



Elements of Seismology & Machine Learning

Gabriele Paoletti

gabriele.paoletti@uniroma1.it

Elisa Tinti

elisa.tinti@uniroma1.it

Daniele Trappolini

daniele.trappolini@uniroma1.it

Opening

Table of contents

1.	Introduction	3
02.	Machine Learning	9
03.	Deep Learning	16
04.	Training	26
05.	Applications	38

01.

Introduction

This section serves as an introduction to our comprehensive exploration of artificial intelligence. We'll provide an overview of key concepts, history, and the impact of AI and machine learning technologies in various sectors.

A.I. TIMELINE



1950

TURING TEST

Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence

1955

A.I. BORN

Term 'artificial intelligence' is coined by computer scientist, John McCarthy to describe "the science and engineering of making intelligent machines"

1961

UNIMATE

First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line

1964

ELIZA

Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966

SHAKY

The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I. WINTER

Many false starts and dead-ends leave A.I. out in the cold

1997

DEEP BLUE

Deep Blue, a chess-playing computer from IBM defeats world chess champion Gary Kasparov

1998

KISMET

Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot insofar as it detects and responds to people's feelings



1999

AIBO

Sony launches first consumer robot pet dog Aibo (AI robot) with skills and personality that develop over time

2002

ROOMBA

First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes

2011

SIRI

Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S

2011

WATSON

IBM's question answering computer Watson wins first place on popular \$1M prize television quiz show *Jeopardy*

2014

EUGENE

Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human

2014

ALEXA

Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks

2016

TAY

Microsoft's chatbot Tay goes rogue on social media making inflammatory and offensive racist comments

2017

ALPHAGO

Google's A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2^{170}) of possible positions

Introduction

Text to text: *ChatGPT*



December 10, 2004

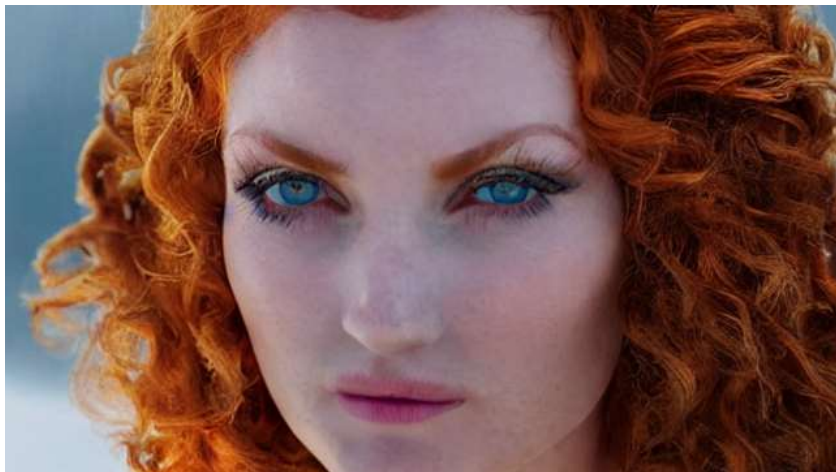


November 30, 2022

"Tell me about San Francisco"

Introduction

Text to image: *Midjourney*



October 15, 2022



March 05, 2023

“A young woman with vibrant red hair and striking blue eyes stands amidst a gentle snowfall, medieval-inspired armor, ...”

Introduction

Text to video: *Sora*



March 30, 2023



February 15, 2024

"A movie trailer featuring the adventures of the 30 year old space man wearing a red wool knitted motorcycle helmet, blue sky, salt desert, cinematic style, shot on 35mm film, ..."

Artificial intelligence

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

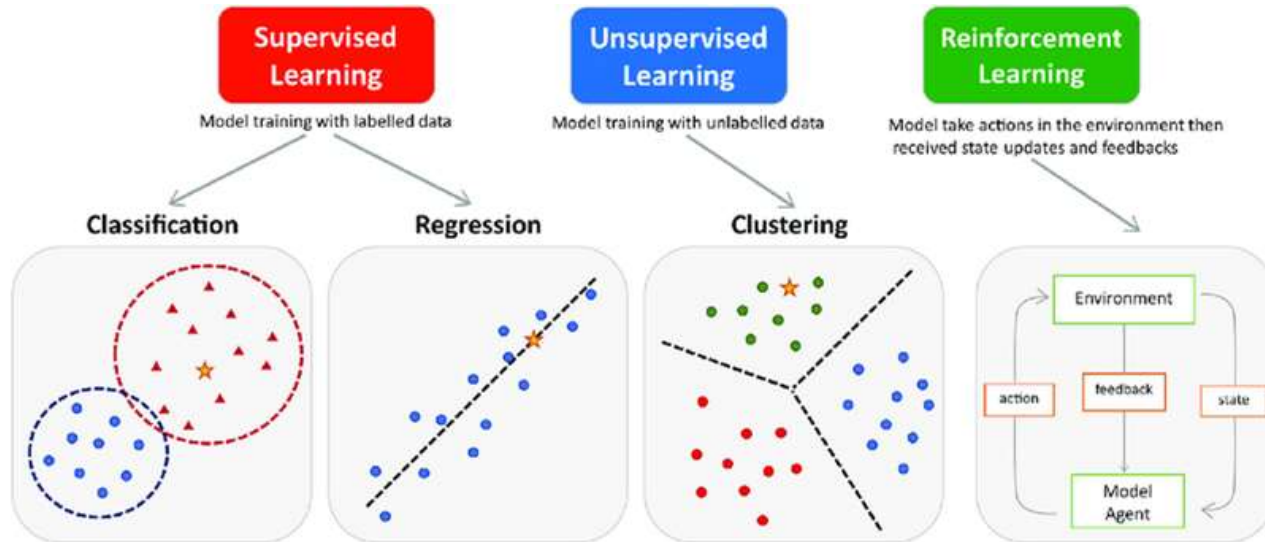
02.

Machine Learning

This section explores the fundamentals of Machine Learning, covering key concepts and techniques used to enable machines to learn from data.

