



Milano R meeting

# MACHINE LEARNING

---

A case of digit recognition based on a shallow neural network implemented in R

Michele Vitali

Milano, 18 Aprile 2013

# Agenda

- Machine learning
- Digit recognition problem and data set
- Artificial neural networks and shallow neural networks
- Description of the experiment
- Architecture in R
- A little bit of R
- Results
- Potential improvements

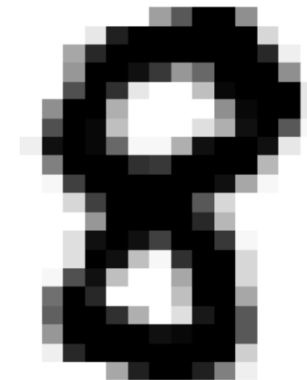
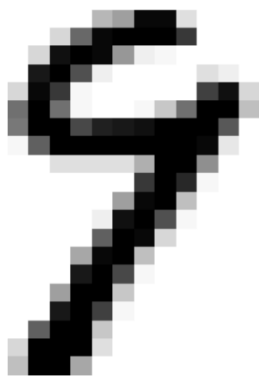
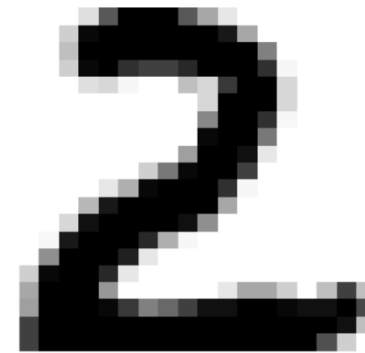
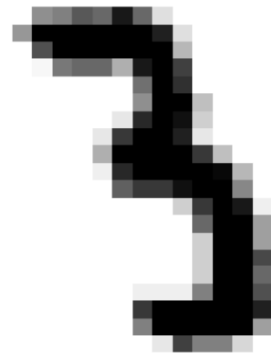
# Machine learning

**Machine learning**: a scientific discipline that is concerned with the design and development of algorithms that allow computers to **evolve behaviors** based **on** empirical **data** without being explicitly programmed

Two types of ML:

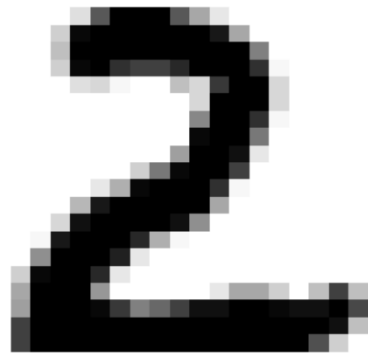
- **Supervised learning**: the machine learning task of inferring a function from **labeled** training **data**
- **Unsupervised learning**: the problem of trying to find hidden structure in **unlabeled data**

# Digit recognition problem



# Digit recognition problem

**Digit recognition:** automatically classify handwritten numbers from 0 to 9



Darkness of each pixel:

- 784 ( $28 \times 28$ ) values from 0 to 255

Digit written in the picture (label)

