

Por qué la Inteligencia Artificial no es solo una aplicación

Cómo la IA puede automatizar lo que haces con tus datos no estructurados para aumentar los ingresos y reducir los costos

@luisriverag

Septiembre de 2023

Diseño y Arquitectura

1

- Computación vs IA
- Tipos de IA
- Mantener tus datos privados
- Bases de datos
- Código abierto

2

IA vs Texto

- Capas
- Memoria
- Usando tus propios datos
- Construyendo un modelo
- Ejemplos B2C

IA vs Imágenes

3

- Con tus datos
- Ejemplo de hierba marina
- Ejemplos B2C

Computing is deterministic, AI is probabilistic.

Computing involves using a specific sequence of steps to solve a problem.

AI uses probability and statistics to make educated guesses, or predictions.



MATH

- 1 In mathematics, there is a strict theorem and proof structure.
- 2 Math deals with the understanding of space, numbers, measures, and structures in a simple form.
- 3 Mathematics is a pure science.

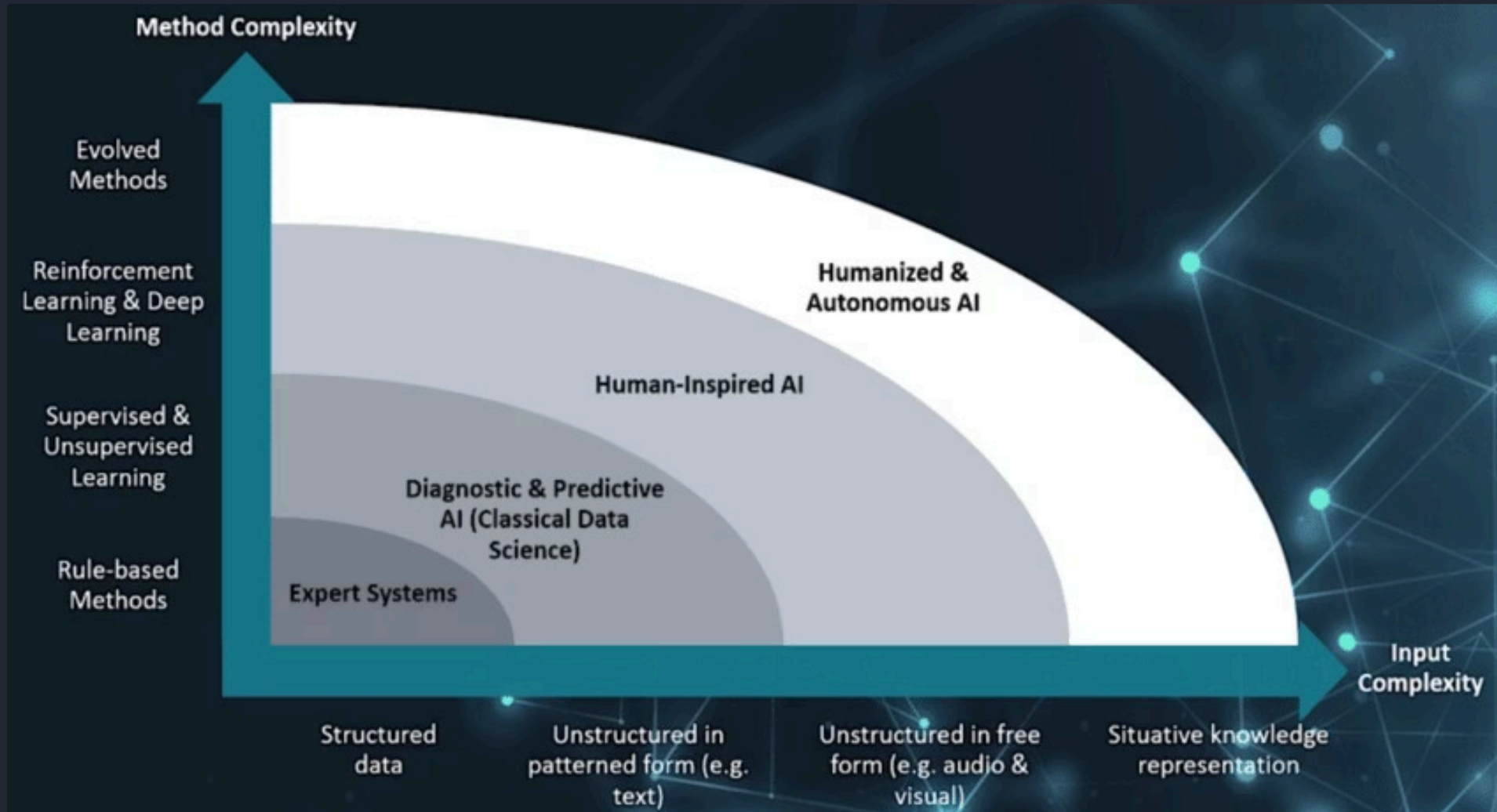
VS



STATISTICS

- 1 Statistics is a field in which people work with real-world data.
- 2 The gathering, sorting, construction, and display of raw data are all covered by statistics.
- 3 Statistics is an applied science.

AI Types



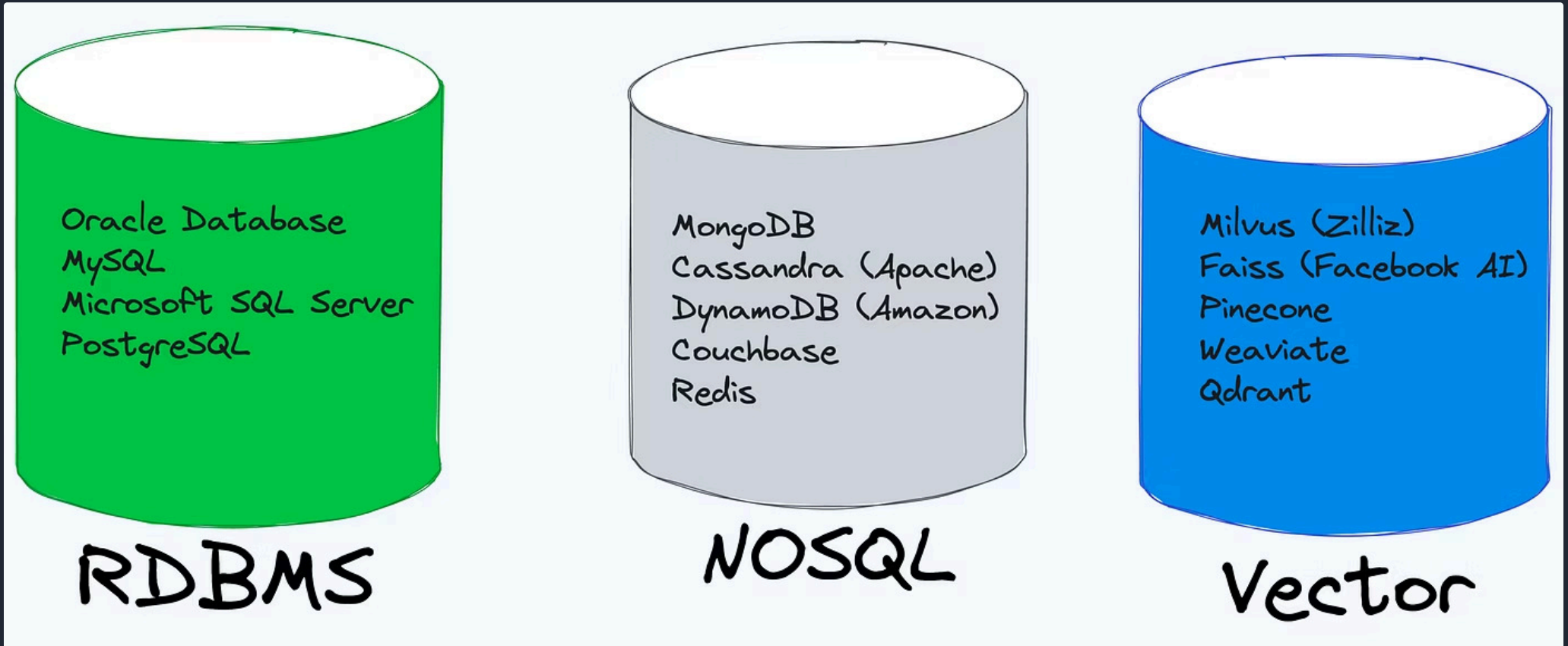
Keeping your data private

You can deploy AI without sharing your secret sauce



Databases

Computation can't work with unstructured data, but AI does



Open source

Stand in the shoulders of giants

H₂O
AutoML



Caffe2



LangChain



Microsoft
Cognitive
Toolkit

LLaMA

∞ Meta + Microsoft



opennn
neural networks

stability.ai

2. AI vs Text: LLMs

¿Qué es LLM?