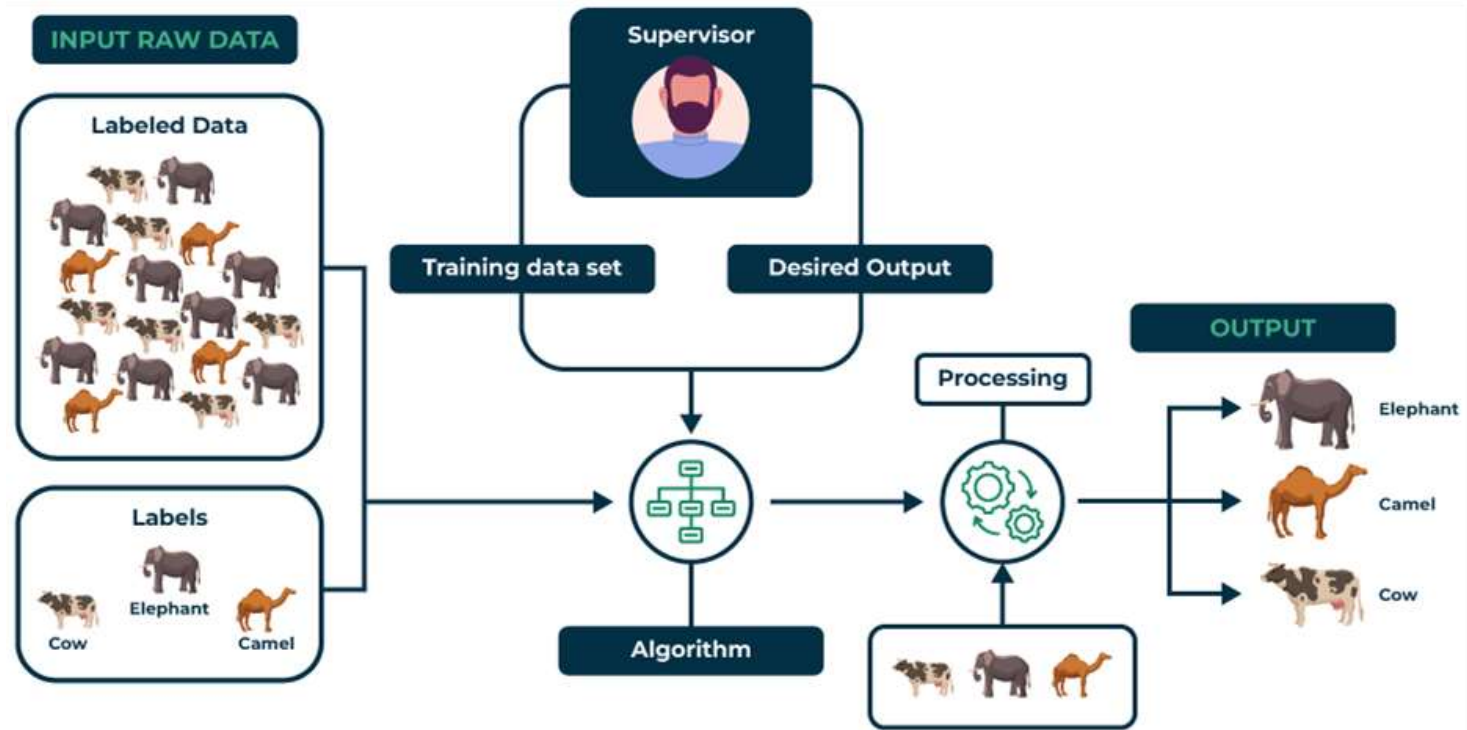
Three main paradigms

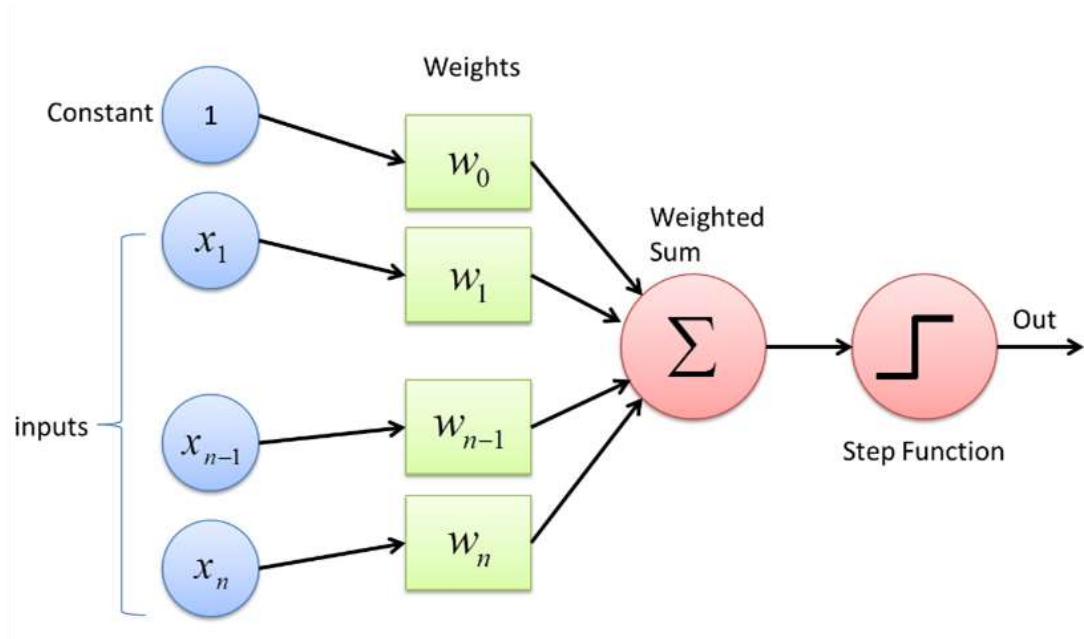
"Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed" ~Arthur Samuel (1959)



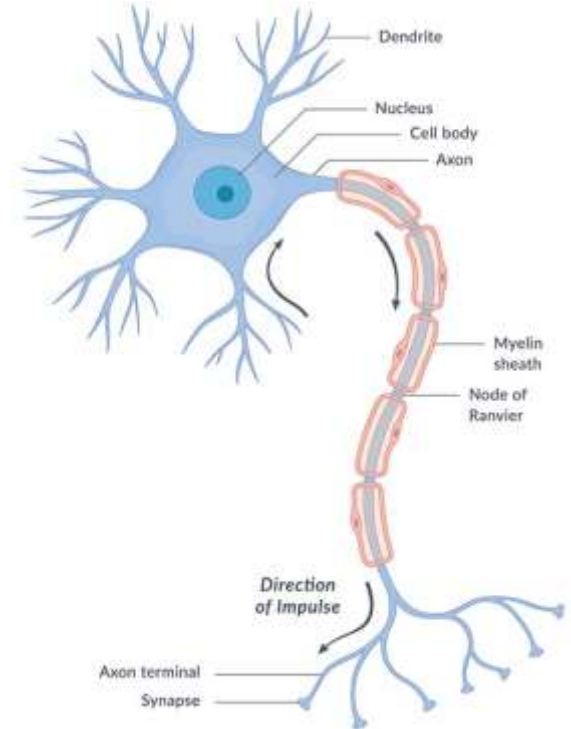
Machine Learning

Supervised learning



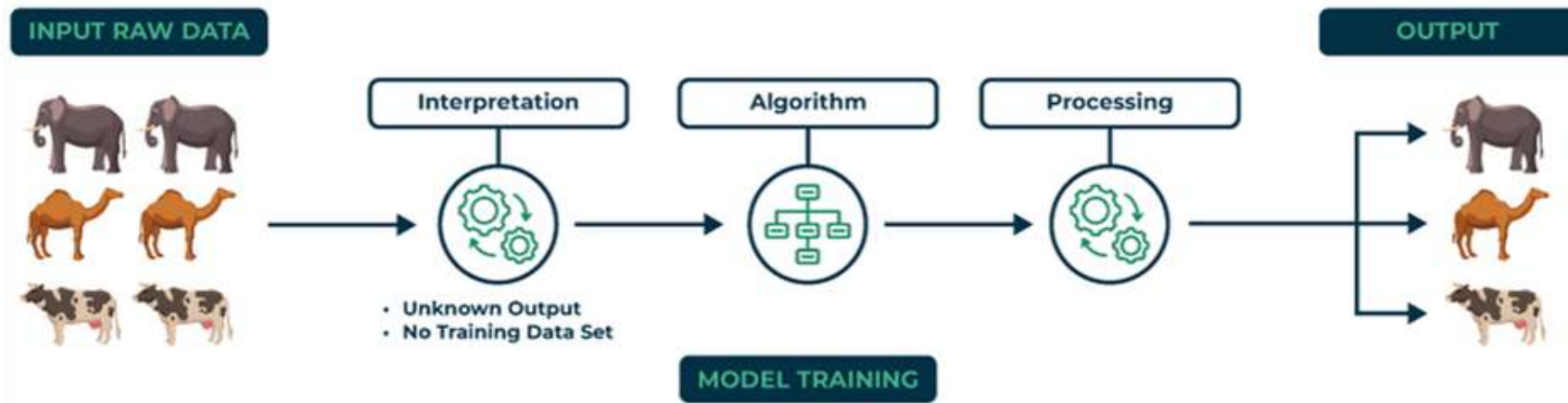


Rosenblatt et al. 1957



Unsupervised learning

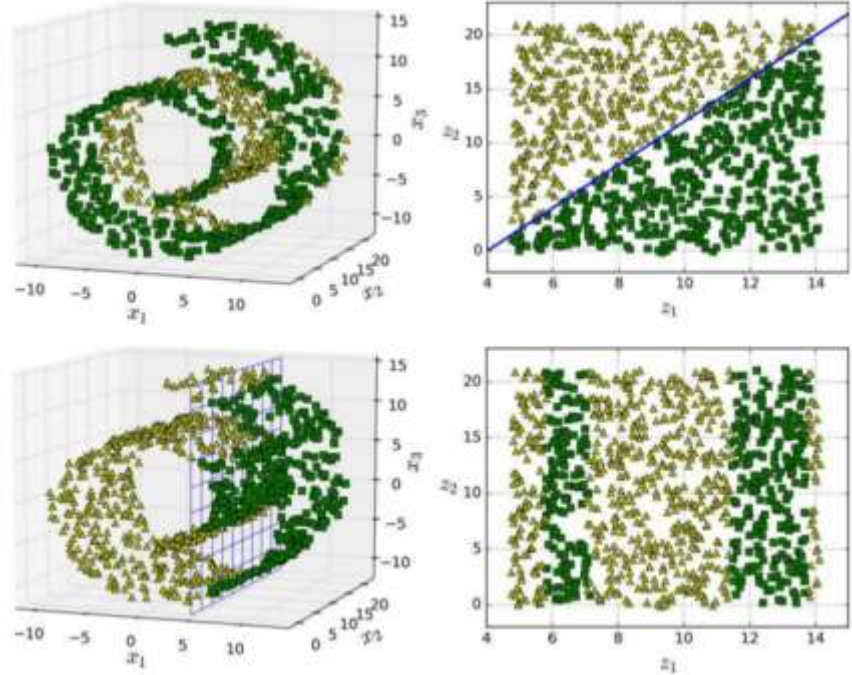
Unsupervised learning is a type of machine learning where models discover patterns in data without pre-existing labels, often used for clustering or dimensionality reduction.



Iteration #01
(inertia: 3622.78)



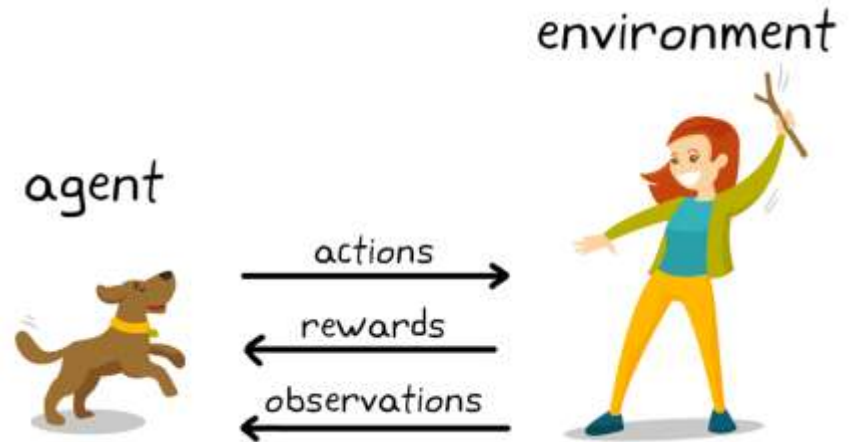
K-Means



PCA

Reinforcement learning

Reinforcement Learning is a machine learning method where an agent learns optimal actions through trial and error to maximize rewards in an environment.



