Advanced School in Artificial Intelligence

Self supervised learning

Ing. Zese Riccardo

riccardo.zese@unife.it

Regione Emilia-Romagna

FERRARIAE UNIVERSITAS · EX LABORE FRUCTUS · 1391

Università degli Studi di Ferrara

Outline

- Introduction to Python

- Introduction to Neural Networks

- Convolutional NN

- Recurrent NN

- Autoencoders and self supervised learning

Università degli Studi di Ferrara

Outline

- Introduction to Python

- Introduction to Neural Networks

- Convolutional NN

- Recurrent NN

- Autoencoders and self supervised learning

Università
degli Studi
di Ferrara

Supervised vs Unsupervised Learning

Supervised Learning

Data: (x, y)
x is data, y is label

Goal: Learn a function to map x -> y

Examples: Classification, regression, object detection, density segmentation

Unsupervised Learning

Data: x
Just data, no labels!

Goal: Learn some underlying hidden structure of the data

A new Idea!

Self-supervised learning obtains supervisory signals from the data itself, often leveraging the underlying structure in the data.
The MODEL SUPERVISES ITSELF!

Autoencoders applies this idea

Università degli Studi di Ferrara

Self-Supervised Learning

- Practically speaking, it's impossible to label everything in the world.

- There are also some tasks for which there's simply not enough labeled data, such as training translation systems for low-resource languages.

- Common sense helps people learn new skills without requiring massive amounts of teaching for every single task.
  - For example, if we show just a few drawings of cows to small children, they'll eventually be able to recognize any cow they see.

- Self-supervised learning (SSL) is one of the most promising ways to build such background knowledge and approximate a form of common sense in AI systems.

Università degli Studi di Ferrara

Self-Supervised Learning

- The general technique of self-supervised learning is to predict any unobserved or hidden part (or property) of the input from any observed or unhidden part of the input.
  - For example, as is common in NLP, we can hide part of a sentence and predict the hidden words from the remaining words.

- Self-supervised learning is in some sense a type of unsupervised learning
  - No labels were given

  … but tries to solve tasks that are traditionally supervised (e.g., image classification) without any labelings available.

Università degli Studi di Ferrara

Example of training

- One of the main applications for SSL is the Natural Language Processing task.

- For example, the system is trained on short text (typically 1,000 words) in which some of the words have been masked or replaced.

- The system is trained to predict the words that were masked or replaced.

- In doing so, the system learns to represent the meaning of the text so that it can choose the "correct" words, or at least guess the most probable words that make sense in the context.

Università
degli Studi
di Ferrara

Which score?

- There are many possibilities, some of them strictly related to the data.

- One general possibility is to consider an an energy-based model, that takes two inputs and returns a score, the energy, telling how much the two inputs are related:
  - Low energy, inputs related
  - High energy, inputs unrelated

In computer vision

- For image recognition, the model takes two images.

- If one is a slightly distorted versions of the other, the model is trained to produce a low energy as output.

- Use of **Siamese networks**, composed of two identical (or almost identical) copies of the same network, each fed with a different image. The networks produce output vectors called **embeddings**, representing the images. Then, a third module takes the two embeddings and computes the energy as the distance between the two embedding vectors.

In Natural Language Processing

- Usually, NLP do not use the joint embedding architecture.

- Use of a predictive architecture in which the model directly produces a prediction for a sentence.

- Given a sentence, first it is corrupted, e.g., by masking some words to produce a second (corrupted) sentence.

- The corrupted input is fed to a neural network trained to reproduce the original text.

- Reconstruction error as the energy:
  - low energy for "clean" text
  - higher energy for "corrupted" text.

Università degli Studi di Ferrara

Interesting readings

- https://amitness.com/2020/05/self-supervised-learning-nlp/

- https://jonathanbgn.com/2020/12/31/self-supervised-learning.html

- https://towardsdatascience.com/self-supervised-keypoint-learning-aade18081fc3

Università degli Studi di Ferrara

Is it the same thing of reinforcement learning?

- In many cases there is the need of other techniques to help generating examples (e.g., genetic algorithms)

- It is similar to some extent to the reinforcement learning task, where the «label» (supervision) is given by the environment.

Università
degli Studi
di Ferrara

Is it the same thing of reinforcement learning?

- Reinforcement learning is learning what to do — how to map situations to actions—so as to maximize a numerical reward signal. The learner is not told which actions to take, but instead must discover which actions yield the most reward by trying them.

- There is no fixed training dataset but one or more objectives to achieve, actions they may perform, and feedback about performance toward the objective.
  - https://www.youtube.com/watch?v=CI3FRsSAa_U

Università degli Studi di Ferrara

Is it the same thing of reinforcement learning?

- Self-supervised learning refers to an unsupervised learning problem that is framed as a supervised learning problem in order to apply supervised learning algorithms to solve it.

- A common example of self-supervised learning is computer vision where a corpus of unlabeled images is available and can be used to train a supervised model, such as making images grayscale and having a model predict a color representation (colorization) or removing blocks of the image and have a model predict the missing parts (inpainting).

Università degli Studi di Ferrara