LLM significa "modelos de lenguaje grandes". Son sistemas de IA que pueden generar texto y escribir de manera similar a como lo hacen los humanos.

¿Qué puede hacer LLM?

LLM puede ayudar a optimizar muchos procesos de escritura, incluyendo la generación de diferentes formas de contenido escrito como correos electrónicos, informes y publicaciones digitales.

¿Cómo funciona LLM?

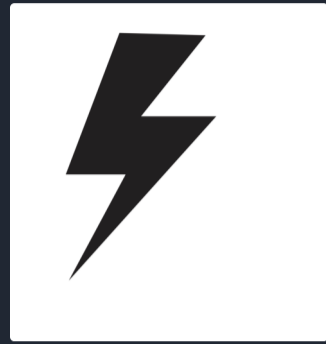
LLM funciona tomando grandes cantidades de texto y utilizando esos datos para crear su propia escritura original que imita los patrones y el estilo del texto que se le ha proporcionado.

¿Por qué usar LLM?

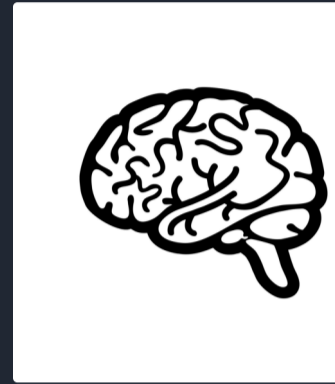
LLM puede ahorrar mucho tiempo y esfuerzo, al tiempo que aumenta la calidad del contenido escrito.

Layers

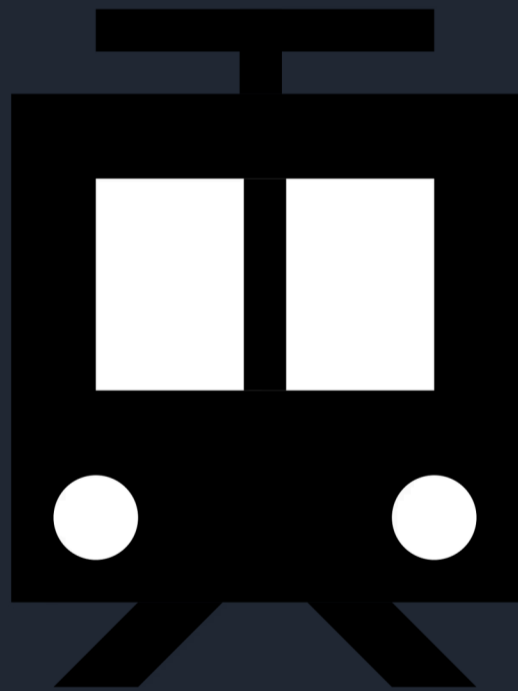
Prompt



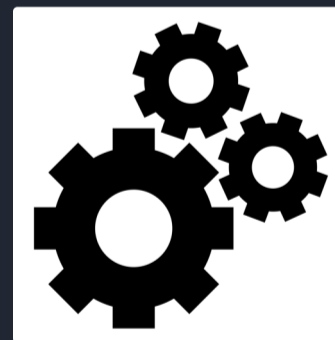
Memory



Rails



Model



"Ingeniería de prompts"

Trucos para obtener los mejores resultados aprovechando los datos en memoria

1 Actúa como si...

Imagínate que estás hablando con una persona real y proporciona contexto e información como lo harías en una conversación normal.

3 Paso a paso

Si quieres generar texto en un orden específico, divide tu prompt en pasos y utiliza el texto generado de cada paso como el prompt para el siguiente paso.

5 Cambia el comienzo

Intenta cambiar el comienzo de tu prompt para ver cómo afecta al texto generado. Por ejemplo, "¿Cuál es el significado de la vida?" vs. "Háblame acerca del significado de la vida".

2 Zero Shot

Intenta utilizar el aprendizaje de cero disparo para generar texto sobre un tema que no se utilizó en los datos de entrenamiento. Ejemplo: "Generar texto sobre la invención de la rueda".

4 Prompts múltiples

Utiliza múltiples prompts para obtener un conjunto de respuestas más diverso. Combina los prompts sumándolos, o utilízalos uno tras otro.

6 Utiliza la personalidad

Añade rasgos de personalidad a tu prompt para generar texto que se ajuste al tono deseado. Por ejemplo, "Escribe una historia divertida sobre un gato" vs. "Escribe una historia aterradora sobre un gato".

Usando tus datos

1 Curación

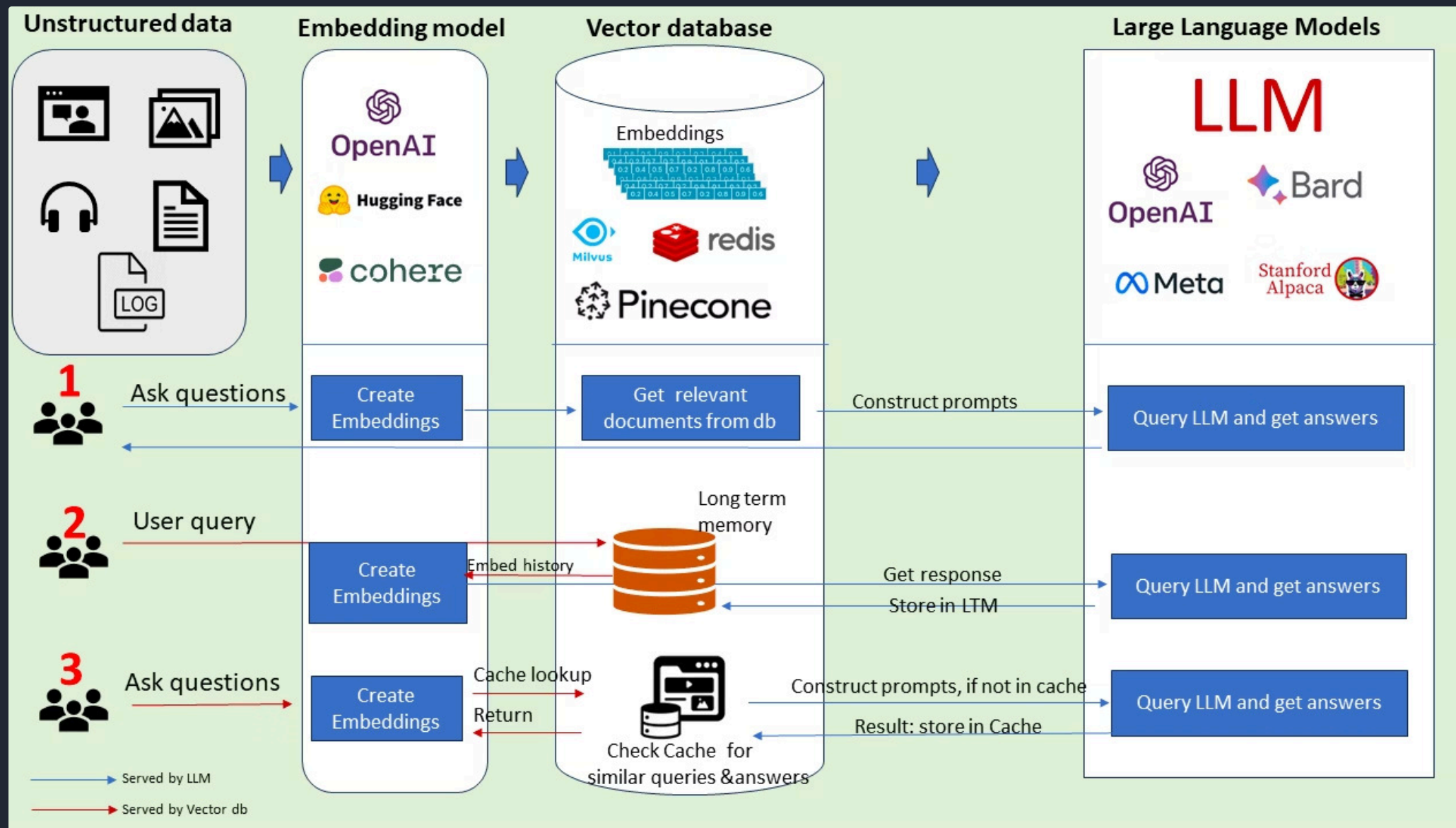
LLM puede curar y personalizar el contenido basándose en los datos del cliente, las preferencias y otra información relevante.