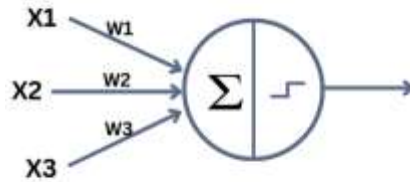
03.

Deep Learning

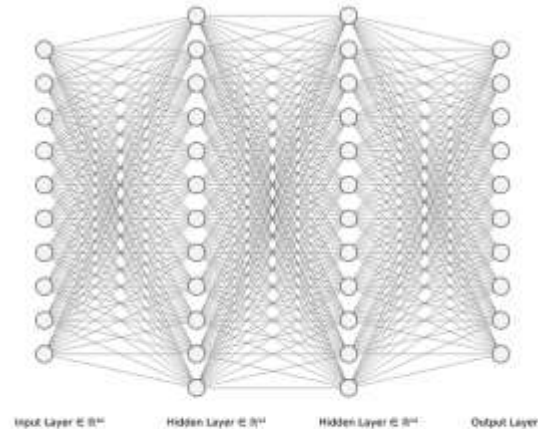
In this section, we focus on Deep Learning, a subset of machine learning that utilizes neural networks with many layers. We'll examine how deep learning models can learn complex patterns and perform tasks like image and speech recognition.

Neural Networks (NNs)

A computational model inspired by the human brain's structure, and consists of layers of interconnected nodes or neurons that process and transmit signals to solve tasks.



Single-layer perceptron

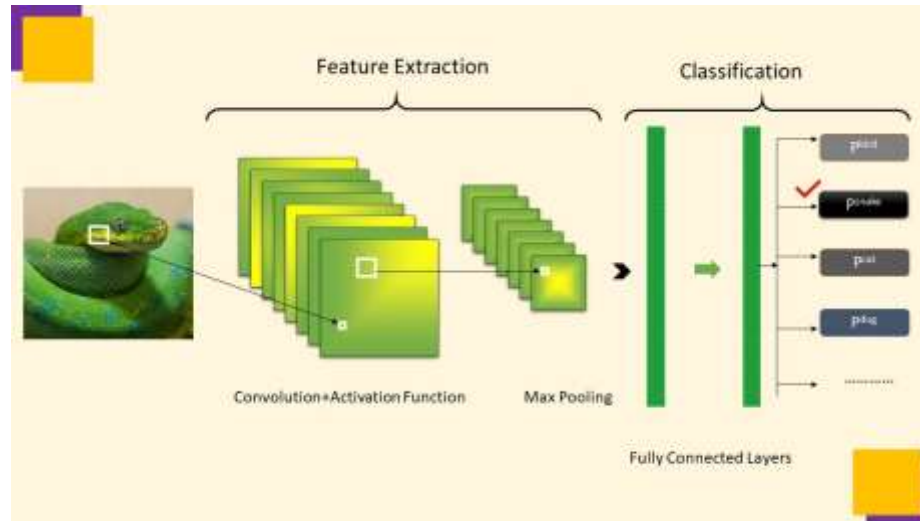


Multi-layer perceptron

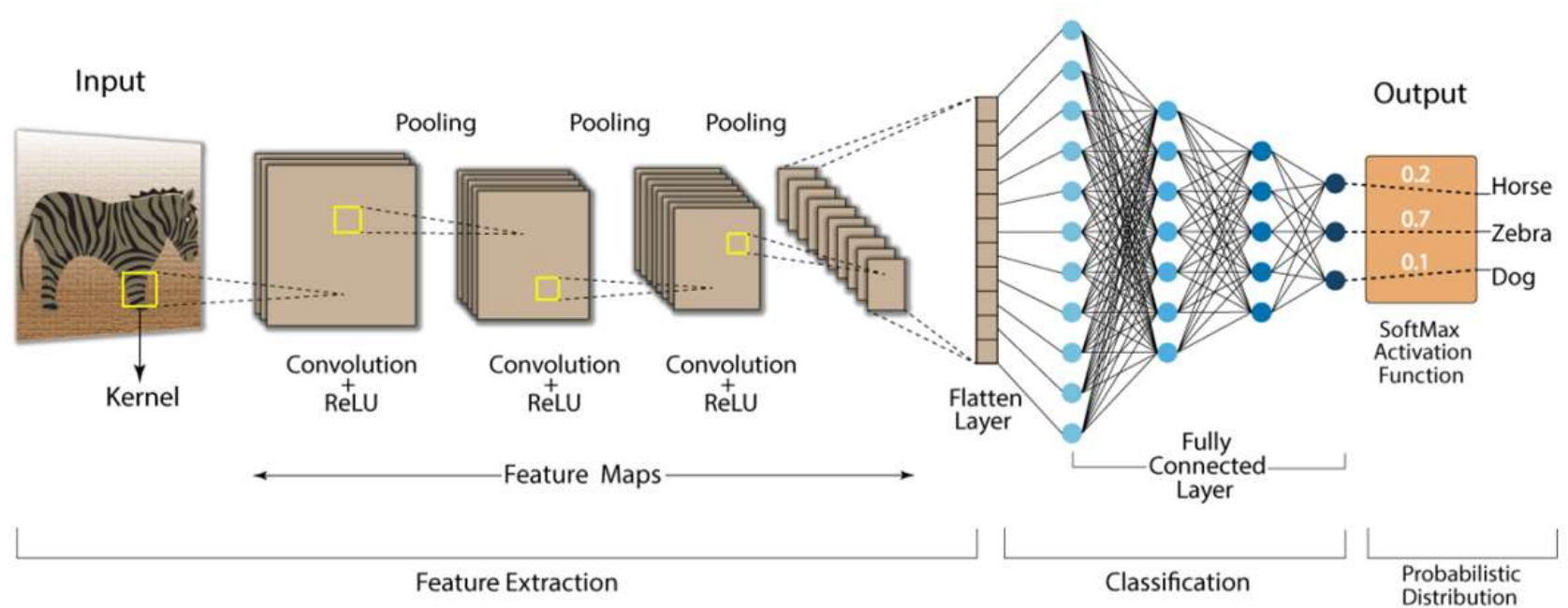
Deep Learning

Convolutional Neural Networks: *What is a convolution*

Convolution involves sliding a smaller array, known as a kernel or filter, over a larger array (the input signal or image) to produce a new array called the convolved feature or feature map.



Convolutional Neural Networks (CNNs)



