

AA 2022/2023

Machine Learning for Modelling: *Supervised Learning*

Simone Bianco

1

AA 2022/2023

Recap

2

L01

- **Basic concepts:** model, data set, feature vector, features, dimension, attributes, instance, labels, generalization, validation, testing, etc.
- **Popular learning algorithms:**
 - LDA
 - Decision Trees
 - Neural Networks
 - K-Nearest Neighbor
 - SVM
- **Algorithms evaluation and comparison:** t-times k-fold cross validation, McNemar and Friedman test, critical difference diagram

3

L02

- **Ensemble methods:**
 - Boosting
 - AdaBoost, multiclass AdaBoost, AdaBoost for regression
 - Bagging
 - Random Tree ensembles
 - Random Forest
 - Random Forest for density estimation
 - Random Forest for anomaly detection
 - Stacking classifiers

4

L03

- **Object instance recognition:**

- Template matching
- Local descriptors
- SIFT
- Ransac

- **Evaluation metric for object detection:**

- Intersection o Union (IoU)
- mean Average Precision (mAP)

- **Applications:** panorama stitching, video stabilization, tracking, Augmented Reality

- **Bag of Words (BoW)**

5

L04

- **Homogeneous coordinates**

- **Viola-Jones Object Detection Framework:**

- Haar-like features
- Integral image
- AdaBoost algorithm
- Cascade classifier

6

L05-L06

- **Convolutional Neural Networks:**

- MLP
- CNN basic blocks
- CNN training
- Diagnose training and techniques to prevent overfitting
- Famous CNN architectures

- **Transfer learning:**

- CNN fine-tuning
- CNN as feature extractor

7

L07

- **Recurrent and Recursive Neural Networks:**

- Computational graph
- Deep RNNs
- Training a RNN (BPTT)
- Vanishing gradients

- **Gated RNN:**

- Long Short-Term Memory (LSTM)
- Gated Recurrent Unit (GRU)

8

L08

Object detection:

- Detection as regression
- Detection as classification
- Region proposals

Two-stage object detection:

- Region-based convolutional neural networks
- R-CNN
- Fast R-CNN
- Faster R-CNN

9

L09

Single-shot object detection:

- YOLO (main concepts, unified detection, network design, training, inference)
- YOLO timeline (YOLO-v2, v3, v4, v5, v6, v7)

10

L10

Transformers:

- Dot-product self-attention
- Values, queries, and keys
- Matrix form
- Positional encoding
- Scaled dot-product self-attention
- Multiple heads
- Transformer layers

Transformers for NLP: Bert, GPT3

Transformers for images: ImageGPT, ViT, Swin, DaViT

11

L11

Self-Supervised Learning:

- Terminology: pretext tasks, downstream tasks, etc.
- Learning visual features from pretext tasks
- Commonly used pretext tasks:
 - Generation-based
 - Context-based
 - Free semantic label-based
 - Cross modal-based
- Commonly used downstream tasks for evaluation

12