Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Суворов Антон Дмитриевич
Должность: Ректор
Дата подписания: 31.08.2026 16:35:32
Уникальный программный ключ:
a39bdb15d680d3b0adbfced0af5c1efb14747dc0

Course Syllabus

Skoltech

Contact Person	Ivan Tyukin
Contact Person's E-mail	I.Tyukin@skoltech.ru
Course Title (in English)	Statistical Learning Theory
Course Title (in Russian)	Статистическая теория обучения
Lead Instructor	Tyukin, Ivan

1. Annotation

Course Description

The course is intended as an introduction into the field of statistical learning with the focus on the understanding of (i) when and why can we learn from empirical data and (ii) when such learning may not be possible or practically feasible. One of the central questions this course will address is: given a model (a neural network, linear classifier, k-NN, etc.) and a training algorithm, how large both the training and testing datasets must be so that upon completion of training we can guarantee an appropriately chosen level of the model's performance, with high probability. In order to be able to answer this important practical question we will review and learn main ingredients of modern statistical learning theory. We will also delve into the topic of complexity (VC dimension, Rademacher complexity, Covering numbers), an important notion for the understanding of learning from empirical data. Upon successful completion of the course the students will be able to make informed decisions about the choice of machine models and learning algorithms given data. Theoretical results will be illustrated with examples, and the students will be offered a number of hands-on exercises and assignments.

Аннотация

Курс представляет собой введение в область статистического обучения, направленное на понимание того, (1) когда и почему возможно обучение на эмпирических данных и (2) когда такое обучение может быть невозможным или практически неосуществимым. Центральный вопрос курса: для заданной модели (нейронная сеть, линейный классификатор, k-NN и т.д.) и алгоритма обучения, какого объема должны быть обучающая и тестовая выборки, чтобы по завершении обучения можно было гарантировать с высокой вероятностью достижение заданного уровня производительности модели. Для ответа на этот важный практический вопрос будут рассмотрены основные компоненты современной статистической теории обучения. Мы также углубимся в тему сложности (размерность Вапника-Червоненкиса, сложность Радемахера, покрывающие числа) — важнейшее понятие для понимания обучения на данных. Успешно завершив курс, студенты смогут принимать обоснованные решения о выборе моделей и алгоритмов обучения в зависимости от данных. Теоретические результаты будут проиллюстрированы примерами, а студентам будет предложен ряд практических упражнений и заданий.

2. Basic Information

Course Academic Level	MSc, PhD

Course Academic Level Master-level course suitable for PhD students

Number of ECTS credits 3

Course Prerequisites / Recommendations

We will assume a significant level of mathematical maturity. This means an understanding of linear algebra, analysis, and probability. Convex optimization and machine learning will be extremely helpful but is not strictly necessary. Programming skills are required (C/C++/Python or Julia).

Type of Assessment Graded

Grading Scale

A:	75
B:	65
C:	55
D:	45
E:	25

Course Term (in context of Academic Year) Term 3

Students of Which Programs do You Recommend to Consider this Course as an Elective?

BSc Programs	Masters Programs	PhD Programs
	AI and Financial Technologies AI in Robotics Data Science	Computational and Data Science and Engineering