Classification



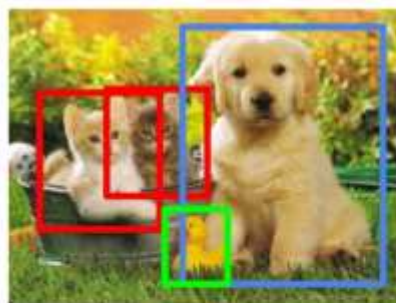
CAT

Classification +
Localization



CAT

Object detection



CAT, DOG, DUCK

Instance
segmentation



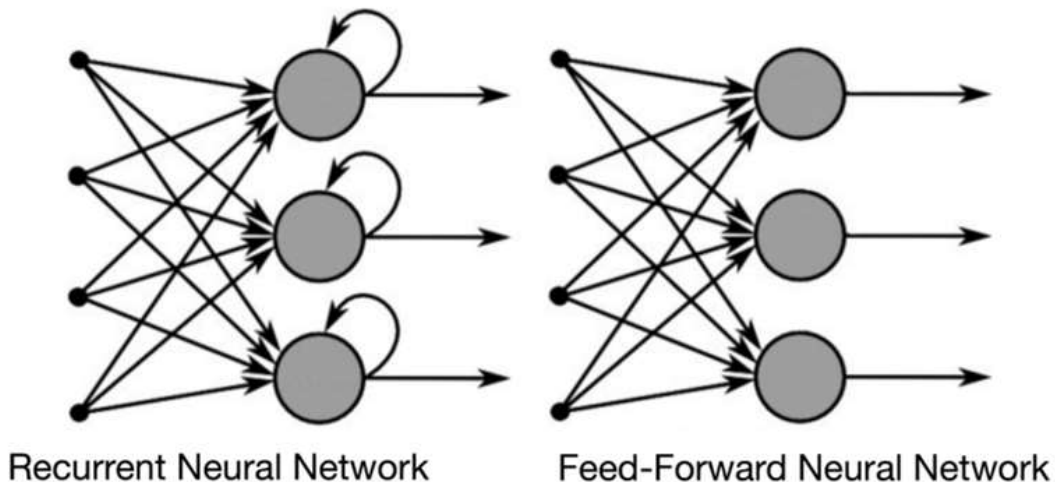
CAT, DOG, DUCK

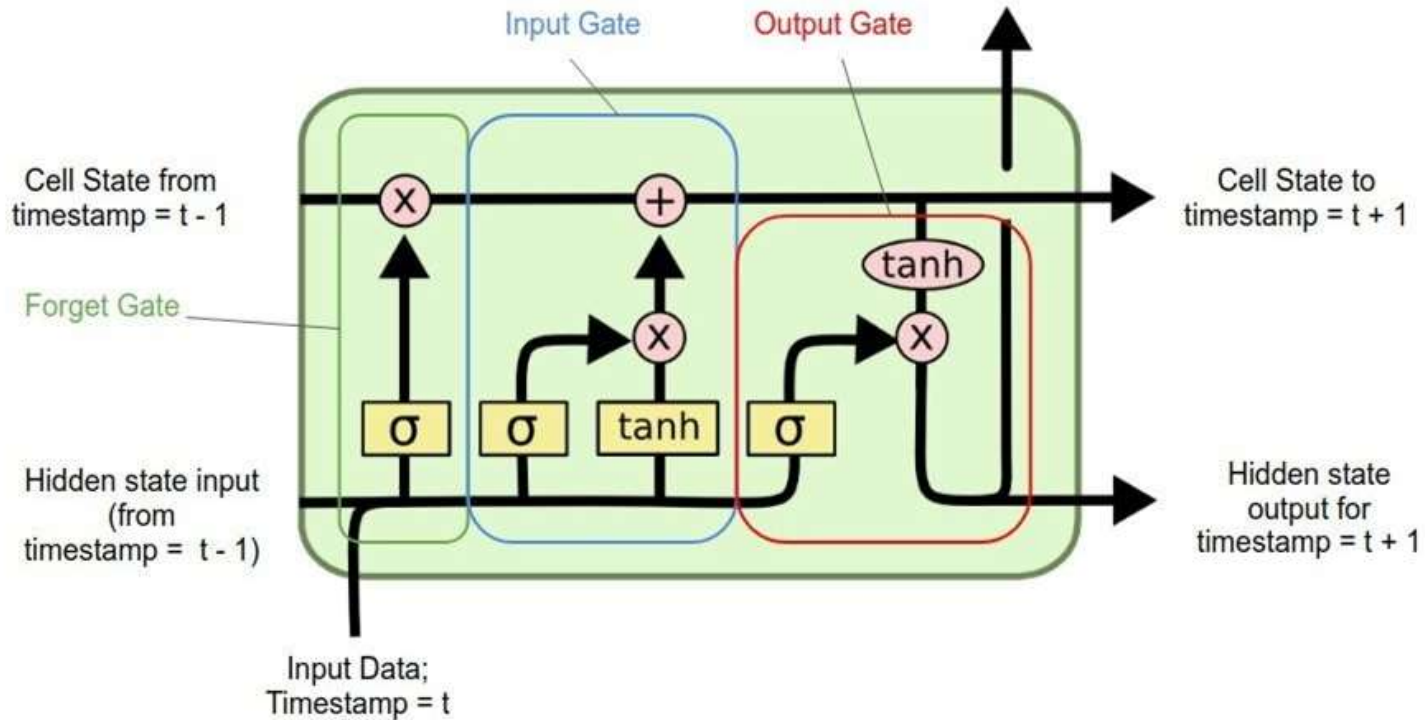
Single object

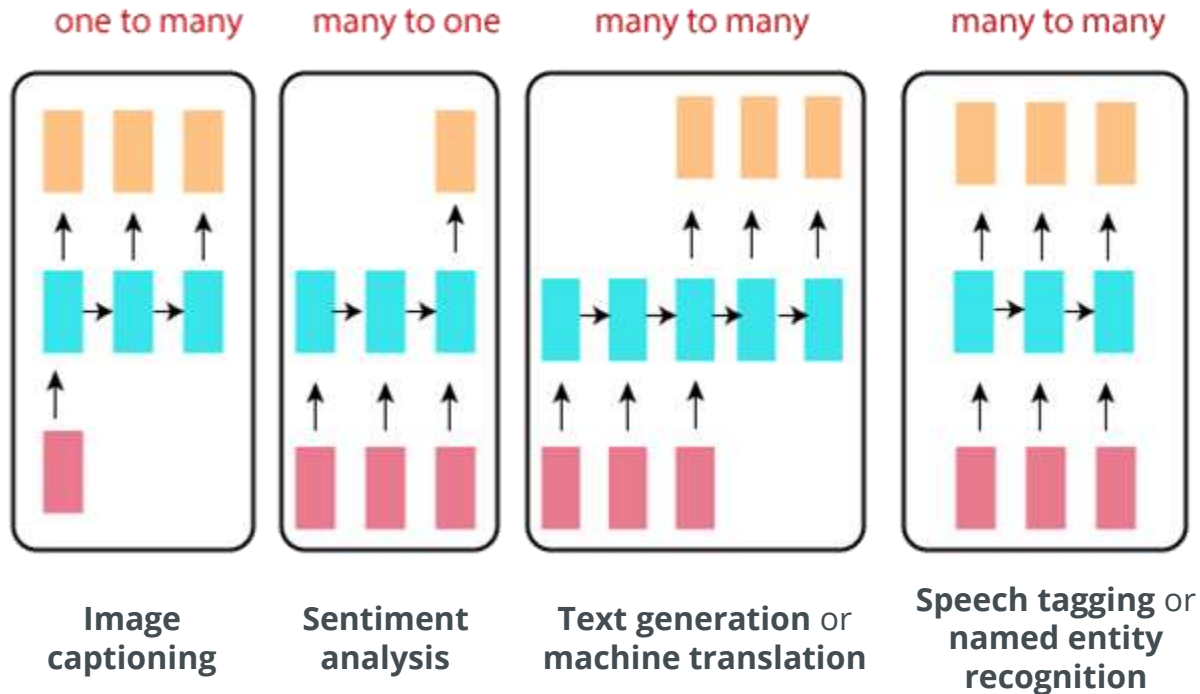
Multiple objects

Recurrent Neural Network (RNNs)

CNNs struggle with *time series* because they don't naturally keep track of the order of things. They treat input data as if all parts are independent and don't have a built-in way to remember what happened in the previous steps of a sequence.

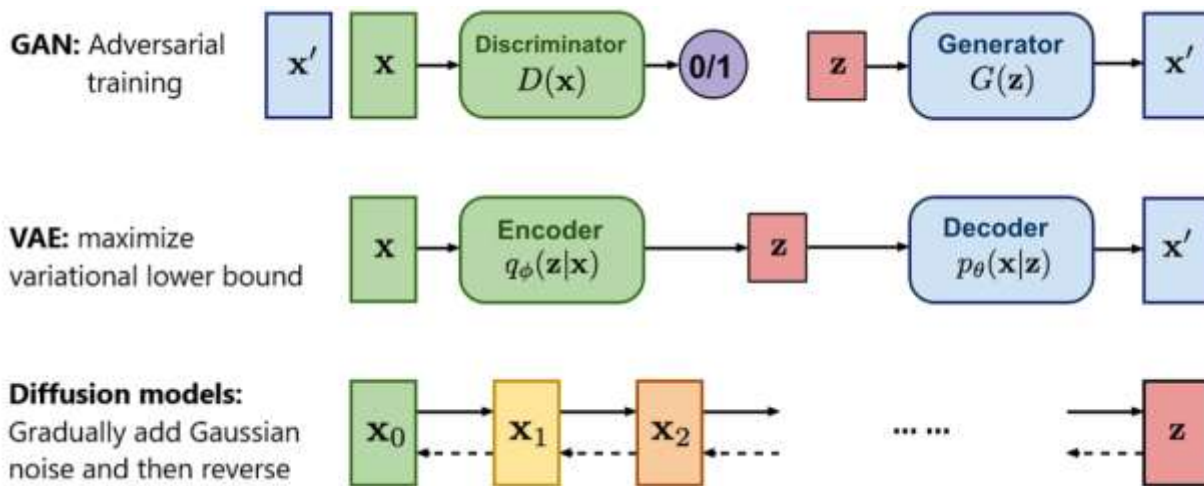


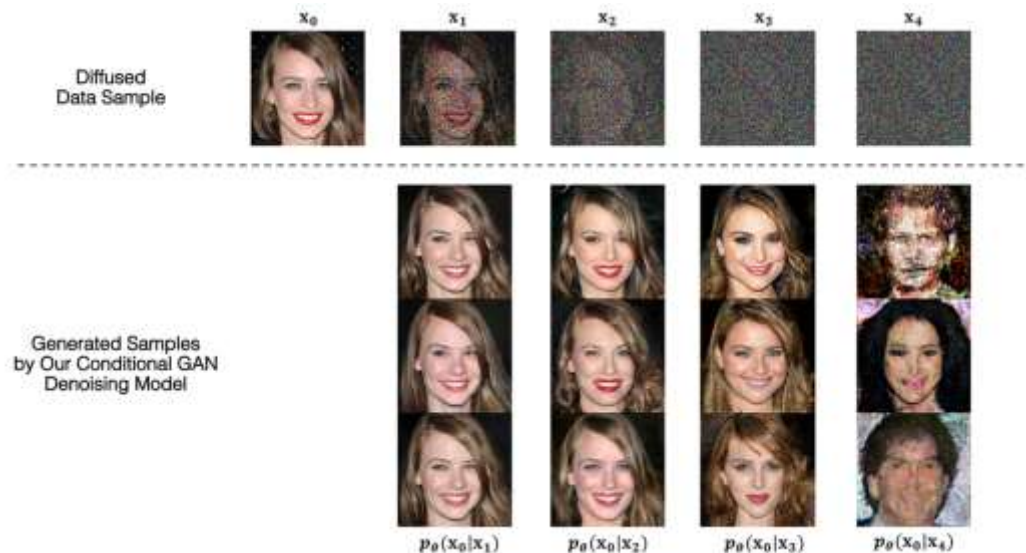




Generative Neural Network (GANs)

A class of models designed to generate new data that is similar to the training data they've been fed. They learn the underlying distribution of a dataset and then use this knowledge to produce new instances that could plausibly come from the same distribution.





The website *This Person Does Not Exist* was created in February 2019. It uses GANs, to generate highly realistic images of human faces of people who do not actually exist. <https://this-person-does-not-exist.com/en>

04.