2 Personalización

LLM puede utilizarse para crear contenido único y personalizado para diversas aplicaciones, como descripciones de productos personalizadas o campañas de marketing por correo electrónico personalizadas.

3 Mejora de la eficiencia

LLM puede hacer que las tareas de escritura sean más eficientes y rentables, liberando tiempo y recursos para centrarse en otras prioridades.



Construyendo un modelo

Datos de entrenamiento

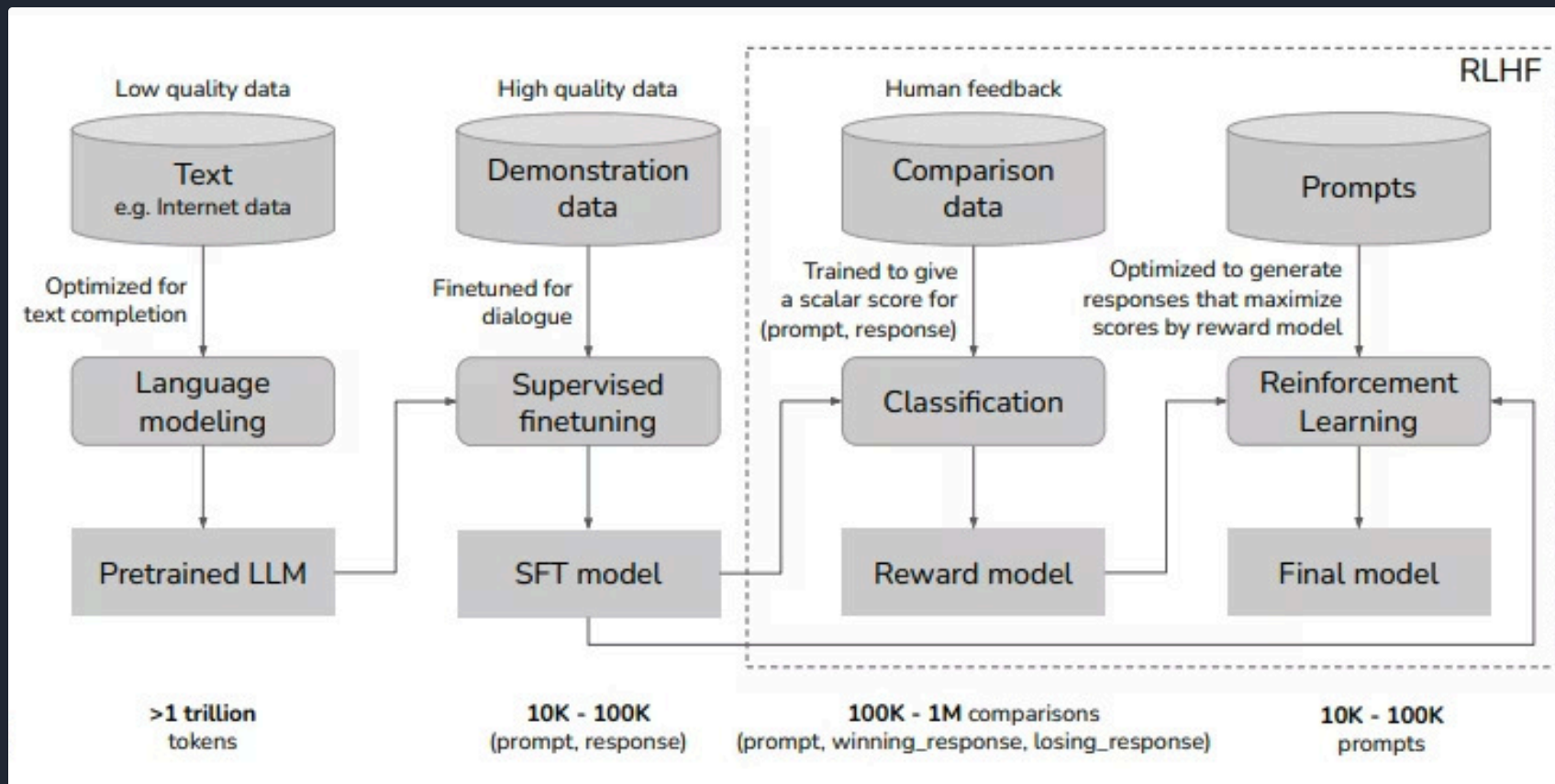
- Recopila y limpia datos de texto de varias fuentes.
- Organiza los datos y elimina duplicados o datos innecesarios.
- Usa los datos preprocesados para ajustar modelos preexistentes o construir un nuevo modelo de lenguaje.

Ajuste fino

- Elige un modelo de lenguaje preexistente a gran escala (como GPT-3) para usar como modelo base.
- Ajusta el modelo base con tus datos para crear un LLM personalizado.
- Prueba y valida el LLM para asegurarte de que genere salidas precisas y útiles.

Implementación

- Integra el LLM en tu flujo de trabajo o aplicación.
- Monitorea continuamente sus salidas y rendimiento, haciendo ajustes donde sea necesario.



Local Chat-bot with company internal documents Example

<https://github.com/PromptEngineer/localGPT/>

Ask questions to your documents without an internet connection, using the power of LLMs. 100% private, no data leaves your execution environment at any point. You can ingest documents and ask questions without an internet connection!

Built with [LangChain](#) and [Vicuna-7B](#) and [InstructorEmbeddings](#)

1 `ingest.py` uses LangChain tools to parse the document and create embeddings locally using InstructorEmbeddings. It then stores the result in a local vector database using Chroma vector store.

2 `run_localGPT.py` uses a local LLM to understand questions and create answers. The context for the answers is extracted from the local vector store using a similarity search to locate the right piece of context from the docs.

Article. II.

SECTION. 1

The executive Power shall be vested in a President of the United States of America. He shall hold his Office during the Term of four Years, and, together with the Vice President, chosen for the same Term, be elected, as follows:

Each State shall appoint, in such Manner as the Legislature thereof may direct, a Number of Electors, equal to the whole Number of Senators and Representatives to which the State may be entitled in the Congress: but no Senator or Representative, or Person holding an Office of Trust or Profit under the United States, shall be appointed an Elector.

[The Electors shall meet in their respective States, and vote by Ballot for two Persons, of whom one at least shall not be an Inhabitant of the same State with themselves. And they shall make a List of all the Persons voted for, and of the Number of Votes for each; which List they shall sign and certify, and transmit sealed to the Seat of the Government of the United States, directed to the President of the Senate. The President of the Senate shall, in the Presence of the Senate and House of Representatives, open all the Certificates, and the Votes shall then be counted. The Person having the greatest Number of Votes shall be the President, if such Number be a Majority of the whole Number of Electors appointed; and if there be more than one who have such Majority, and have an equal Number of Votes, then the House of Representatives shall immediately chuse by Ballot one of them for President, if they have a Majority, then from the five

the said House shall in like Manner chuse the President.

B2C Examples



3. AI vs Images

1 Visual Recognition

Image AI can recognize visual elements of images, allowing it to automatically add captions, tags, and descriptions, making it easier to manage large photo libraries.

2 Photo Editing

Image AI can help automate various photo editing tasks like resizing, cropping, and adjusting color and brightness, saving time and streamlining the editing process.

3 Creative Suggestions

Image AI can suggest unique and creative ways to present visual elements, offering insights and inspirations which can be used in designing campaigns or developing brands.

Images with Your Data



Interpretation

Image AI can recognize and analyze visual patterns like seagrass in aquatic environments. Using this data, scientists are able to monitor the presence and health of seagrass, which is critical for understanding marine ecosystems.

