

Machine Learning and Data Mining

Intro and Course overview

Majid Sohrabi

National Research University Higher School of Economics



MMCP

Sept 04, 2024

Contacts

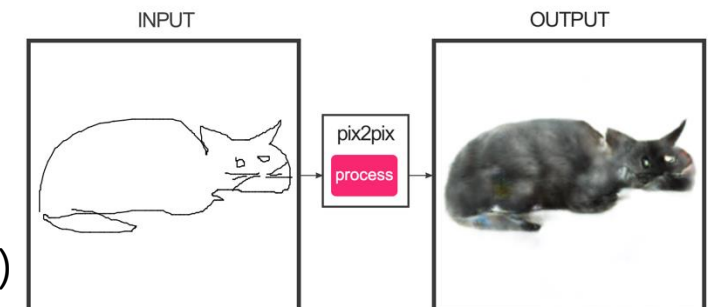
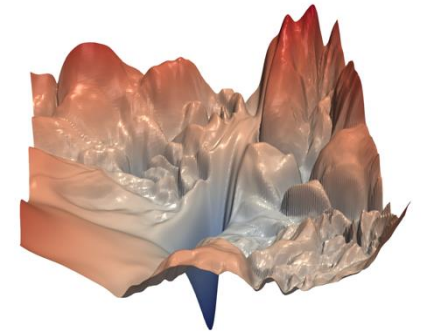
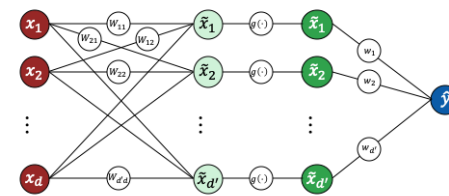
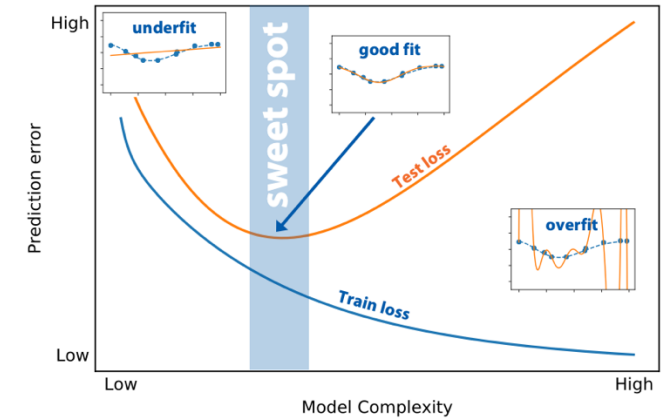
- ▶ Majid Sohrabi (lectures)
 - Email: msohrabi@hse.ru
 - Telegram: @MSohrabi_CS
- ▶ Alexandra Kogan (seminars)
 - Email: akogan@hse.ru
 - Telegram: @horror_in_black

Repository with course material

<https://github.com/Majid-Sohrabi/MLDM-2024/>

Course content

- ▶ Introduction to classical machine learning
 - Basic algorithms and their application (metric methods, linear algorithms, SVM and kernel trick)
 - Bias-variance error decomposition, regularization techniques
 - Ensembling, bagging, boosting
- ▶ Introduction to deep learning
 - MLP, backpropagation, optimization techniques
 - Regularization layers, CNNs
- ▶ Advanced techniques
 - Deep generative models, notable architectures
 - Bayesian optimization, bayessian NNs, ELBO
 - Misc. topics (imbalanced datasets, importance sampling, etc.)