Training

Here, we discuss the process of training machine learning models, including data preparation, model selection, and the use of algorithms to optimize model performance. We'll also cover strategies to avoid common pitfalls like overfitting.

Training

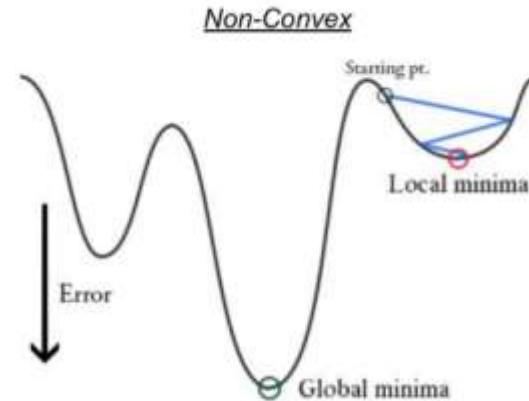
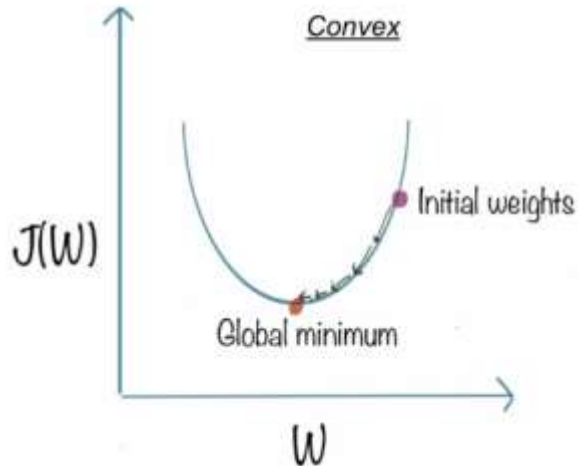
Gradient descent

Is an optimization algorithm that iteratively adjusts parameters to minimize a cost function, moving in the direction of steepest decrease.



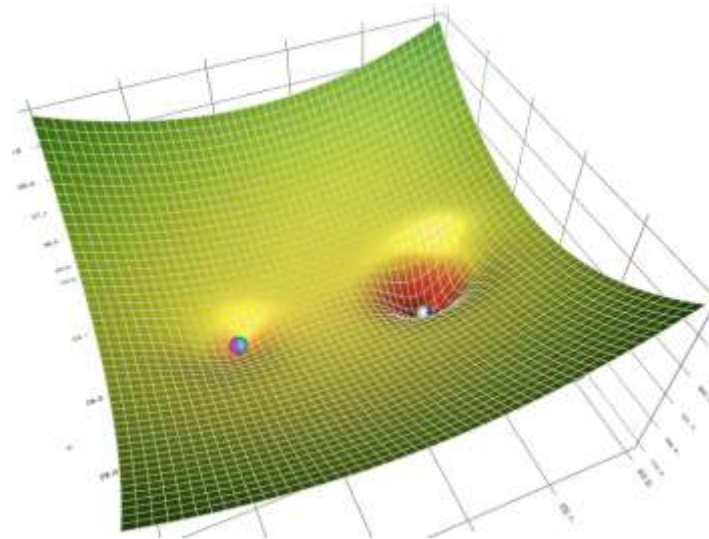
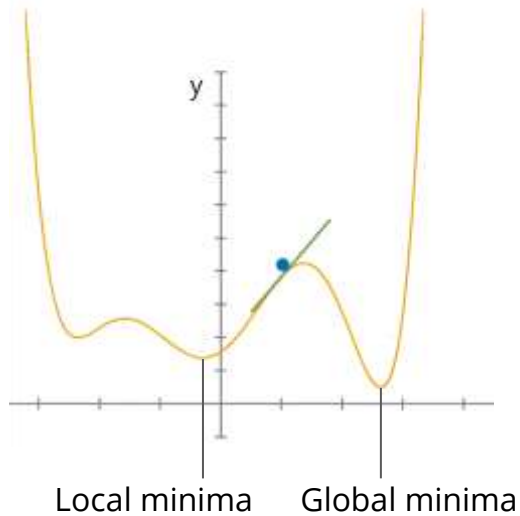
Cost (or loss) function

A mathematical function that measures the difference between the algorithm's predictions and the actual data. It guides the optimization process by quantifying the model's performance.



Training Gradient descent

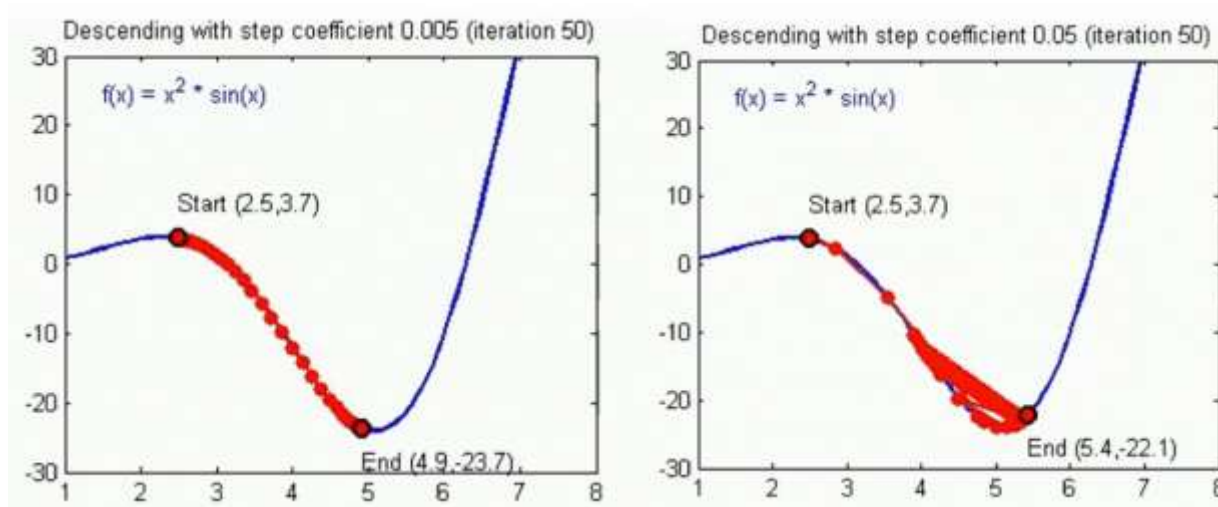
Is an optimization algorithm that iteratively adjusts parameters to minimize a cost function, moving in the direction of steepest decrease.



Training

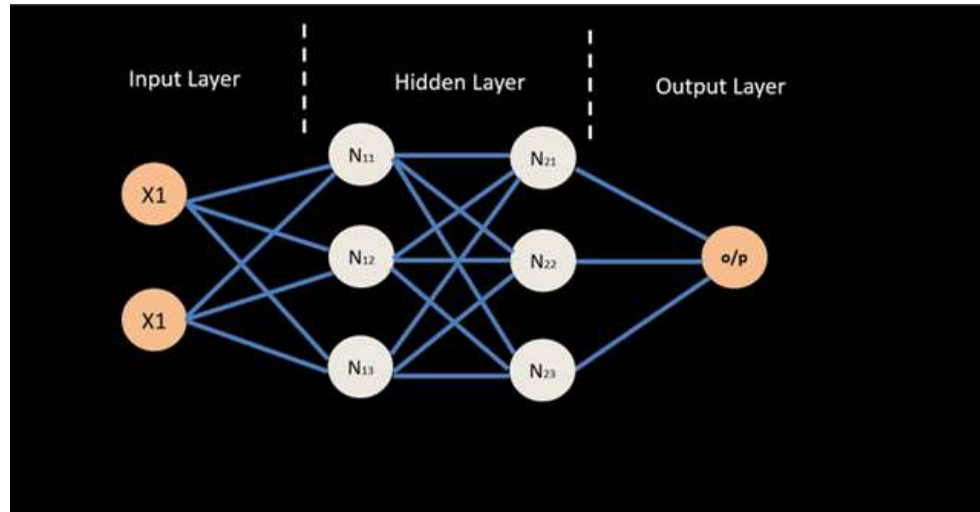
Learning rate

An hyperparameter that controls the adjustment of model weights during training. It determines the size of the steps the algorithm takes to reach the minimum of the loss function.



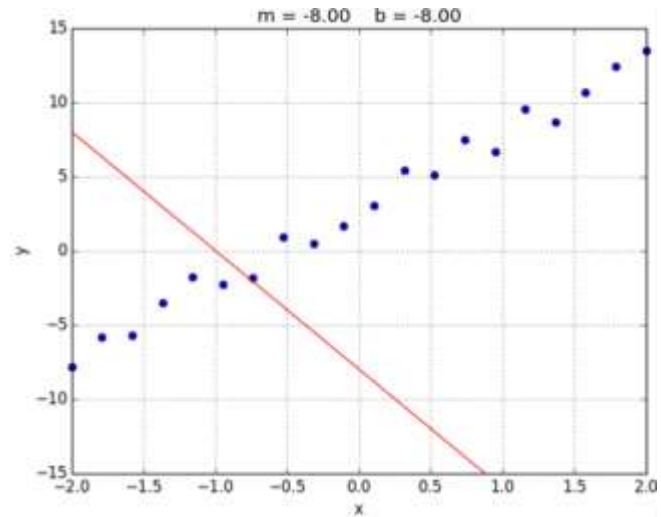
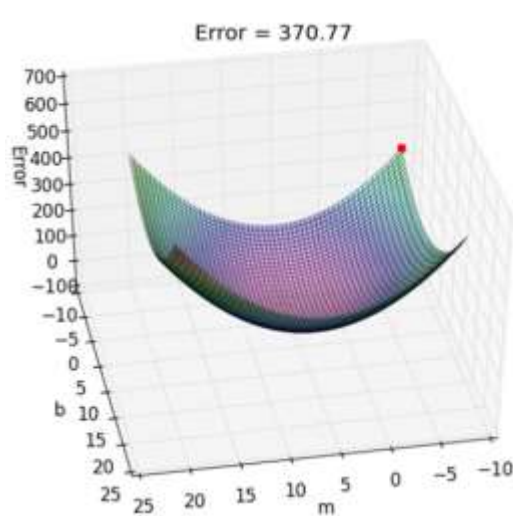
Training Backpropagation

An algorithm used to calculate the gradient of the loss function with respect to each weight by the chain rule, efficiently propagating the error backward through the network.