# Digit recognition problem

## Data set

- 42000 handwritten numbers with the associated label

## Training set

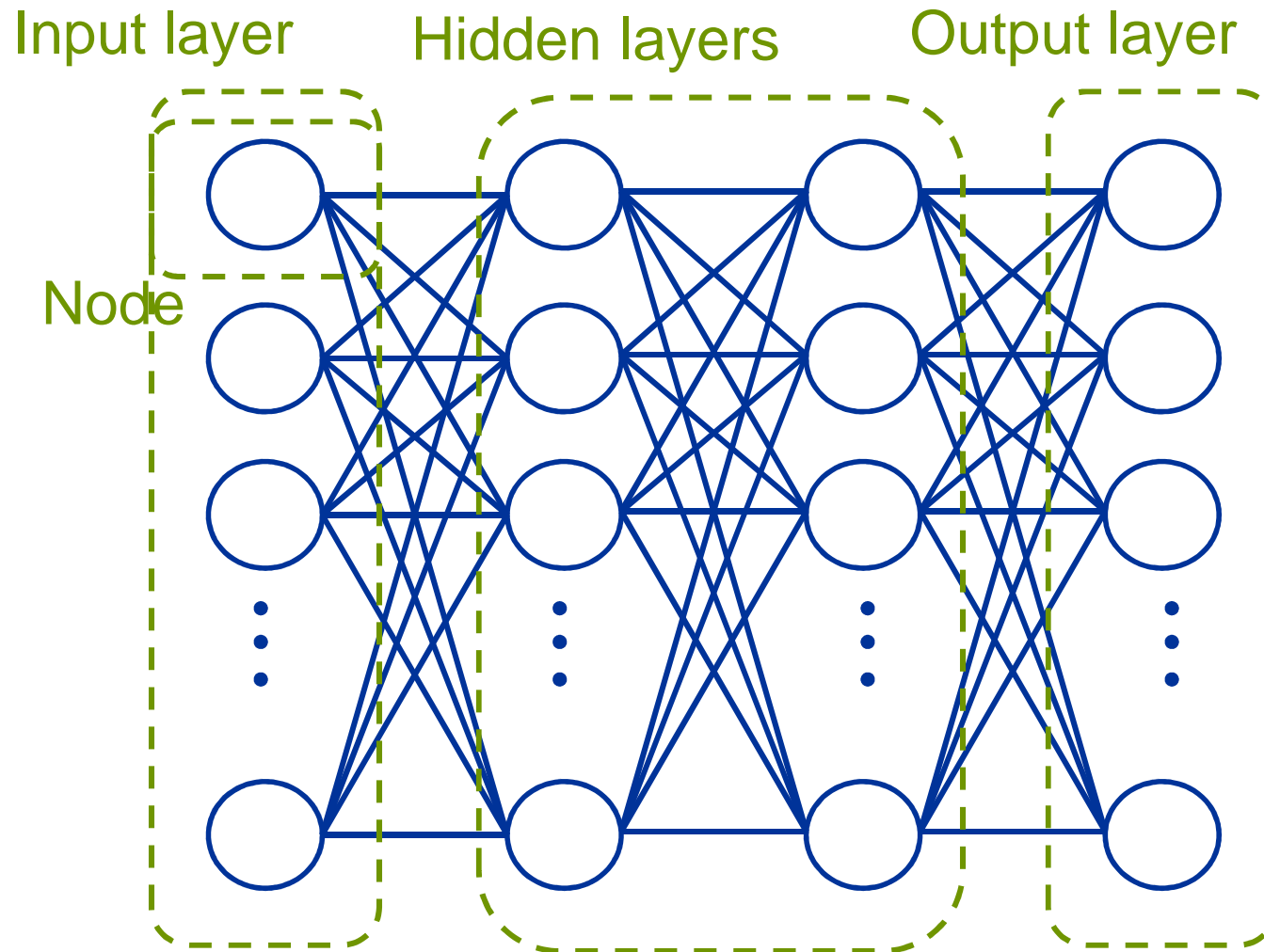
- 25200 records (60%)

## Cross validation set

- 8400 records (20%)