Maximum Number of Students

	Maximum Number of Students
Overall:	15
Per Group (for seminars and labs):	15

3. Course Content

ECTS Credit System – Reference Tool

Total Number of ast. Hours in Your Course 81

Topic	Summary of Topic	Contact Hours: Lectures	Contact Hours: Seminars	Contact Hours: Labs	Non-contact Hours: Student's Independent Study
Introduction into the problem of learning.	Optimality of Bayesian decision-making. Glivenko-Cantelli theorem, Dvoretzky–Kiefer–Wolfowitz–Massart inequality. Examples of application to improving classifiers.	1.5	3	0	2
The problem of learning from empirical data.	Risk and Empirical risk functionals. Generalization error and overfitting. The concept of PAC learning. Sample complexity. Concentration inequalities (Markov, Chebyshev, Hoeffding).	1.5	3	0	2
Bounds for binary classification & finite/countable classes.	PAC learnability and no free lunch theorems. Bias-variance trade-off.	1.5	3	0	2
Bounds for uncountably large hypothesis classes.	Growth functions. Vapnik-Chervonenkis (VC) dimension. Sauer-Shelah lemma. PAC learnability with finite VC dimension.	1.5	3	0	2
Computing VC dimension for basic models.	kNN classifiers, support vector machines, neural networks.	1.5	3	0	2
Complexity measures beyond VC dimension.	Covering numbers, Rademacher complexity, fat-shattering dimension.	1.5	3	0	2
The boosting problem.	Limitations of boosting, the AdaBoost algorithm, and why boosting works.	1.5	3,5	0	2
Stability and generalization.	The challenge of verifiable accuracy and robustness. Stability certificates through randomized smoothing.	1.5	3,5	0	2
Elements of statistical learning in high-dimension.	Sub-gaussian random variables, waist concentration, Johnson-Lindenstrauss lemma. Learning from few examples.	1.5	3,5	0	2

Topic	Summary of Topic	Contact Hours: Lectures	Contact Hours: Seminars	Contact Hours: Labs	Non-contact Hours: Student's Independent Study
Independent Work & Examination.	Preparation and writing of the reflective essay. Final preparation for the oral exam.	0	0	0	15
Essay Defense.	Individual meeting for discussing and defending the reflective essay.	0	3	0	0
Oral Examination.	Individual oral assessment covering all course topics.	0	3	0	0

4. Learning Outcomes

Please specify course intended learning outcomes using
S koltech Learning Outcomes Framework
.

1. FUNDAMENTAL KNOWLEDGE

1.2. KNOWLEDGE OF APPLIED SCIENCE AND ENGINEERING SCIENCE INCLUDING CONTEMPORARY METHODS AND TOOLS

2. PERSONAL AND GENERAL PROFESSIONAL SKILLS AND ATTRIBUTES:

2.1. COGNITION AND MODES OF REASONING

2.1.1. Analytical reasoning and problem solving

2.1.5. Critical thinking and meta-cognition

2.2. ATTITUDES AND LEARNING PROCESS

2.2.5. Self-awareness and a commitment to self-improvement, lifelong learning and educating

3. INTERPERSONAL SKILLS

3.1. COMMUNICATIONS IN INTERNATIONAL ENVIRONMENTS

3.1.5. Communications in English in scientific, business and social settings

4. LEADING THE INNOVATION PROCESS

4.2. VISIONING — INVENTING NEW TECHNOLOGIES THROUGH RESEARCH

4.2.1. The research process — hypothesis, evidence and defense

5. Assignments and Grading

In-person Attendance Requirement 80

In-person Attendance Requirement Comment

Regular attendance and active participation in lectures and seminars are mandatory and form a significant part of the final grade. Absences are only excused for valid reasons (e.g., medical, with documentation).

Assignment Type	Assignment Summary	% of Final Course Grade
Class participation	Regular attendance and active, high-quality engagement in lectures and seminar discussions. Evaluated on both presence and contribution.	30
Essay	A critical written analysis (10-15 pages) of open problems in Statistical Learning Theory, demonstrating understanding, critique, and research thinking.	20
Other	A comprehensive individual oral exam covering all course topics, assessing depth of knowledge, reasoning, and application skills.	50

6. Assessment Criteria

Select Assignment 1 Type

Class participation

Input Sample of Assignment 1 or Share a Link to Assignment 1

Regular, active, and thoughtful participation in all lectures and seminar discussions throughout the course. Attendance is recorded and forms the basis of this assessment component.

Assessment Criteria for Assignment 1

Weight: 30% of the final grade. The grade is based on two components:

1. Attendance Record (15 points): Students are required to attend at least 80% of all scheduled classes. Each unexcused absence beyond the allowed 20% threshold will result in a proportional deduction of points.

2. Quality of Engagement (15 points): Evaluated based on active contribution to seminar discussions, asking relevant and insightful questions, and constructive participation in problem-solving activities during classes. Passive attendance is insufficient for a high score.

Select Assignment 2 Type

Essay

Input Sample of Assignment 2 or Share a Link to Assignment 2

A written reflective essay (approximately 10-15 pages) on open problems and current research frontiers in Statistical Learning Theory. The essay must critically analyze at least two major open questions discussed in the course, propose well-reasoned potential research directions, and reflect on the limitations of existing theoretical frameworks.

The use of AI-assisted tools (e.g., for help with phrasing or brainstorming) is permitted. In passages where text or ideas are directly borrowed from AI-generated content, please provide a footnote or reference indicating the specific tool used. This is for academic transparency. The core arguments, critical analysis, and conclusions must reflect your own understanding.

The written essay is followed by a mandatory oral defense. Each student will schedule an individual meeting with the instructor to discuss their essay. The defense is a discussion aimed at exploring the essay's ideas in depth and confirming the student's mastery of the topic.

Assessment Criteria for Assignment 2

Weight: 20% of the final grade. The assignment is graded out of 100 points according to the following criteria, then scaled to 20%:

1. Depth of Understanding & Accuracy (20 pts): Correct and insightful explanation of the chosen problems and relevant concepts.

2. Critical Analysis & Originality of Thought (20 pts): Quality of critical evaluation and creativity of

proposed directions.

3. Clarity, Structure & Presentation (15 pts): Logical flow, professional presentation, clarity of writing. Adherence to the policy on referencing AI use where applicable.

4. Literature Review & Referencing (15 pts): Appropriate use and citation of core and external literature.

5. Oral Defense & Mastery of the Topic (30 pts): This is the most heavily weighted criterion. Evaluation based on the student's performance during the individual oral defense. The student must demonstrate a confident command of the essay's content, engage in a substantive discussion of the ideas, and respond clearly to the instructor's questions.

Select Assignment 3 Type

Other

Input Sample of Assignment 3 or Share a Link to Assignment 3

A comprehensive individual oral examination conducted at the end of the course, covering all major topics from the syllabus. The student will be asked to explain key concepts, outline proofs of fundamental theorems, and apply theoretical results to analyze simple learning scenarios or models.

Assessment Criteria for Assignment 3

Weight: 50% of the final grade. The oral exam is graded out of 100 points according to the following criteria, then scaled to 50%:

1. Accuracy & Depth of Theoretical Knowledge (40 pts): Correctness and completeness in explaining definitions, theorems, and core results. Demonstration of a deep, interconnected understanding of the subject matter.

2. Articulation & Conceptual Reasoning (30 pts): Ability to present ideas clearly and logically under examination conditions, to connect different concepts, and to think critically in response to follow-up questions.

3. Problem-Solving & Application (30 pts): Skill in applying abstract theoretical guarantees to concrete examples, analyzing given scenarios, and performing correct derivations or calculations when prompted.

The examination lasts approximately 20-30 minutes per student.

7. Textbooks and Internet Resources

You can request at most two required textbooks. Additionaly, you can suggest up to nine recommended textbooks.

Required Textbooks	ISBN-13 (or ISBN-10)
Anthony, M., and Bartlett, P. L. Neural network learning: Theoretical foundations. Cambridge university press, 2009.	978-0-521-57353-5

Papers	DOI or URL
Bousquet, O., Boucheron, S., Lugosi, G. (2004). "Introduction to Statistical Learning Theory". In: Bousquet, O., von Luxburg, U., Rätsch, G. (eds) Advanced Lectures on Machine Learning. ML 2003. Lecture Notes in Computer Science, vol 3176. Springer, Berlin, Heidelberg.	https://doi.org/10.1007/978-3-540-28650-9_8

8. Facilities

9. Additional Notes

Is this syllabus complete?	Yes, the syllabus is a final draft waiting for approval by Education Department
The proposed course 1) has explicit academic content and requirements for receiving credits, 2) is in alignment with the program's learning outcomes, 3) adheres to policies and Skoltech regulations.	Lead Instructor confirms
Syllabus status	Approved by the Education Department