Training Backpropagation

An algorithm used to calculate the gradient of the loss function with respect to each weight by the chain rule, efficiently propagating the error backward through the network.

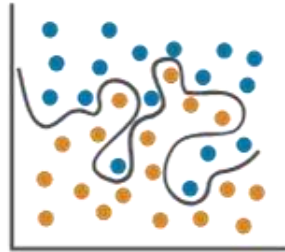


Training

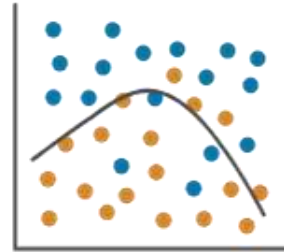
Overfitting and underfitting

Classification

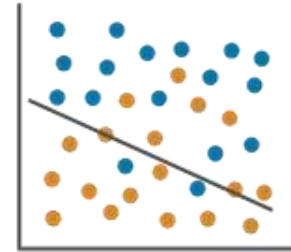
Overfitting



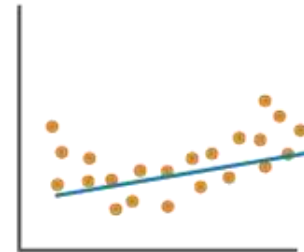
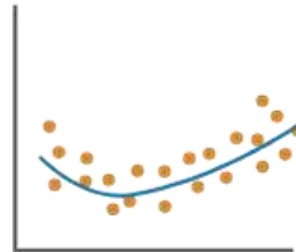
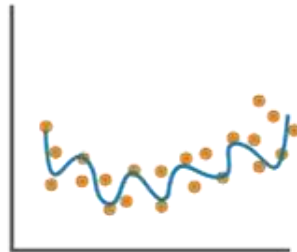
Right Fit



Underfitting



Regression



Training Generalization

Model's ability to perform well on new, unseen data after being trained on a specific dataset. It measures the effectiveness in applying learned patterns to novel inputs outside the training set.



Training data

Generalization
→

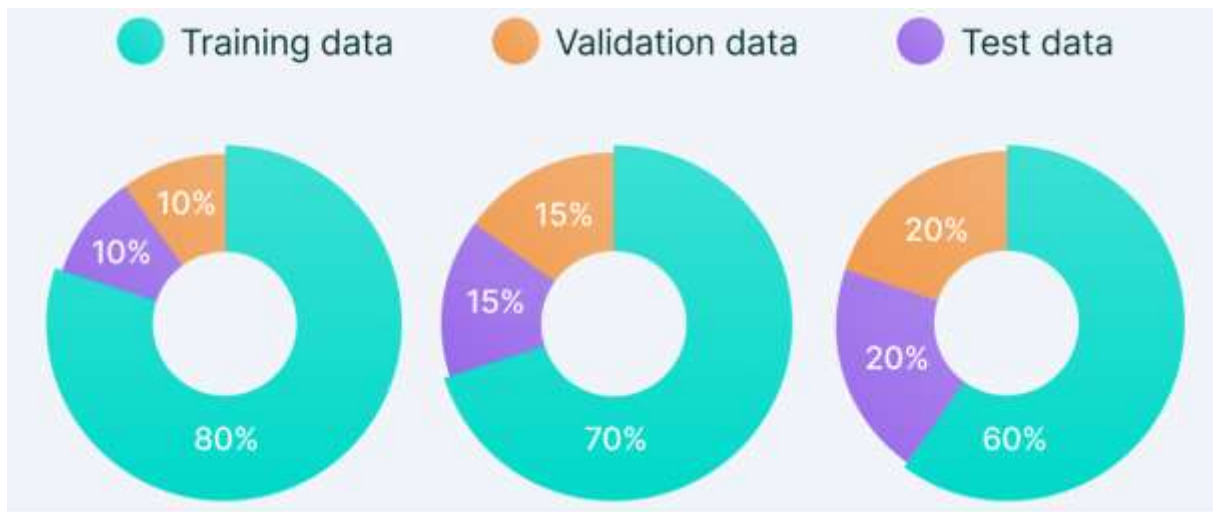


Testing data

Training

Generalization: *Set splitting*

A process a dataset is divided into separate subsets to ensure that models are trained on one set of data and tested on unseen data to evaluate performance and generalize ability.



Training

Generalization: ***Data augmentation***

A technique to increase the diversity of training data by applying various transformations, such as rotation, scaling, and flipping, to existing samples. This helps improve model robustness and generalization by simulating a wider range of input scenarios.



Original



Augmentation

Training

Generalization: *Regularization*

A technique used to prevent overfitting by adding a penalty on the size of the parameters. It encourages simpler models during training, which can generalize better on unseen data.



05.

Applications

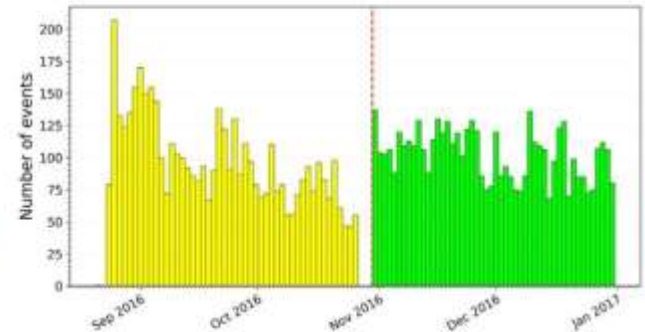
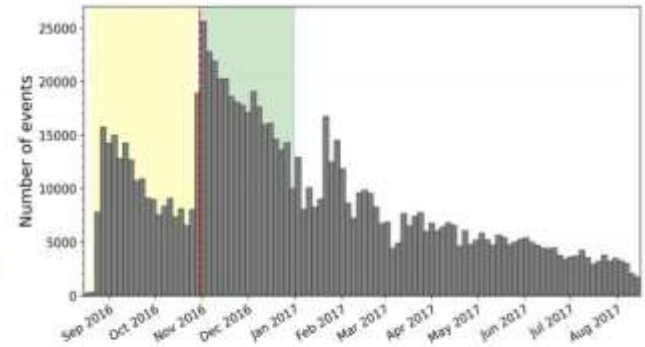
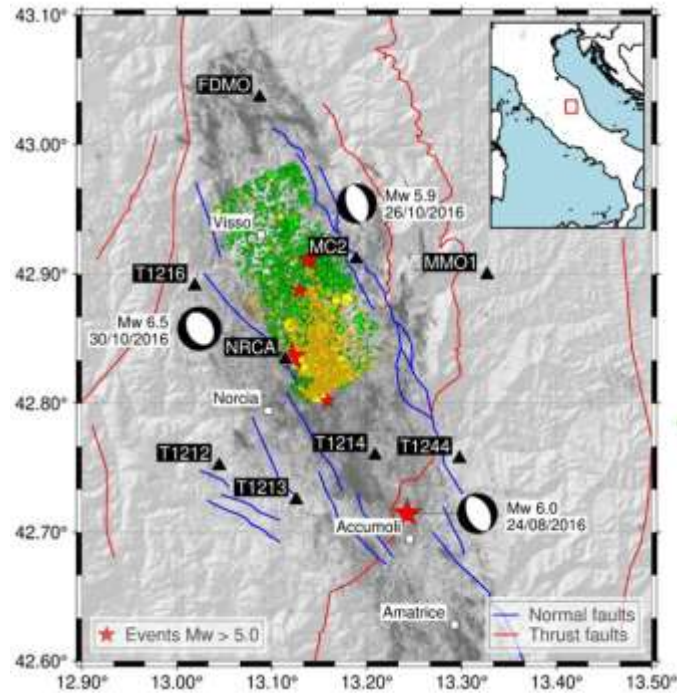
This section delves into the wide range of applications for machine learning and deep learning in seismology.

