AA 2023/2024

Machine Learning for Modelling:
Supervised Learning

Simone Bianco

AA 2023/2024

Exam Project

Valid for June/July/Sept 2024 and Jan/Feb 2025 sessions

Task and dataset

- **Task:** Image classification using SIFT/Bag of words coupled with traditional classifiers and CNNs

- **Dataset**: A variant of XYZ dataset (TBD)
    - input size 64x64
    - ~100 classes
    - ~1000 images per class
    - Train, Val, Test splits: Test will not be used, your test set is the validation set
    → validation set must be extracted from the training set

Tasks

1. Data Preprocessing: The student should preprocess the images by resizing them to a common size, converting them to grayscale or RGB, and normalizing the pixel values.

2. Feature Extraction using SIFT and Bag of Words: The student should extract features from the images using Scale-Invariant Feature Transform (SIFT) (or similar approaches) and Bag of Words (BoW) representation and train a traditional classifier (e.g., Support Vector Machines (SVMs), Random Forests, or k-Nearest Neighbors (k-NN)) to classify them.

3. Convolutional Neural Networks: The student should train a convolutional neural network (CNN) on the preprocessed images to classify them into different categories. The student should experiment with different architectures such as VGG, ResNet, Inception, etc.

4. The students have to cast the problem as supervised learning and self-supervised learning.

5. Model Evaluation: The student should evaluate the performance of the models using metrics such as accuracy, precision, recall, and F1-score.

6. Hyperparameter Tuning: The student should fine-tune the hyperparameters of the models to achieve better performance.

The project should include a written report that describes each step of the project in detail, presents the results, and discusses the limitations and future directions of the project. The report should also include visualizations to support the analysis.

What must be delivered

- All Python/Matlab software codes associated with the Tasks, use the standard practice you learnt during the lab lectures to describe and comment the software code.

- A technical report consisting of no more than 8 pages (references excluded) where you describe how you analyzed the given dataset. The template to be used is available on the course page in the "Exam" section

What is allowed

- You can develop the project yourself or you can develop the project with one of your colleagues

- You can access any source of high quality to develop your project, other courses, papers, books, ... but it must be properly cited

What is not allowed

- No plagiarism in any form, including chatGPT and similar (you must add a statement to your report that clearly exclude the use of any form of plagiarism)

5

Deadlines

When it must be delivered

- On Friday before the exam (usually), deadline at 23:59:
  - 18-06-2024 for the 21-06-2024 exam
  - 19-07-2024 for the 25-07-2024 exam
  - 13-09-2024 for the 16-09-2024 exam

How it must be delivered

- Write an email to me with the link to a shared folder on Google Drive

- After the deadline no modifications are allowed to the code and the report

- Only (small) changes to the presentation are allowed (e.g., typos correction, etc.)

6

Assessment

- Ongoing Exam (Max 33 points):
  - Lab reports   10 pts
  - Project   12 pts
  - Interview   11 pts

- Full Exam (Max 30 points):
  - Project   20 pts
  - Interview   10 pts

7

Project grading

- **Software code**   (**up to 6 points**)
  - Correctness   (up to 2 points)
  - Efficiency   (up to 2 points)
  - Clarity of comments   (up to 2 points)

- **Technical report**   (**up to 6 points**)
  - Correctness, rigor, accuracy of the methodological approach   (up to 2 points)
  - Clarity of expression and communication of ideas   (up to 2 points)
  - Tables, figures, captions, are they appropriate? Well motivated? Explained?   (up to 2 points)

8

Interview and rejecting the grade

Interview

— about 20 minutes

- 10 minutes for the oral presentation of the project, slides to be used, both members must present

- 10 minutes of interview on methodological aspects of the project and of what presented in the course

Rejecting the Grade

- A new version of the project, substantially different from the previous one (even if on the same task) must be prepared and submitted

- Lab Reports expire after 1 year (they are valid until February 2025 exam session)

- New interview

9