measure of the discrepancy between expected output ŷ and obtained output y.
Example: output should be [0, 1] but is [0.2, 0.8].
Parameters
f(·, W) contains parameters W to be trained,
In most cases, an ideal f(·, W) exists but is hard to find in practice,
Learning is a regression ill-posed problem.

Loss: it's a way to tell the model when it is wrong and train the model accordingly. The model contains parameters (model weights and bias) and usually, given a task, an optimal set of parameters exist but again finding it is ill posed problem (many solutions exist)

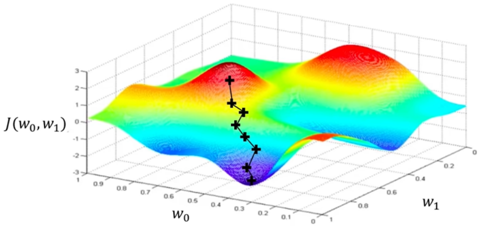
Global formalism

- Loss: $J(\mathbf{W}) = \sum_i \mathcal{L}(f(\mathbf{x}^{(i)}, \mathbf{W}), \mathbf{y}^{(i)}), i = examples$
- Model parameters: $\mathbf{W}^* = argmin(J(\mathbf{W}))$

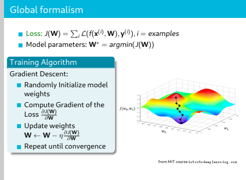
Training Algorithm

Gradient Descent:

- Randomly Initialize model weights
- Compute Gradient of the Loss $\frac{\partial J(\mathbf{W})}{\partial \mathbf{W}}$
- Update weights $\mathbf{W} \leftarrow \mathbf{W} - \eta \frac{\partial J(\mathbf{W})}{\partial \mathbf{W}}$
- Repeat until convergence

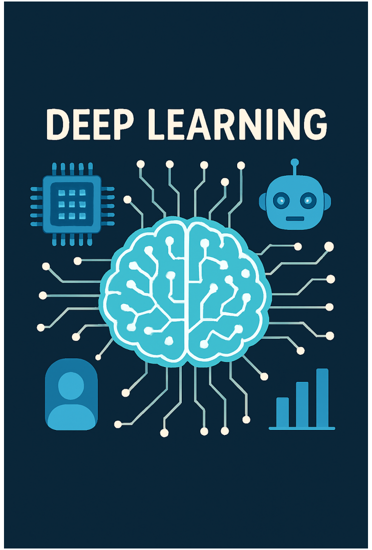


from MIT course `introtodeeplearning.com`



The total loss J (Empirical Risk, Objective function) is the average of Loss for each input/example and the optimal model parameters are those that minimize it. But how to find them? In other words, how to train the model? Here is a simplified description of the training algorithm at the base of modern DL, gradient descent. Repeat until reaching a local minimum (as illustrated in the figure for a simple example where we have only 2 parameters. We'll see that the function becomes much more complicated for millions of parameters -modern neural networks.)

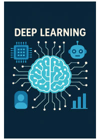
Deep Learning



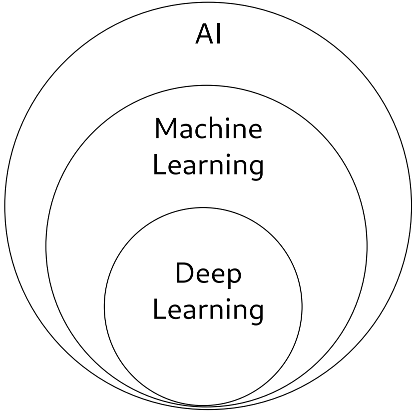
2025-10-02

Course 1: Generalities about AI

└Deep Learning



Deep Learning: A Particular Case of Machine Learning

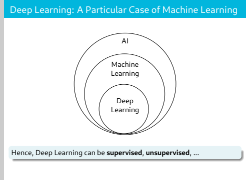


Hence, Deep Learning can be **supervised**, **unsupervised**, ...