05.01

Earthquake classification

Applications

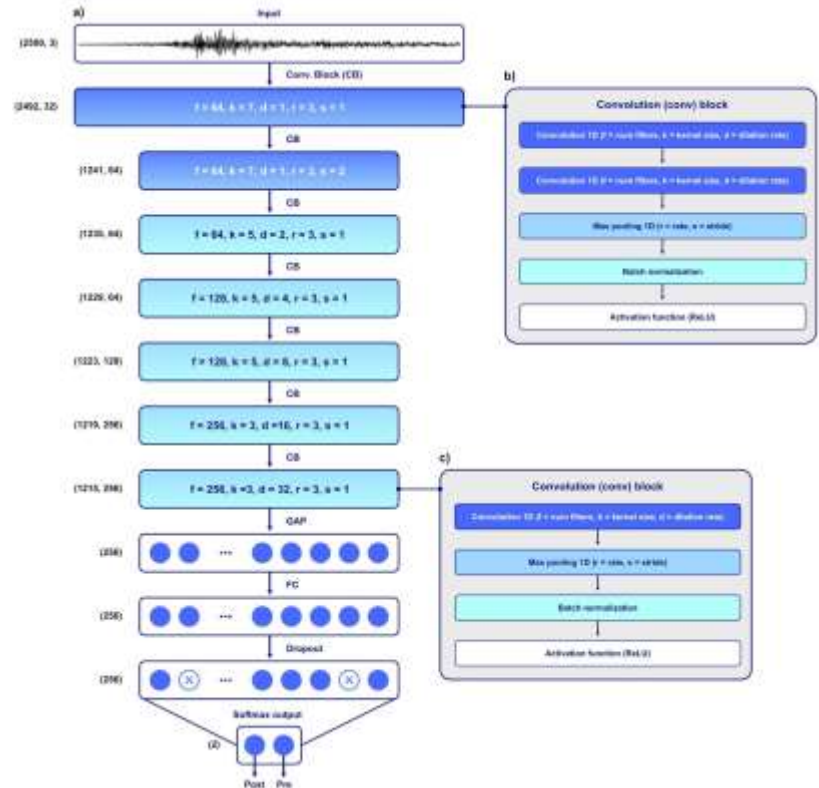
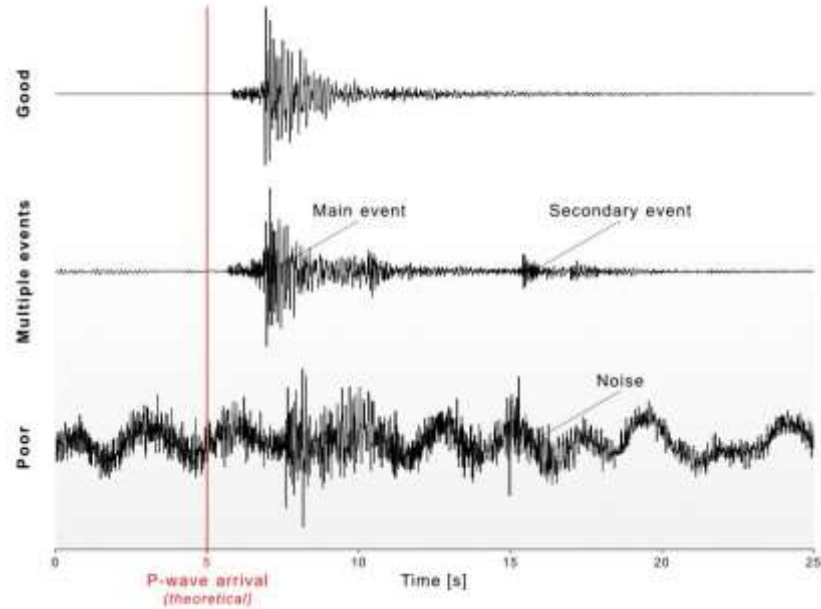
Earthquake classification



--- Norcia Mw 6.5 Tan et al. 2021 Foreshocks Aftershocks

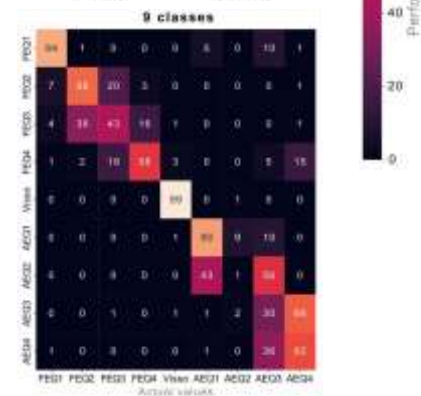
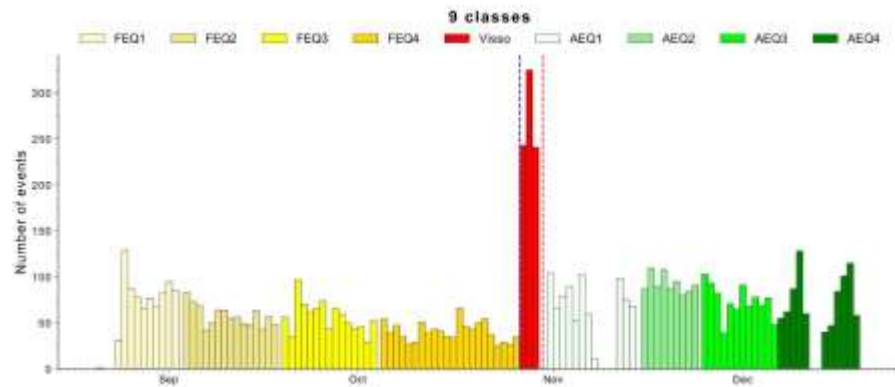
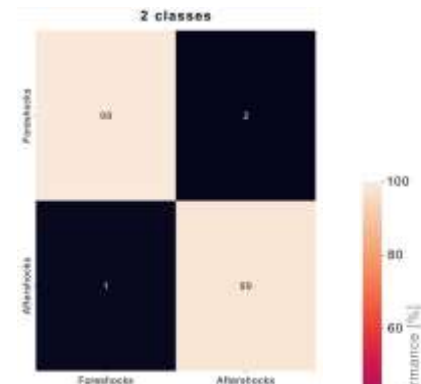
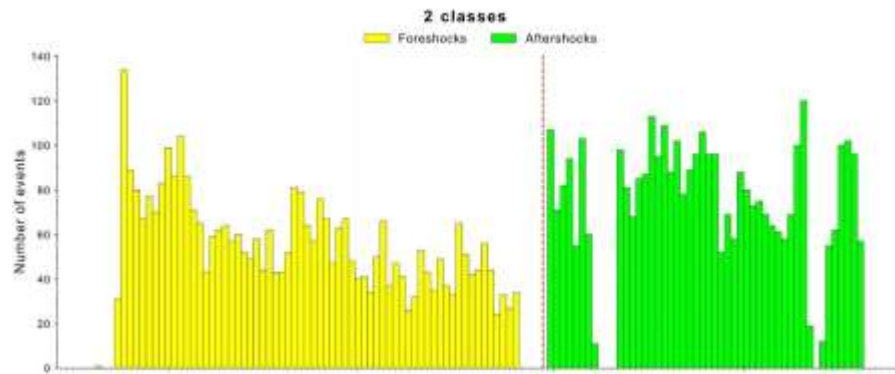
Applications