Design

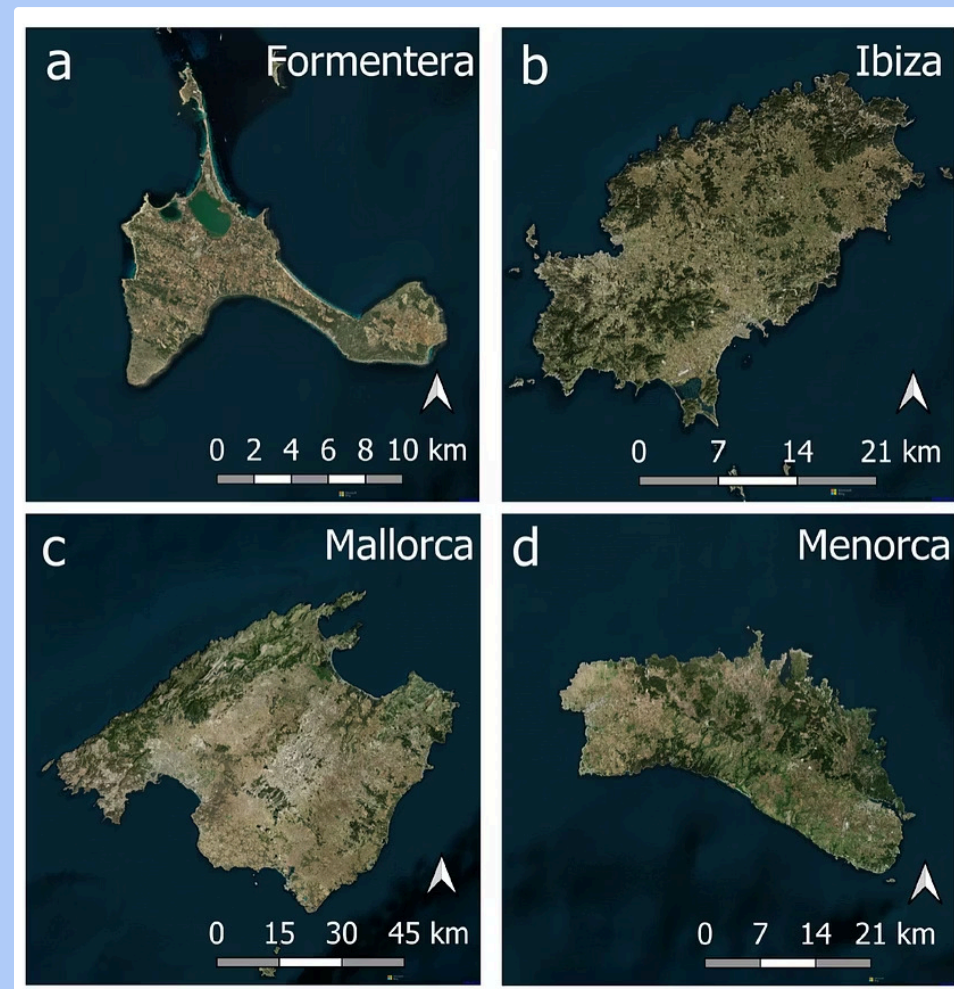
Image AI can help designers generate unique ideas by analyzing user preferences, colors, and space layouts to make suggestions and even create 3D visualizations of potential designs.



Efficiency

Image AI can analyze visual data to provide insights on solar panel efficiency, helping designers and installers optimize placement and alignment to maximize energy production.

Seagrass Example

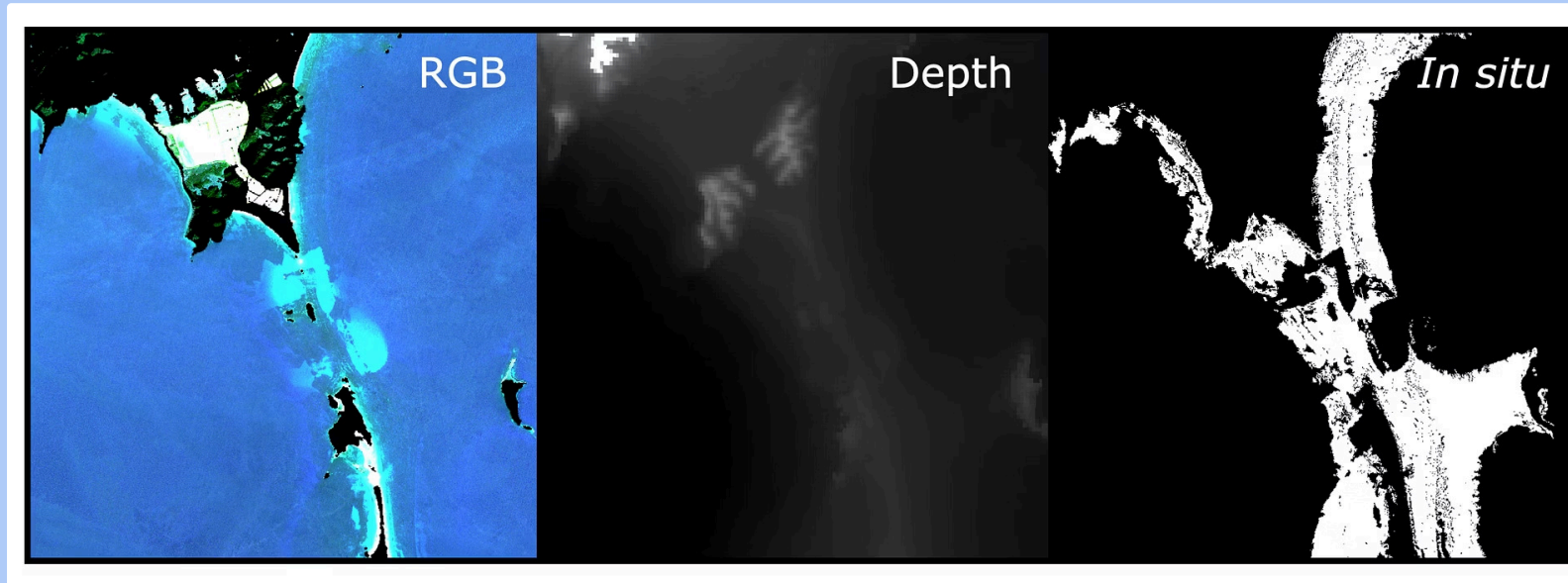


In this real example the aim was to develop an algorithm to detect the endangered seagrass *Posidonia Oceanica* from satellite images. Real maps of *Posidonia* exist from surveys done on site by biologists, however these are slow to collect and expensive. The Sentinel 2 satellites revisit the Balearic Islands every five days, their data is free and publicly available, so the aim was to monitor the evolution of *Posidonia* using satellite images.

We began with satellite images of the Balearic Islands from the Sentinel 2 satellites. The images consist of geologicalized images using the red, green and blue bands (RGB). We will attempt to classify each pixel as having *Posidonia*, or not having *Posidonia*.

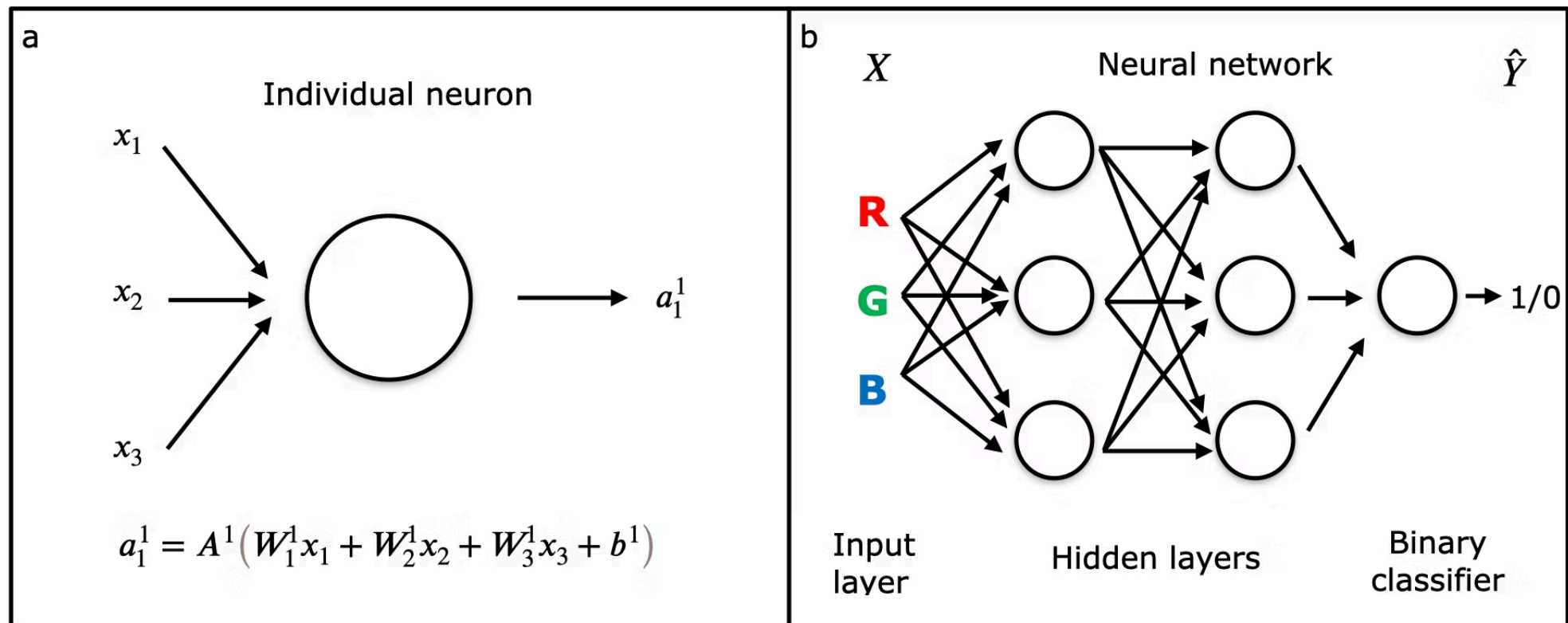
The individual images were not found to be good enough for individual images to be used. It was decided to combine the data from several satellite visits, so to merge ("stack") the data from different dates. This is common in data science: we often need to curate the data before using it. Inexperienced data scientists often use bad quality data, compromising the results. Since machine learning is driven by data, it is often said that "junk in, junk out", meaning that no matter how sophisticated the algorithm, if data of poor quality is used (or worse, statistically biased data), the final result will be meaningless.

Seagrass Example



The RGB image was combined with other data, such as seabed depth (bathymetry) and real maps of Posidonia seagrass made by biologists (*In situ* data). This data shows the true distribution of the seagrass, and it is used to train the deep learning neural network. Data science involves putting together different datasets, often from different sources, and this can require a significant amount of work. The inability to combine all the available data is one of the reasons why a project can fail.

Seagrass Example



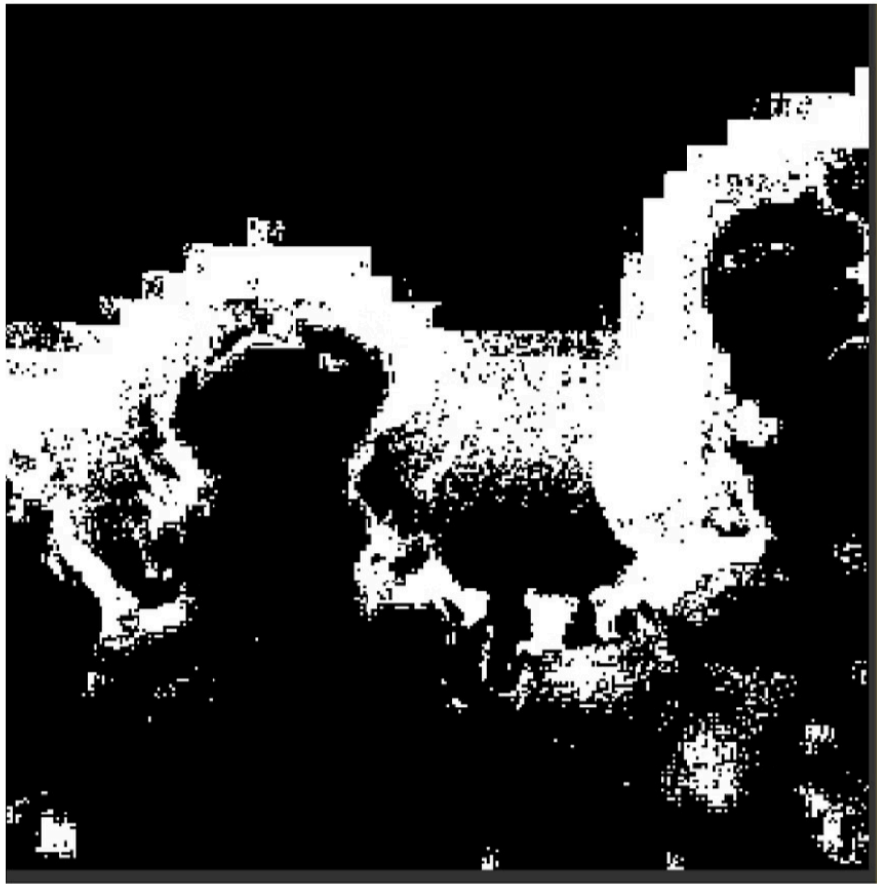
A deep neural network consists of several layers, each of them containing a number of "neurons". These neurons are nodes which have as inputs all the outputs of the previous layer, or in the case of the first layer, the actual data. The neurons perform a linear combination of the inputs using coefficients. The last layer has a different neuron that assigns a class to the data. In this case the class is 1 (Posidonia) or 0 (not Posidonia).

As we train the neural network with real data (the seagrass maps, in this case), the coefficients in each neuron are updated to improve the match between the class (predicted from the RGB images) and the correct class (known from the *In situ* data).

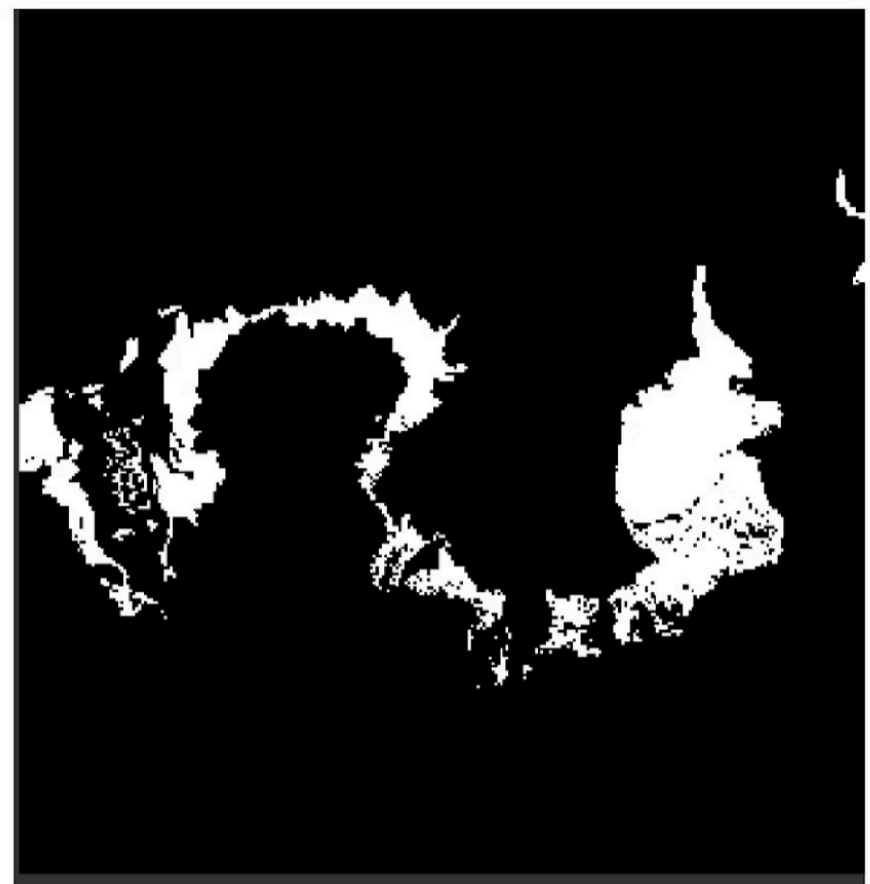
The neural network must have enough neurons and enough layers to be able to reproduce the complexity of the data: a network that is too simple might never be able to classify correctly the data. However, a network that is too complex might be too slow to finish the project on time.

Seagrass Example

Cost function: quantifies difference between predicted and real

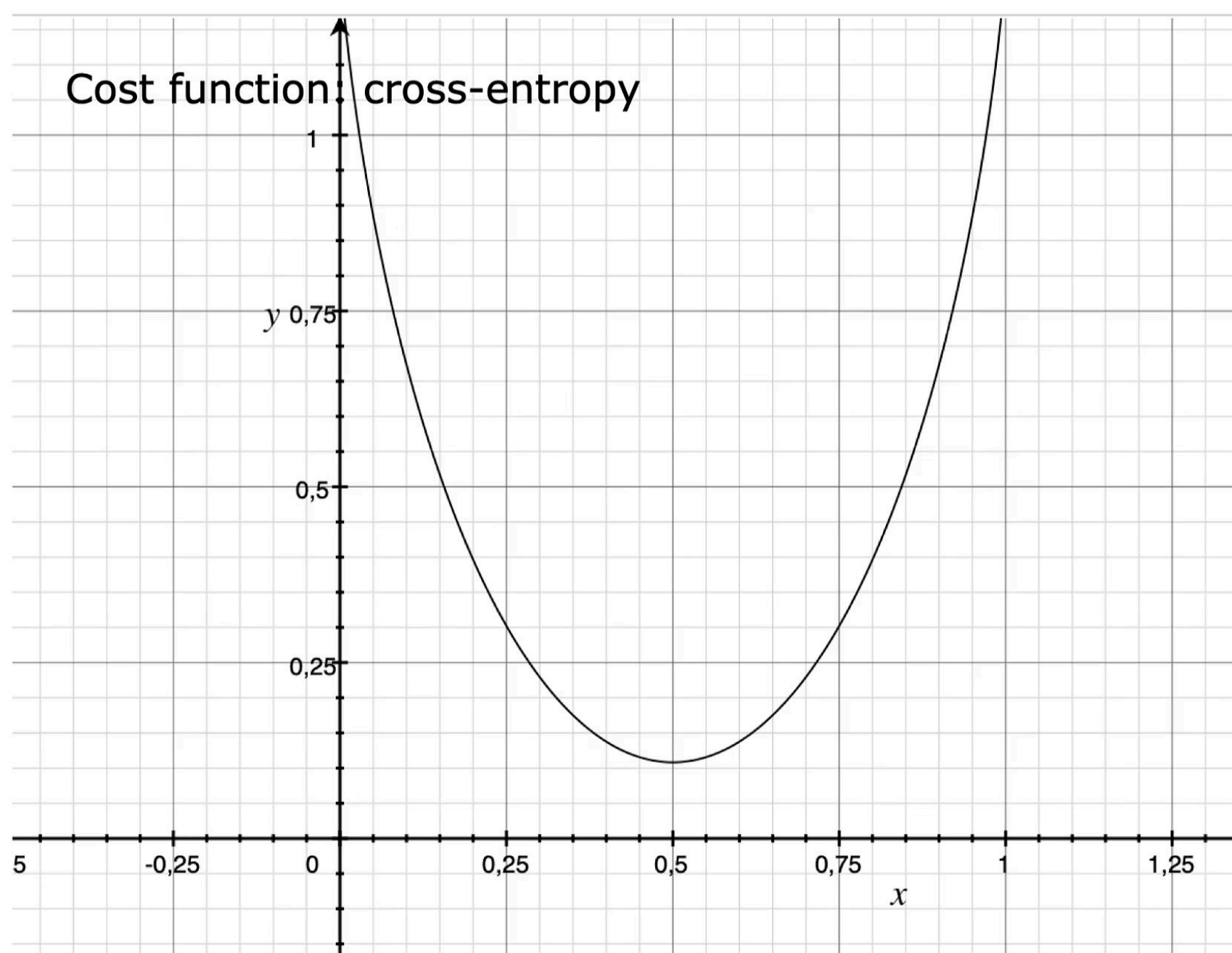


Predicted



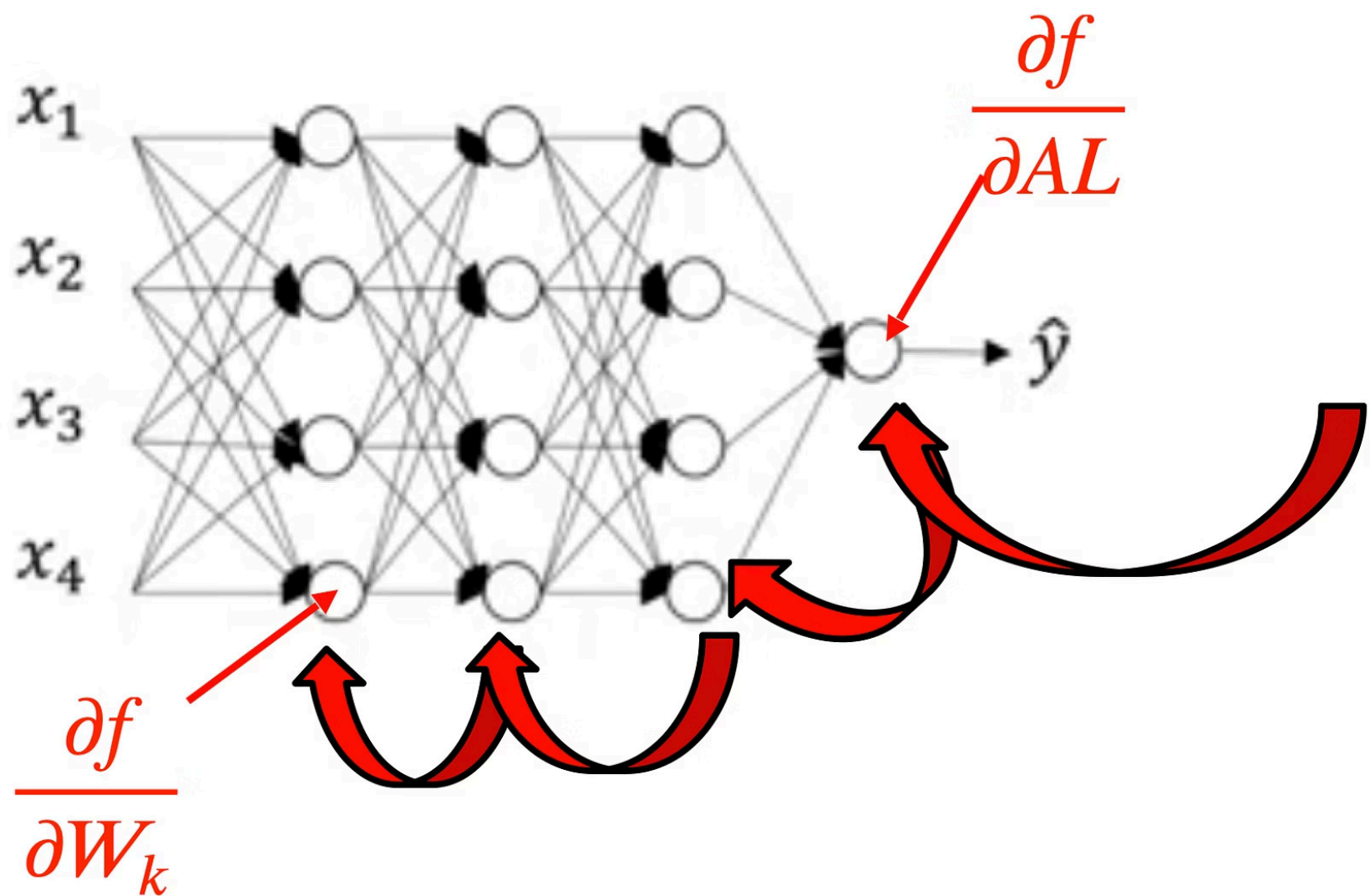
Real

It is the cost function that quantifies how different the predicted classes are from the correct classes, for all the available data together. When splitting data into classes (such as Posidonia/not Posidonia), the appropriate cost function is the cross-entropy. For regression problems (such as predicting numeric values), a least squares cost function is appropriate.