2025-10-02

Course 1: Generalities about AI

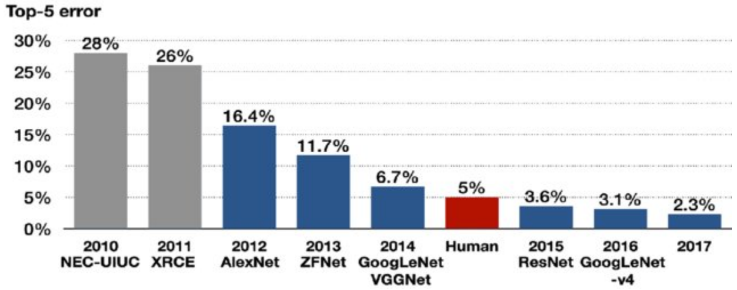
└─Deep Learning: A Particular Case of Machine Learning



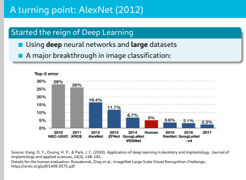
A turning point: AlexNet (2012)

Started the reign of Deep Learning

- Using **deep** neural networks and **large** datasets
- A major breakthrough in image classification:



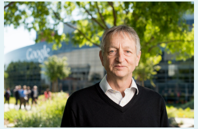
Source: Kang, D. Y., Duong, H. P., & Park, J. C. (2020). Application of deep learning in dentistry and implantology. Journal of implantology and applied sciences, 24(3), 148-181.
Details for the human evaluation: Russakovsk, Dieg *et al.*. ImageNet Large Scale Visual Recognition Challenge, https://arxiv.org/pdf/1409.0575.pdf



The landscape of Machine Learning changed in 2012: Deep Neural Networks, a technique used in a minority of cases until then, suddenly won the Image Classification contest on ImageNet, a standard image classification dataset. From this point, Deep Neural Networks became mainstream, and the performance of Deep Neural Network models skyrocketed.

The great elders of Deep Learning (Turing Prize 2018)

Geoffrey Hinton



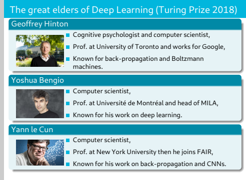
- Cognitive psychologist and computer scientist,
- Prof. at University of Toronto and works for Google,
- Known for back-propagation and Boltzmann machines.

Yoshua Bengio

- Computer scientist,
- Prof. at Université de Montréal and head of MILA,
- Known for his work on deep learning.

Yann le Cun

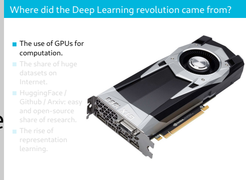
- Computer scientist,
- Prof. at New York University then he joins FAIR,
- Known for his work on back-propagation and CNNs.


The great elders of Deep Learning (Turing Prize 2018)
Geoffrey Hinton
Cognitive psychologist and computer scientist,
Prof. at University of Toronto and works for Google,
Known for back-propagation and Boltzmann machines.
Yoshua Bengio
Computer scientist,
Prof. at Université de Montréal and head of MILA,
Known for his work on deep learning.
Yann le Cun
Computer scientist,
Prof. at New York University then he joins FAIR,
Known for his work on back-propagation and CNNs.

The reason why we mention them is because their work has mostly enabled to get out of the last two AI winters. It is also worth noting that they won the Turing prize in 2018, which is the highest distinction in computer science.

Where did the Deep Learning revolution came from?

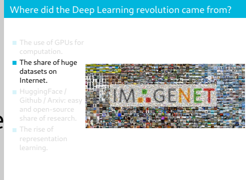
- The use of GPUs for computation.
- The share of huge datasets on Internet.
- HuggingFace / Github / Arxiv: easy and open-source share of research.
- The rise of representation learning.



DL training algorithms are highly parallelizable, and can benefit from modern GPU

Where did the Deep Learning revolution came from?