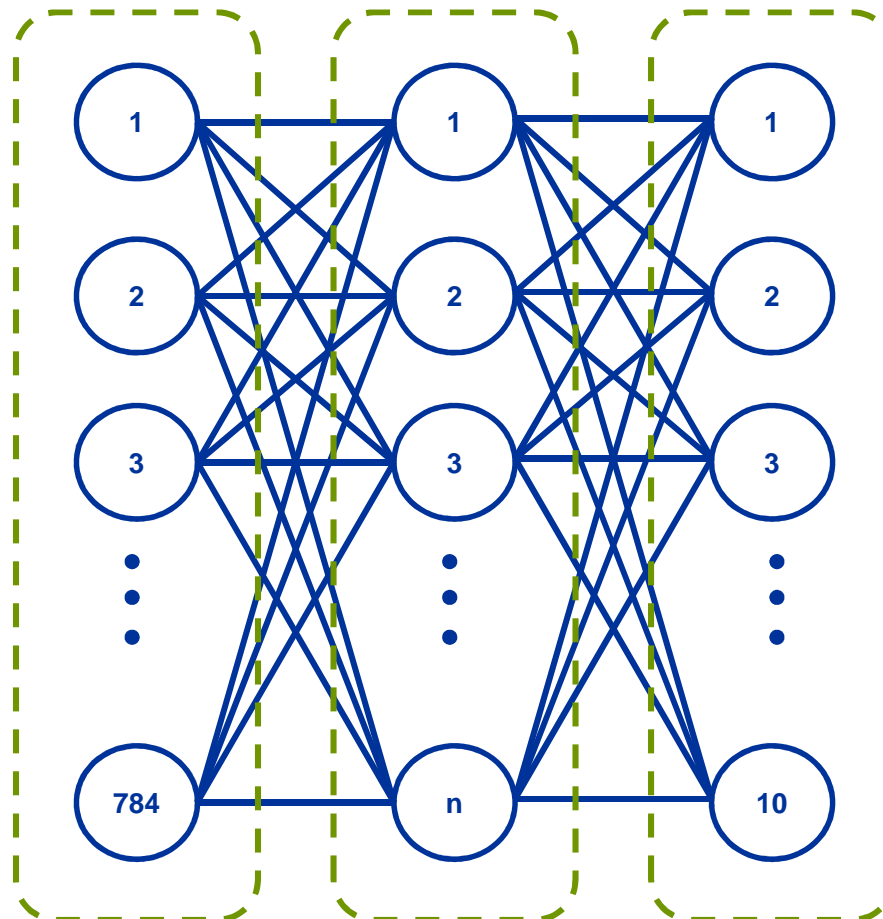
## Test set

# Artificial neural networks



# Shallow neural networks

Image pixel    Probability estimation