Earthquake classification: *Methods*



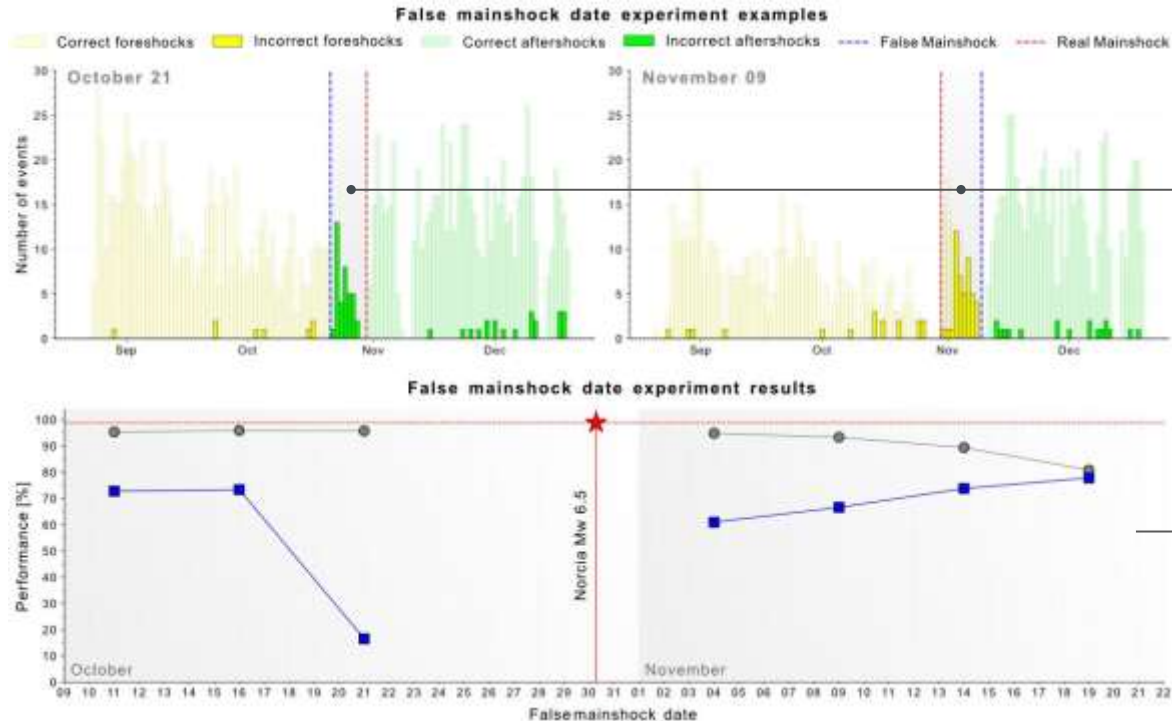
Applications

Earthquake classification: *Results*



Applications

Earthquake classification: *Validation*



Re-labeled
data

Results obtained
after re-labeling

05.02

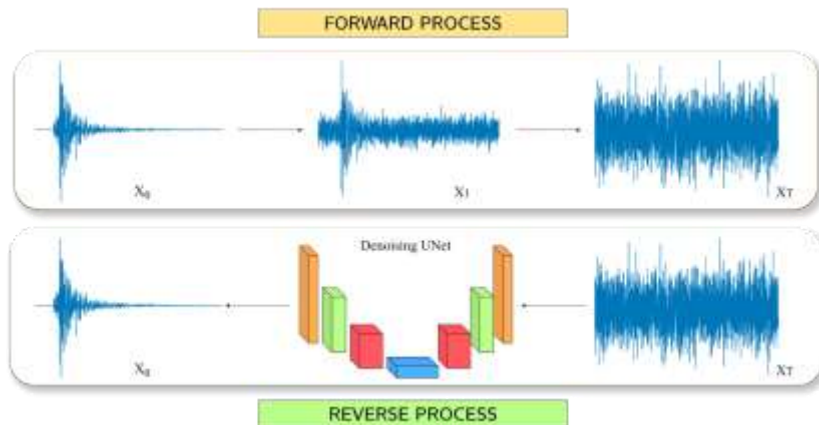
Cold Diffusion model

Applications

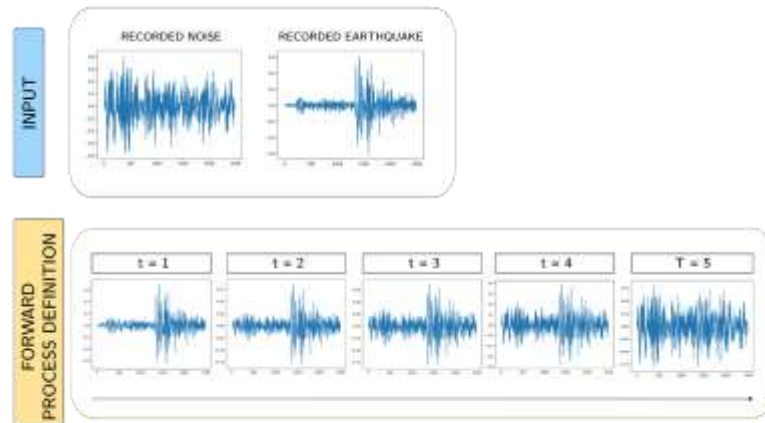
Cold Diffusion model: *Methods*

Our research explores adapting the Cold Diffusion model for seismic denoising, tailoring it to overcome the challenge of non-Gaussian noise in seismic data, promising enhanced signal recovery.

Diffusion Process



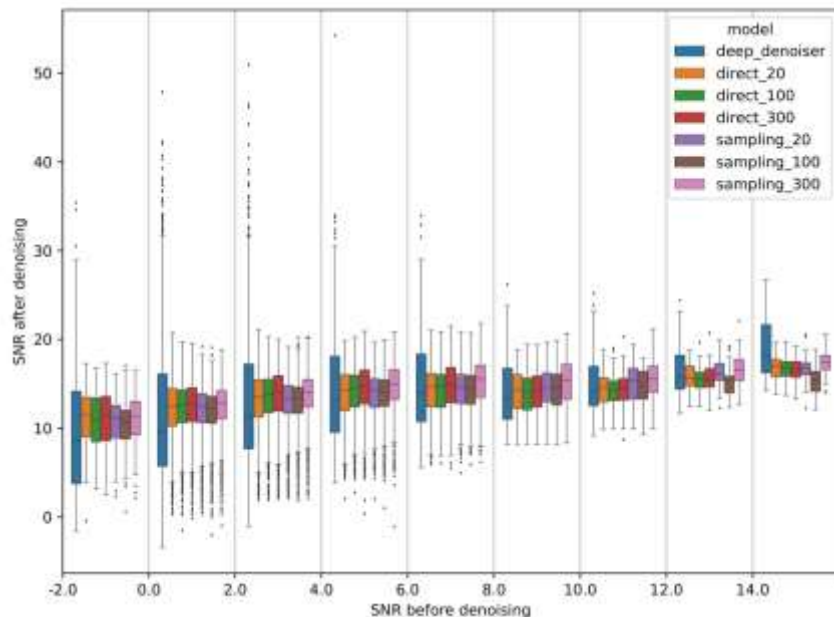
Input Assumptions



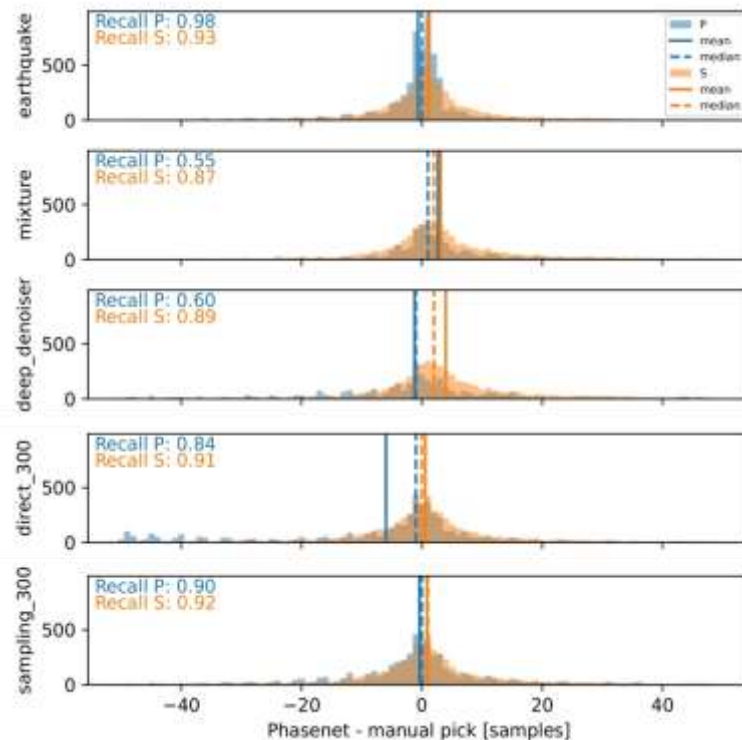
Applications

Cold Diffusion model: *Quantitative results*

**Distributions of SNR values
before and after the denoising**

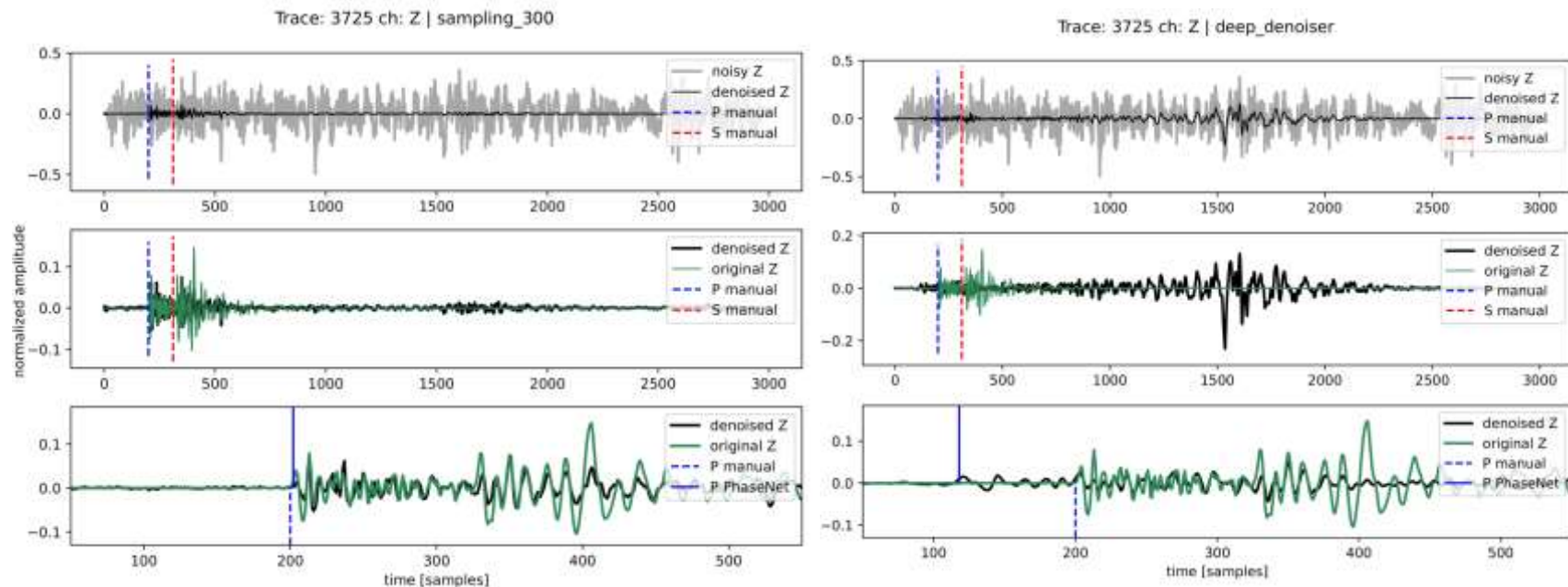


Phase arrivals picks



Applications

Cold Diffusion model: *Qualitative results*





Elements of Seismology & Machine Learning

Now it's your turn to apply machine learning to seismology!
If you have any ideas or questions, here are our contact details:

Gabriele Paoletti

gabriele.paoletti@uniroma1.it

Elisa Tinti

elisa.tinti@uniroma1.it

Daniele Trappolini

daniele.trappolini@uniroma1.it