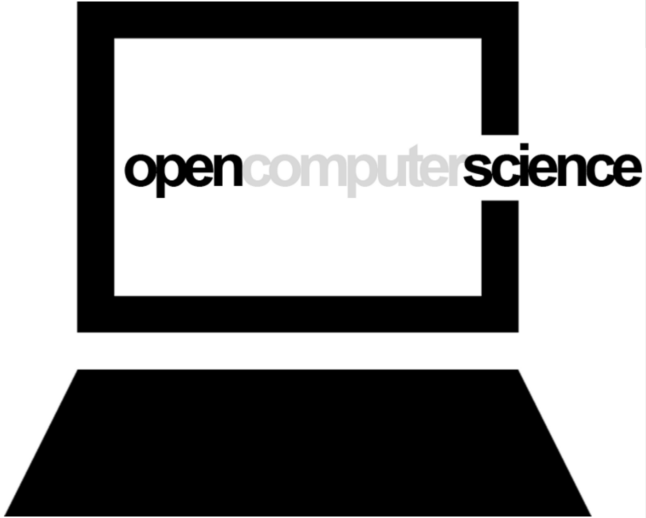
- The use of GPUs for computation.
- The share of huge datasets on Internet.
- HuggingFace / Github / Arxiv: easy and open-source share of research.
- The rise of representation learning.



We are in a data pervasive era, massive amount of digitale data is available and has been shared. This benefits DL algo that are extremely data hungry

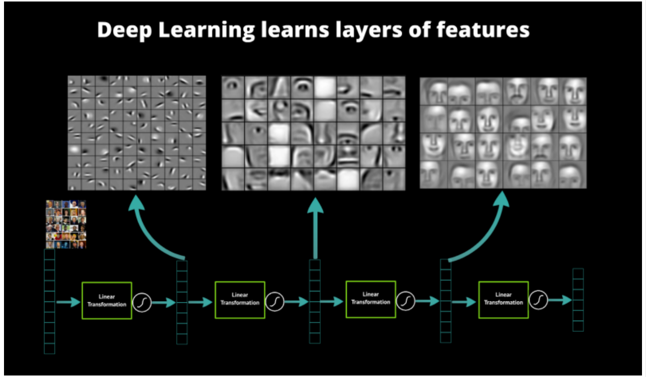
Where did the Deep Learning revolution came from?

- The use of GPUs for computation.
- The share of huge datasets on Internet.
- HuggingFace / Github / Arxiv: easy and open-source share of research.
- The rise of representation learning.

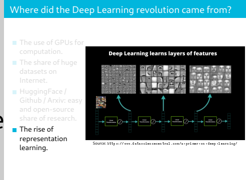


The availability of open source toolboxes such as pytorch and tensorflow and the practice of sharing research content and tools makes implementing and training DL models much easier, and you will find out in this course

- The use of GPUs for computation.
- The share of huge datasets on Internet.
- HuggingFace / Github / Arxiv: easy and open-source share of research.
- The rise of representation learning.

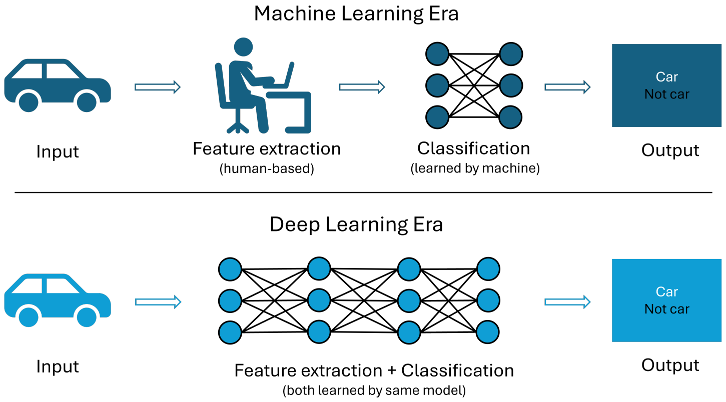


Source: `https://www.datasciencecentral.com/a-primer-on-deep-learning/`



A few more explanation on Representation learning. As we will see in the next lessons, finding a good representation of the data is a very difficult thing. Deep Learning has enabled to search / decompose the data automatically to find such representations. In traditional pattern recognition, features were extracted using a priori expert knowledge on the data (eg looking for orange round objects to recognize oranges, by filter the colour orange, and extracting round objects). In mainstream pattern recognition, mathematical functions (eg wavelets) are used to automatically decompose images in sets of features that are more abstract, but still expertly chosen depending on the data. In Deep Learning: the features are trained end to end.