# Description of the experiment

## Objective of the experiment:

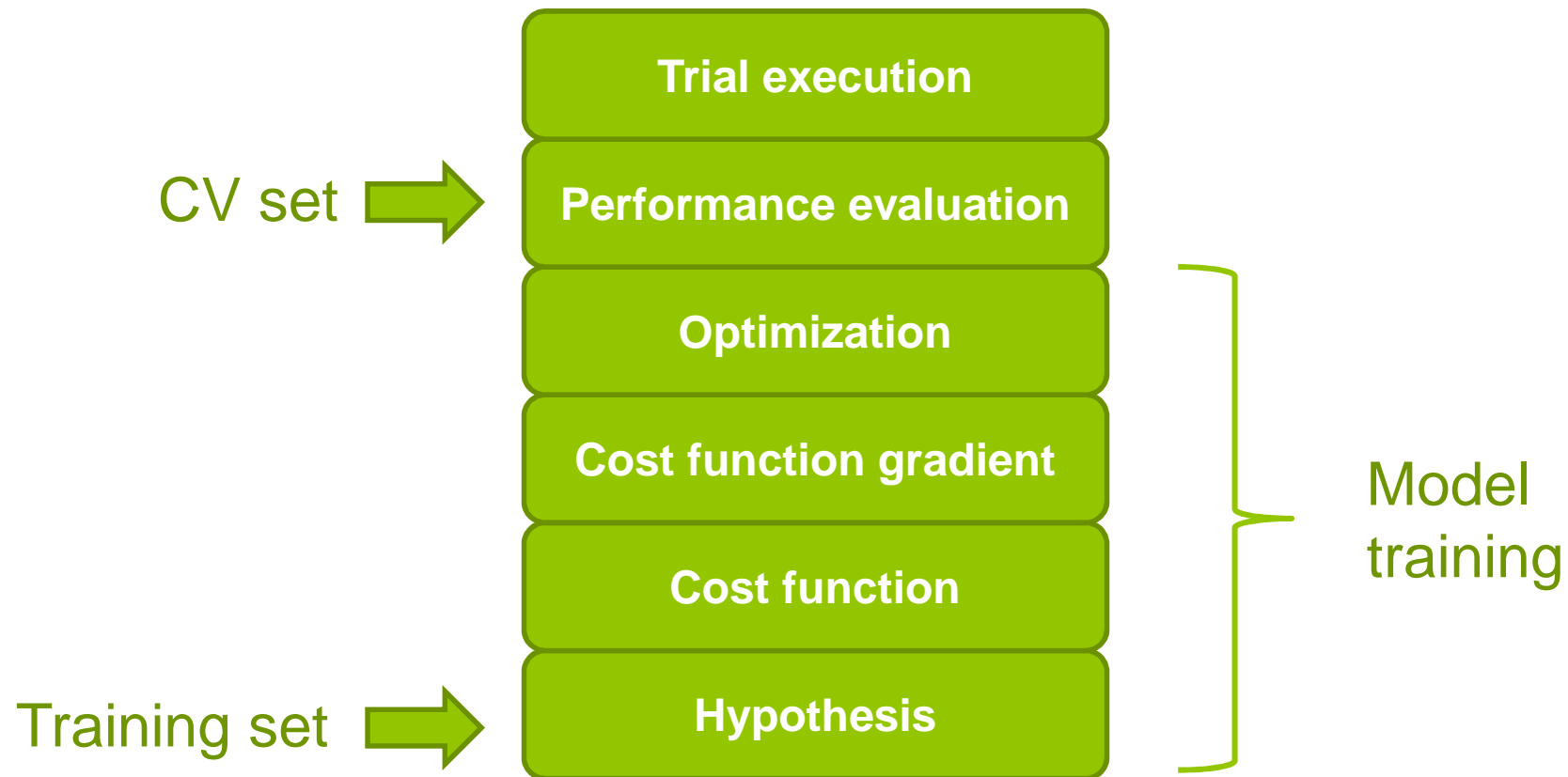
- Observe model performance for different dimensions of the hidden layer

