FEO3350 Information Theory for Statistics and Learning 12.0 credits

Informationsteori för statistik och lärande

This is a translation of the Swedish, legally binding, course syllabus.

Establishment

Course syllabus for FEO3350 valid from Autumn 2020

Grading scale

P, F

Education cycle

Third cycle

Specific prerequisites

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

A student who has passed this course should be able to:

- understand what concepts in information theory are most important in statistical decision theory, signal processing and learning theory
- understand and recapitulate several of the most important proofs that build the theory
- apply and synthesize the proof techniques and approaches in fundamental exercises
- acquire, present, and discuss the knowledge of advanced research papers in the field

Course contents

Lecture 1: Information theory fundamentals: Entropy, mutual information, relative entropy, and f-divergence. Total variation and other distance metrics. Inequalities.

Lecture 2: Rate-Distortion theory: Cost versus information. Bounds. The Blahut algorithm.

Lecture 3: Limits on information flow and processing: Conditional mutual information and relative entropy. Data processing inequalities. Sufficient statistic and the information bottleneck. Rate-distortion interpretation

Lecture 4: Foundations of statistical decision theory: Parameter estimation. Bayes and minimax risk. Binary hypothesis testing

Lecture 5: Information bounds on error probability and risk: Sample complexity. The mutual information method and rate-distortion. Fano inequalities.

Lecture 6: Learning and generalization: Information bounds on generalization error. VC dimension and complexity.

Lecture 7: Classical estimation theory: Maximum likelihood, Fischer information, information bounds, Cramér-Rao, Hammersley-Chapman-Robbins.

Lecture 8: Packing, covering, Fano & minimax risk, metric entropy

Lecture 9: Le Cam's method, mutual information method continued. Density estimation. Functional estimation.

Lecture 10: Dimension reduction and denoising: Sparsity, compressed sensing, denoising, almost noiseless analog compression

Lecture 11: Variational methods: Variational inference and the ELBO, variational characterization of divergence and information. Donsker-Varadhan. Gelfand-Yaglom-Perez.

Lecture 12: The method of types

Lecture 13: Information theory and large deviations, Stein, Chernoff and Sanov. Total variation and hypothesis testing.

Lecture 14: The geometry of information: Information geometry, information projection, iterative methods, Expectation-Maximization

Examination

- EXA1 - Written examination, 12.0 credits, grading scale: P, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

If the course is discontinued, students may request to be examined during the following two academic years.

Ticking in problem solving seminars: Students have to tick before every problem solving seminar the problems that they have solved. The presented solution will be discussed in the seminars with the teachers and other students.

Oral examination of having aquired sufficient knowledge of basic concepts.

Other requirements for final grade

The student has to have ticked at least 70% of the problems and has successful presented a solution when called. Sufficient knowledge has been demonstrated in the oral examination.

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.