Overview

- ▶ Compulsory course for ISSA spec. year 2
- ▶ Elective for other specializations, year 1
- ▶ Inter-campus course
 - The groups are merged (classes take place simultaneously)
- ▶ Duration: 1st half of the academic year (modules 1 and 2)
- ▶ Assessment elements:
 - Homework assignments (50% weight)
 - Exam (50% weight), in the form of a project, with **progress tracked during the semester** (topic choice deadline, preliminary results deadline, final result deadline)
- ▶ Format:
 - Online (lectures & seminars)

The formula

Final grade = $0.5 \cdot \text{Homework score} + 0.5 \cdot \text{Exam score}$

$0 \leq \text{Homework score} \leq 10$

$0 \leq \text{Exam score} \leq 10$

Rounding to the closest integer

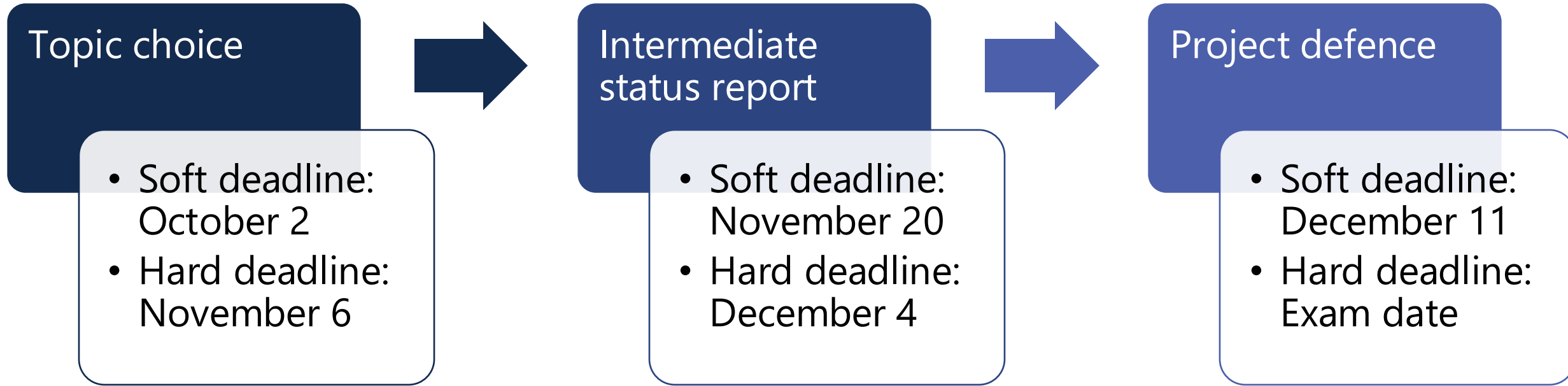
Homework

- ▶ A small set of tasks each week (jupyter notebooks)
- ▶ Solve tasks to earn points
- ▶ Deadline: 2-3 weeks per task
- ▶ Homework grade = $10 \cdot \min\left(1, \frac{\sum \text{points}}{\text{total}}\right)$

Exam

- ▶ Exam in the form of project defense
- ▶ The project is either:
 - Participation in a machine learning competition (on www.kaggle.com or similar)
 - Teams of up to 3 people are OK (roles of all members of a team should be clear and significant)
- ▶ Or:
 - Implementation of some technique or study from an advanced machine learning paper
 - I'll provide some suggestions later,
 - or you can find something that interests you by yourself
- ▶ **Please discuss your choice with me**

Exam project timeline



- ▶ Missing any of the hard deadlines adds a -0.5 points penalty to the exam grade (for each of the missed deadlines)
- ▶ Meeting any of the soft deadlines adds a $+0.5$ points bonus to the exam grade (for each of the met deadlines)

Deep learning course

- ▶ There's an elective course just on the deep learning:
 - <https://www.hse.ru/en/edu/courses/894121223>
 - Highly recommended!
 - Homework scores for overlapping topics can be transferred to our course

Transferring grades from your previous ML courses

- ▶ Those of you for whom our course is compulsory (i.e., ISSA spec. year 2) may be able to transfer grades from previously taken machine learning courses if their program overlaps with ours
- ▶ Contact me in telegram for more details

Thank you!



msohrabi@hse.ru



@MSohrabi_CS



@MSOHRABI_CS