## Measurement of model performance

- Train the model on the training set
- Make estimation on the cross validation set
- Model performance: percentage of correct classification on the cross validation set

# Architecture in R

The stack of functions:

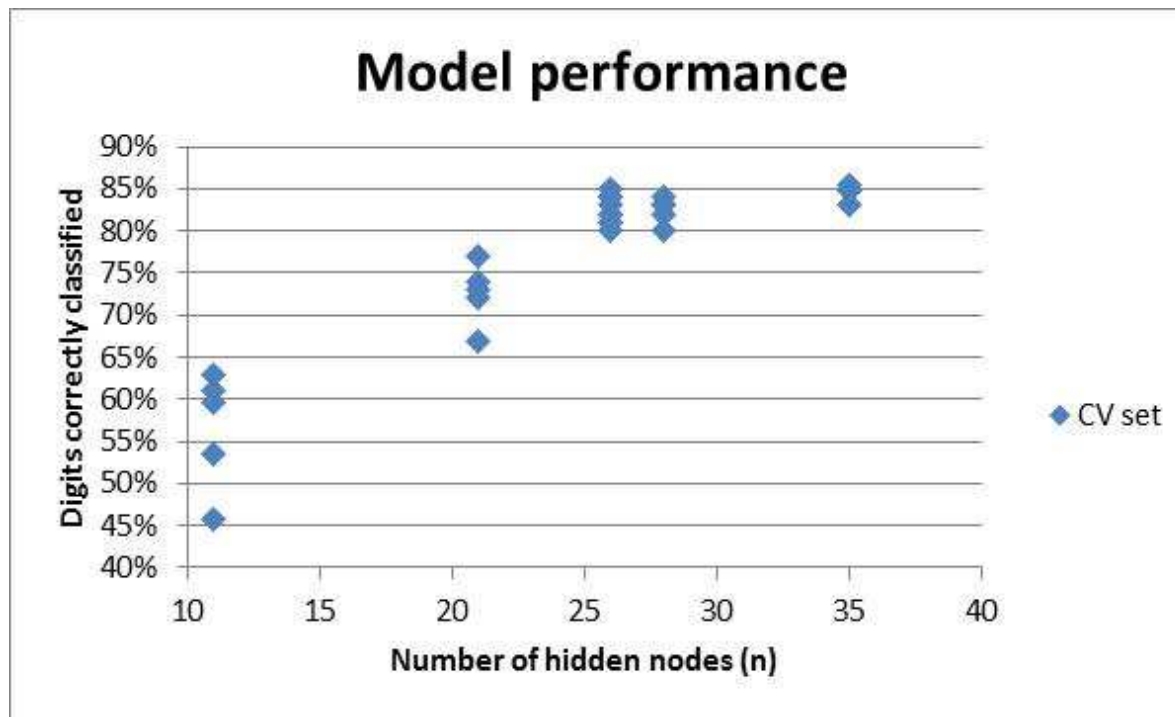


# A little bit of R



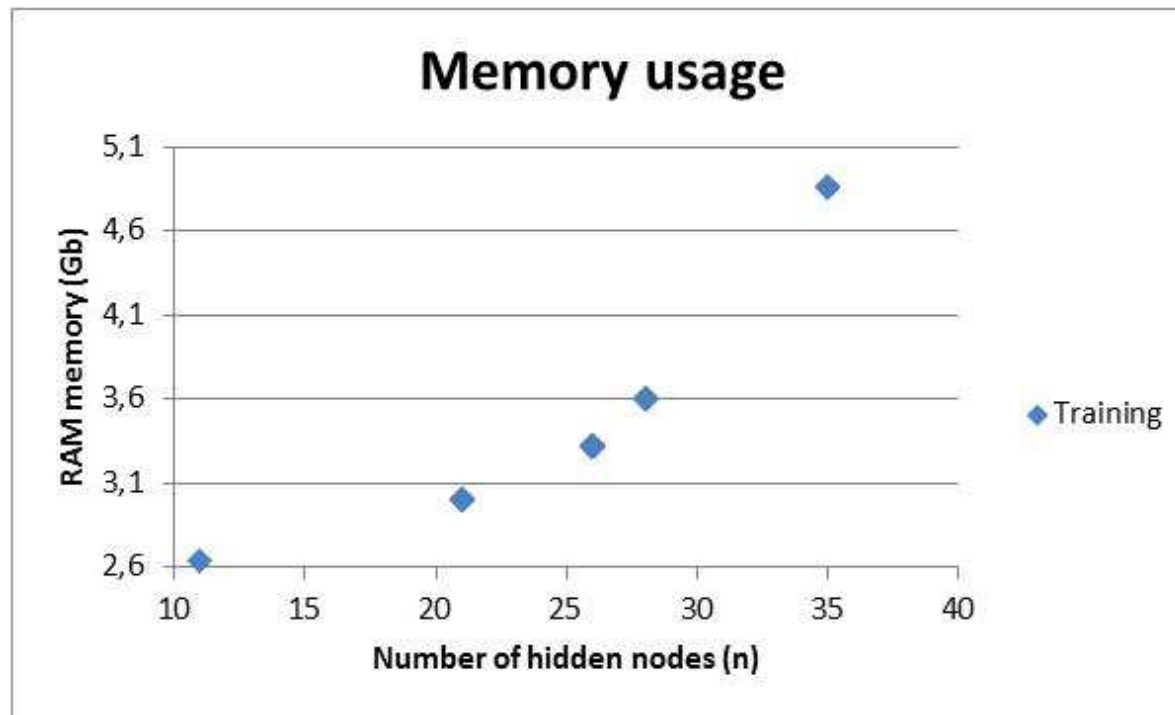
# Results

- Performance improves with the increase of the number of hidden nodes and then flattens at around 85%



# Results

- Usage of RAM memory increases more than linearly with the increase of the number of hidden nodes



# Potential improvements

**Deep neural networks** with more than one hidden layer can model high-level features

Better results are expected: even more than 95% of correct estimations

More complex models and optimization algorithms need to be implemented

More expensive in terms of RAM memory and computation

# Thank you for your attention!



- Michele Vitali
- e-mail: [michele.vitali.it@gmail.com](mailto:michele.vitali.it@gmail.com)

• Follow me on:

