Exam

Organization and Q&A

Exam is oral, i.e. we would like to discuss with you course material in the form of discussion. In order to organize it and give points, the exam is structured in the two parts: Theoretical Minimum Questions ($\{0, 4\}$ points) and Problems ($\{0, 1, \ldots, 6\}$ points).

Theoretical Minimum Questions We select any two questions from the list Sec (1). These questions are light and short. The aim is to check definitions and course material in general. The student is supposed to answer without preparation. We have the following possibilities:

1. Both question are answered successfully. Then the student gets +4 points, the Broad question and 20 minutes for preparation.
2. Both question are answered un-successfully. Then the students gets 0 points and the long summer for preparation.
3. One question is answered successfully. Then the students gets one more question from the list and answers it ...:
   (a) ... successfully. Then the student gets +4 points, the Broad question and 20 minutes for preparation.
   (b) ... un-successfully. Then the students gets 0 points and the long summer for preparation.

Problems These questions check more in-depth knowledge of the topics discussed on the lectures. See examples at 2 and list of the examinable Weeks of the course.

Q&A

Where will it be? When?
Answer: We provide two possibilities: you can either use Zoom or come to the HSE CS building. The format and rules are the same, but we suppose it is easier to communicate in person. The web-camera, microphone and stable Internet connection is your responsibility. The web-camera should be turned on during the whole exam. Date is 24 June.

Can I use Russian?
Answer: Yes, you can choose between Russian and English languages, we will use the same as you.

Can I get 0 points?
Answer: Yes, you can.

Can I get 1,2,3 points?
Answer: No, you can not. For Theoretical Minimum part you can get 0 or 4 points. Hence, you can get 0 or ≥ 4.

What Can I use during exam?
Answer: During the answer you can use only your hand-written notes. You should be alone during the exam.

Can I firstly come to exam and then decide take auto-grade?
Answer: No, you should decide before 19 June: take auto-grade or come to the exam.

1 Theoretical Minimum Questions

1. Name types of ML problems (at least 2 supervised and 2 unsupervised), give examples of such problems. For each problem discuss required data and cost function.
2. What is gradient descent? Write down the algorithm.
3. Is it true that gradient descent converges for any step size? Justify your answer.
4. Which problems can we face, when using closed form solution for linear regression with MSE loss? How to overcome them?
5. What is overfiting? How to detect it? What is underfitting? How to detect it?
6. What is K-fold cross-validation and Leave-one-out-cross validation. How the latter can be effectively used with linear regression?
7. In which situation, accuracy is a bad metric for classification task? What are the alternatives?
8. Explain what is ROC-curve and PR-curve, how are they ussed to assess classification models?
9. Define Logisitc Regression (Training data, model, objective). How is it connected to the minimization of error rate?
10. Define Support Vector Machine (Training data, model, objective). How is it connected to the minimization of error rate?
11. What is a support vector in SVM model?
12. What is an impurity criterion? How is it used for trees construction? Give examples of such criterion for regression and classification.
13. Consider multi-class classification problem with K classes. You've trained a decision tree to predict a class. How to get vector of class probabilities for a new given object using this tree?
14. Consider a trained decision tree. Are we guaranteed that this tree is optimal (for a given dataset)? Why?
15. Write down Bias-variance decomposition (without derivation), explain what each component means
16. What is Bagging? What is the influence of baggining on bias and variance?
17. What is random forest? Are there any differences between bagging and random forest?
18. How deep trees for gradient boosting should be? Justify your answer using bias-variance decomposition.
19. Let us train each base algorithm of gradient boosting on all the features and on the full training dataset. Does it make sense to use such algorithm for regression task? For classification task?
20. Name three types of anomalies, give examples.
21. When k-mean algorithm is used? Explain how it works (training procedure).
22. Give examples of 3 anomaly detection methods. Explain how they work.
23. Define component of the bayesian inference: what is prior, posterior and predictive distribution?
24. Explain the difference between bayesian inference and maximum likelihood
25. Is it always easy to find posterior distribution? Why? What are the options to deal with it?
26. Define artificial neuron. When is it identical to logistic regression?
27. Give examples of activation functions.

2 Problems

Examinable Weeks: $2, 3, 4, 5, 6, 7, 8, 9, 11, 13$.
Topics from the lectures $1, 10, 12, 14, 15$ will not be covered.

Example of Questions

- Derive Bias-Variance decomposition for regression with MSE loss
- Consider Linear Regression problem with MSE loss. For which design matrix X the solution of the corresponding optimization problem is unique? What if this condition is violated, how to regularize the optimization problem?
- Consider Gradient Boosting model. Let us use as residuals difference between predicted and real values: $s_n = y_n - b(x_n)$. Will such method work for classification?