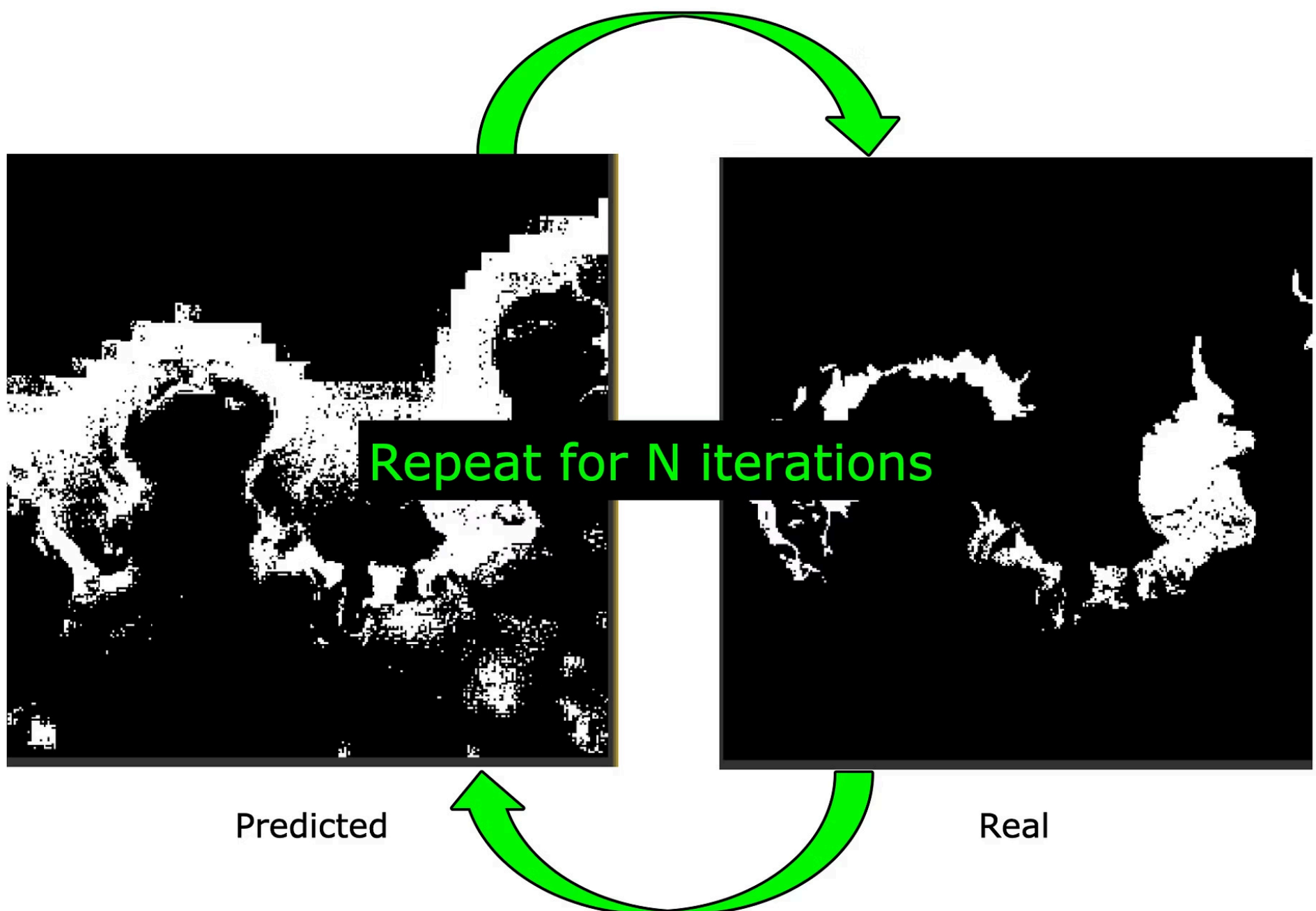


The cost function decreases as the difference between predicted class and real class decreases. Hence finding the minimum value of the cost function gives us the coefficients of the neural network that best fit the real data. This is done iteratively: in each iteration we find a step towards the minimum, and after enough steps we expect to have reached the minimum.

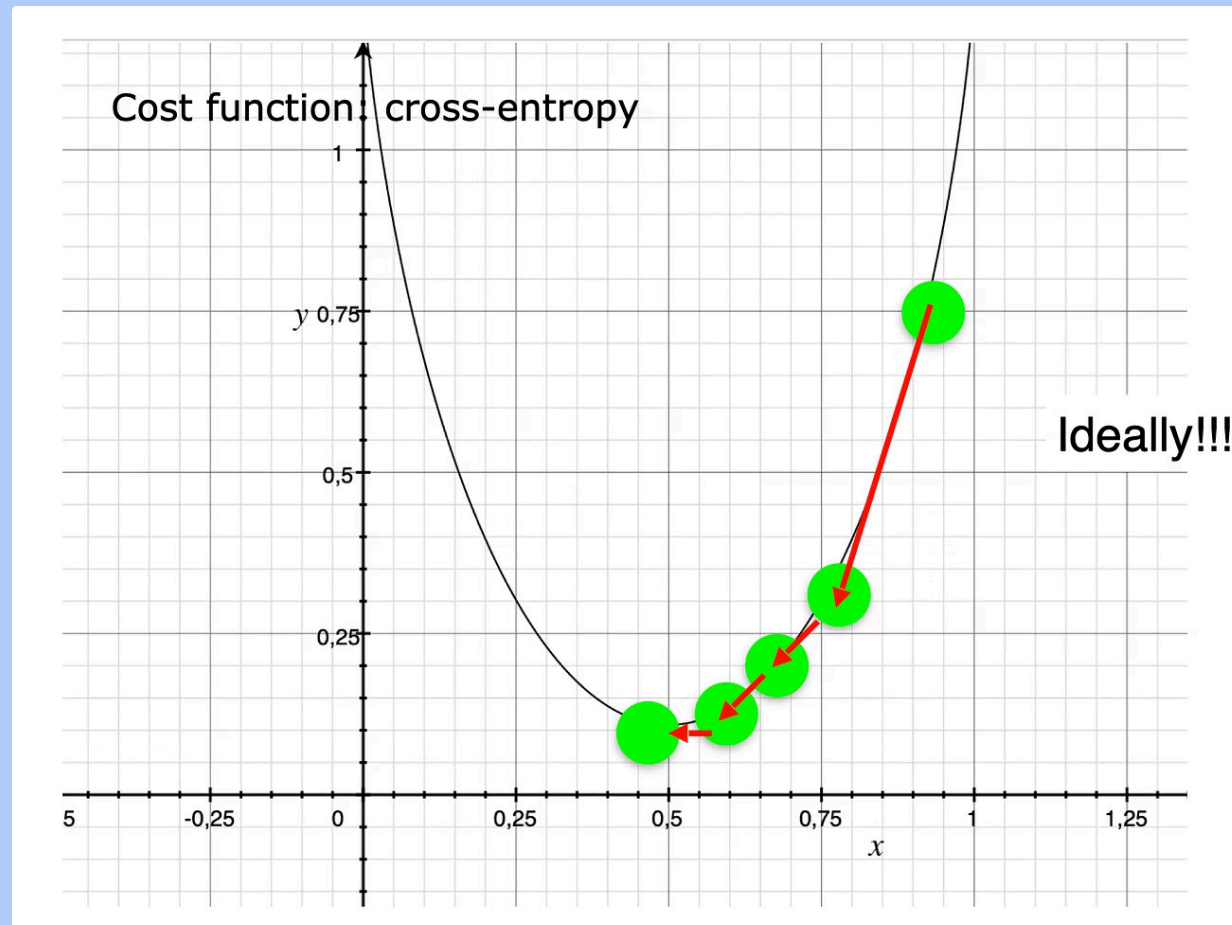
Seagrass Example



The differences between the predicted and real data are called the residuals, and these are back-propagated into the neurons. This means we use partial derivatives to calculate how much each neuron needs to be updated in order to reduce the cost function. This is an iterative process: in each iteration we produce predicted data with the current values of the coefficients inside the neurons, compare these to the real data, compute a cost function and back propagate the residuals to update the coefficients. This is repeated thousands or millions of times.



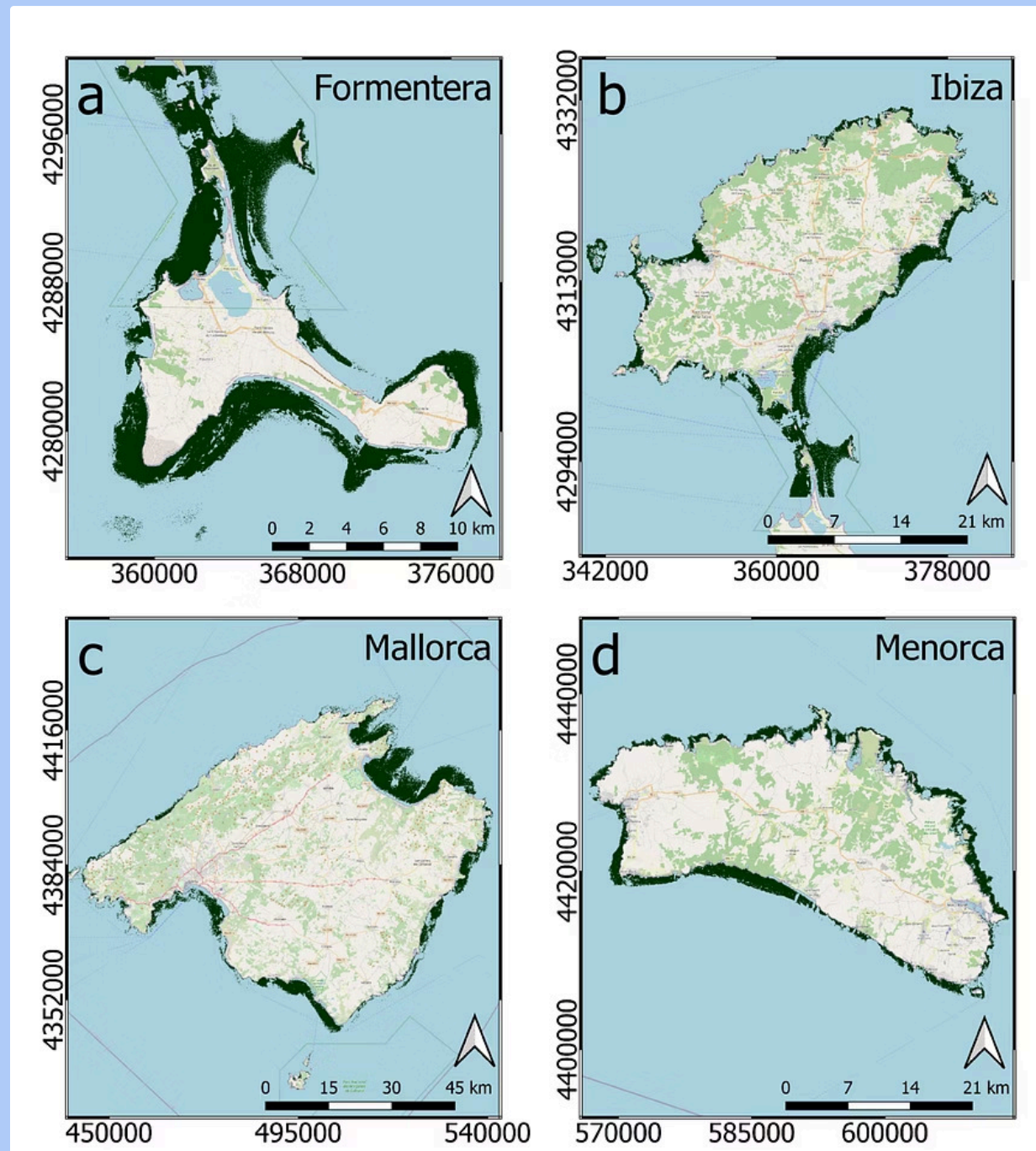
Seagrass Example



As we iterate, the cost function decreases as the difference between predicted class and real class decreases. Hence finding the minimum value of the cost function gives us the coefficients of the neural network that best fit the real data. This is done iteratively: in each iteration we find a step towards the minimum, and after enough steps we expect to have reached the minimum.

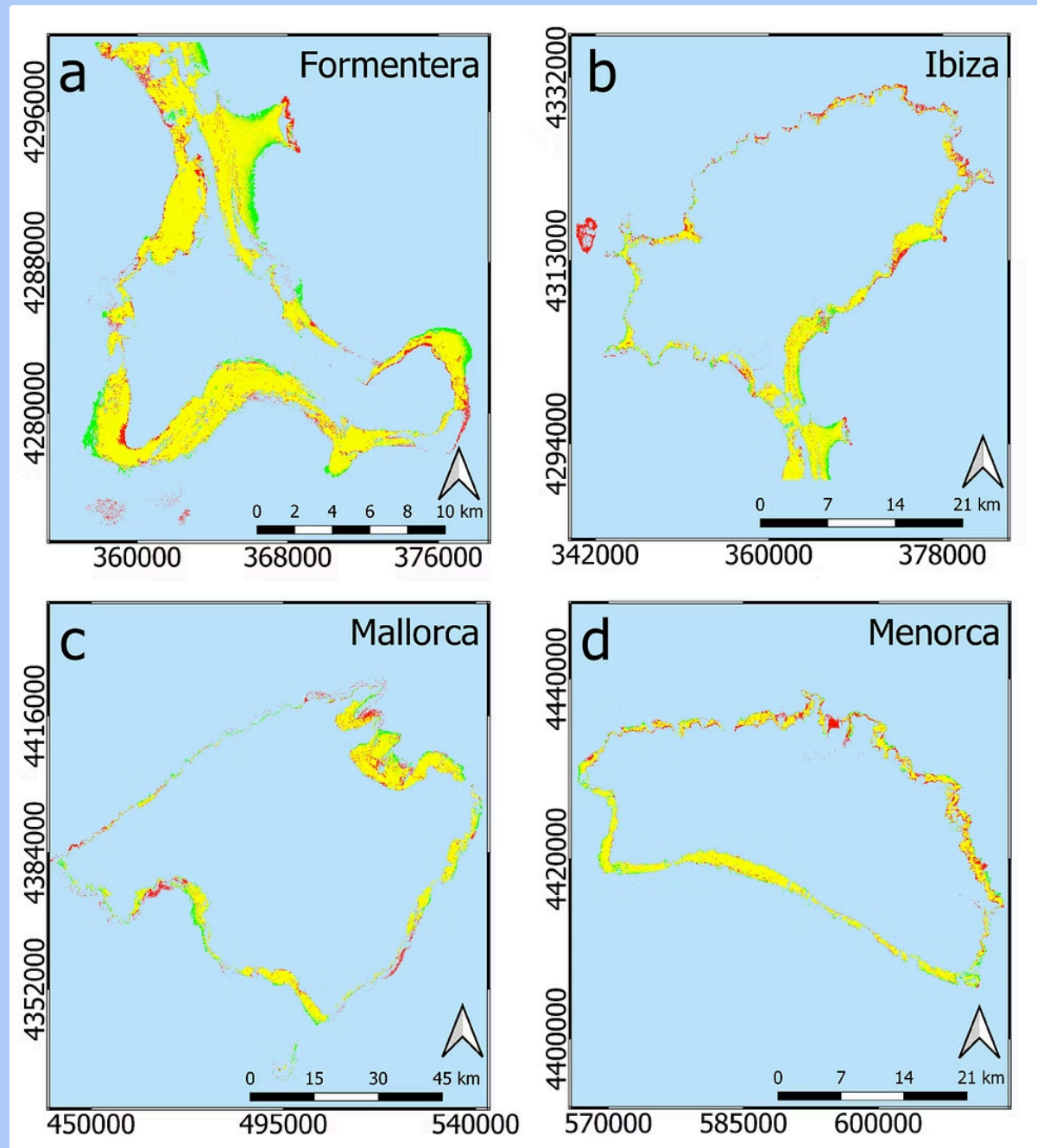
This is the part of the project in which we train the neural network, and hence the data with real classes is known as the training set. There are a lot of mathematical details here, and it is not trivial to reach the global minimum, and problems with the cost function, the step size, and with local minima can all lead to poor results.

Seagrass Example



Here are the *Posidonia Oceanica* seagrass maps produced by our neural network in the Balearic Islands. The next slide compares the prediction to the real (*In situ*) data.

Seagrass Example



This image uses a colour coding to compare the predicted results to the real results. Yellow means the model correctly predicted Posidonia to be present, and blue means the model correctly predicted Posidonia to be absent. Red means the model incorrectly assigned Posidonia (whereas the real data had none), green means the model incorrectly did not assign Posidonia (the real data did have Posidonia).

This is a good example of how AI is probabilistic. Most of the pixels have been correctly labelled, and the overall outline of the islands can be seen clearly. However, there a small percentage of the pixels have their class incorrectly assigned, and these can be false positives (red) or false negatives (green).

B2C Examples



Midjourney



Leonardo.Ai