Traditional Deep Learning (before 2020)

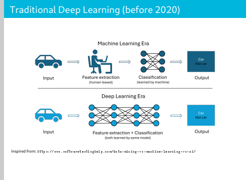


Inspired from: `https://www.softwaretestinghelp.com/data-mining-vs-machine-learning-vs-ai/`

2025-10-02

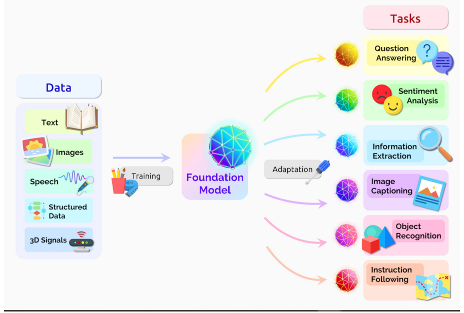
Course 1: Generalities about AI

└Traditional Deep Learning (before 2020)



Deep Learning today - Foundation models (1/3)

- Model trained on an Internet scale dataset
- Self-Supervised training: pretext tasks
- Generalization is not a problem anymore! All is about **particularization**



Source: `https://blogs.nvidia.com/blog/what-are-foundation-models/`

2025-10-02

Course 1: Generalities about AI

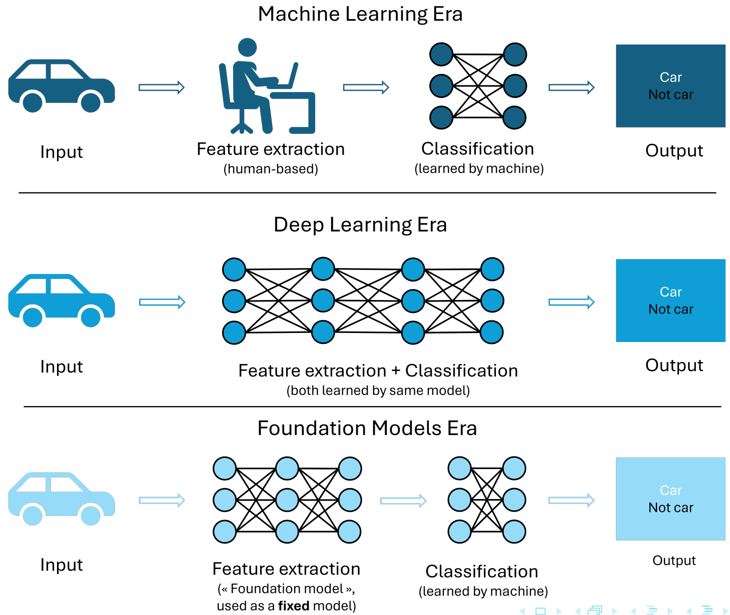
└ Deep Learning today - Foundation models (1/3)

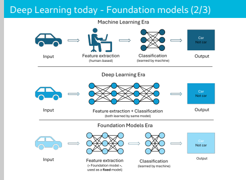
Deep Learning today - Foundation models (1/3)

- Model trained on an Internet scale dataset
- Self-Supervised training: pretext tasks
- Generalization is not a problem anymore! All is about **particularization**

Source: https://blogs.nvidia.com/blog/what-are-foundation-models/

Deep Learning today - Foundation models (2/3)





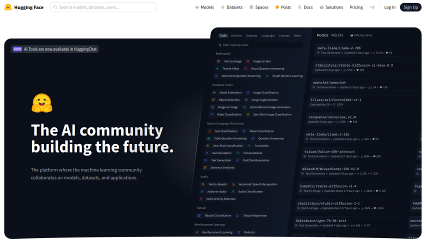
Deep Learning today - Foundation models (3/3)

- Foundation models are really hard to train...
  - Requires a lot (a lot!) of data
  - Requires a lot of computing power
  - Requires a lot of time
- ... But they can be used as powerful feature extractor!
  - Some open-source Foundation models exist, and are available on HuggingFace `https://huggingface.co/`
  - We will use such algorithms during our labs sessions.

