# VU Machine Learning

Summer semester 2023

# Exercise 2

Nysret Musliu (nysret.musliu@tuwien.ac.at)
Andreas Krystallidis (andreas.krystallidis@tuwien.ac.at)

# Exercise 2

- Groups of 3 students
- Implement techniques for classification
- Compare to existing techniques
- Submit the source code
- Prepare around 10-20 slides
- Individual discussions (30 min) for each group (all members must be present)
- Submission: May 29
- Discussions: after May 29 (slots will be available in tuwel)

# Exercise 2 – Techniques

- Implement a framework for neural networks
  - Implement standard forward and backward propagation methods
  - Investigate and experiment with different configurations for the NN including two activation functions, different number of layers and number of nodes per layer
    - Implement also a simple method to automatically find good values for these hyperparameters
  - Your algorithm should work for classification data sets that include nominal or/and numeric attributes (features)
- You should implement the methods from scratch
- Please do not use any part of existing code
- You can use existing code/functions for general parts like
  - Mathematical calculations (derivatives…), code for reading the input and testing the algorithm (cross- validation, performance metrics…)

# Exercise 2 – Data Sets

- Pick 3 classification data sets
  - Data sets from the previous assignments can be used
  - Or other data sets from UCI ML Repository, Kaggle…
- Should have different characteristics
  - number of samples – small vs. large
  - number of dimensions – low vs. high dimensional
- Pre-process the data set if needed (scaling, missing values …)

# Comparison

- Compare your implemented techniques on selected datasets with
  - Existing NN implementation
    - You can use the default parameters for the existing technique
  - + Another classification technique
- Apply cross-validation
- Conclusions
  - How efficient is your algorithm
  - Performance of your algorithm regarding performance metrics for classification
  - Impact of hyperparameters
  - Other findings

# Submission

A zip file with

- **Source code:**
  - You can use any programming language: Python, Matlab, R…
  - Provide the information for the packages needed to run you code
- **Slides**
  - Around 10- 20 slides
  - No report needed
- Submission deadline: May 29, 12:00

# Slides

- A brief discussion of implementations
- Comparison with the existing implementations/other algorithms
- Discussion of experimental results
- Conclusions/lessons learned

# Individual discussions

- Length of discussion: 30 minutes
- You will have questions about
  - Source code
  - Techniques
  - Comparison with the existing techniques
- All members of the group should be able to explain the code/experiments
- The evaluation will be based on your code, discussion, comparison, and conclusions/lessons learned