Deep Learning today - Foundation models (3/3)

- Foundation models are really hard to train...
  - Requires a lot (a lot!) of data
  - Requires a lot of computing power
  - Requires a lot of time
- ... But they can be used as powerful feature extractor!
  - Some open-source Foundation models exist, and are available on HuggingFace `https://huggingface.co/`
  - We will use such algorithms during our labs sessions.

- Foundation models are really hard to train...
  - Requires a lot (a lot!) of data
  - Requires a lot of computing power
  - Requires a lot of time
- ... But they can be used as powerful feature extractor!
  - Some open-source Foundation models exist, and are available on HuggingFace `https://huggingface.co/`
  - We will use such algorithms during our labs sessions.



2025-10-02

Course 1: Generalities about AI

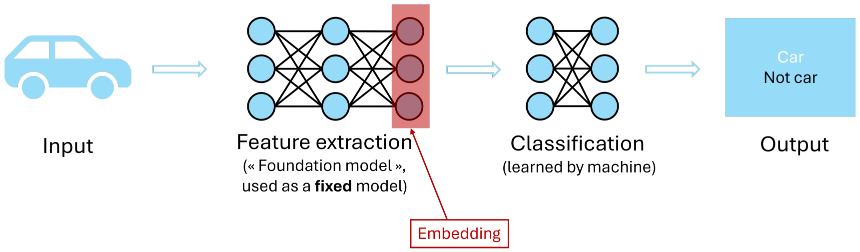
└─Deep Learning today - Foundation models (3/3)

Deep Learning today - Foundation models (3/3)
- Foundation models are really hard to train...
  - Requires a lot (a lot!) of data
  - Requires a lot of computing power
  - Requires a lot of time
- ... But they can be used as powerful feature extractor!
  - Some open-source Foundation models exist, and are available on HuggingFace https://huggingface.co/
  - We will use such algorithms during our labs sessions.

Take-home messages (for now):

- Foundation models represent data with meaningful and rich information (“**embeddings**”) ;
- They can be used to represent a set of data, and extract features.

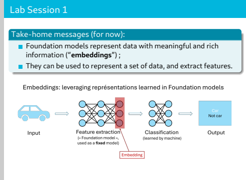
Embeddings: leveraging représentations learned in Foundation models



2025-10-02

Course 1: Generalities about AI

└─Lab Session 1



Lab Session 1

Take-home messages (for now):

- Foundation models represent data with meaningful and rich information (“**embeddings**”) ;
- They can be used to represent a set of data, and extract features.

Lab assignments:

- Introduction to Python, Environments, Numpy, etc,
- Introduction to data visualization for Machine Learning,
- Choose one modality (Text, Image, Audio)
- Tests on some examples of each modality (embeddings are pre-computed for you).

Link to the lab: `https://mee-labs.gitlab-pages.imt-atlantique.fr/intro2ai/lab1/`

Lab Session 1
Take-home messages (for now):
- Foundation models represent data with meaningful and rich information (“**embeddings**”) ;
- They can be used to represent a set of data, and extract features.
Lab assignments:
- Introduction to Python, Environments, Numpy, etc,
- Introduction to data visualization for Machine Learning,
- Choose one modality (Text, Image, Audio)
- Tests on some examples of each modality (embeddings are pre-computed for you).
Link to the lab: https://mee-labs.gitlab-pages.imt-atlantique.fr/intro2ai/lab1/