STOCKHOLM UNIVERSITY
Department of Statistics
Spring term 2021

Mattias Villani and Frank Miller
2021-04-10

Course description for Machine Learning, 7.5 HEC, ST5401

COURSE CONTENTS

The course covers a number of machine learning methods with a focus on prediction. The course deals with supervised and unsupervised machine learning as well as semi-supervised and active learning. Contents include flexible regression and classification, regularization, methods for predictive model performance evaluation, deep learning, clustering algorithms and mixture models, and methods for semi-supervised and active learning.

The course consists of two modules:

1. Machine learning exam, 4.5 ECTS credits
2. Machine learning assignments in the R programming language, 3 ECTS credits

LEARNING GOALS

After completing the course, the student should be able to:

- Formulate and structure solutions to practical machine learning problems
- Identify and estimate suitable machine learning models for prediction and clustering
- Evaluate and select among machine learning models and learning algorithms
- Implement machine learning models and algorithms in a programming language

COURSE LITERATURE

- Lindholm, A., Wahlström, L., Lindsten, F. och Schön, T. B. (2021). *Machine Learning - A First Course for Engineers and Scientists*. Cambridge University Press. Freely available in PDF at: http://smlbook.org/.
- Additional material distributed during the course, e.g. lecture notes, exercises, etc., will be posted on the course web site at `https://github.com/mattiasvillani/MLcourse` or on the learning platform Athena.

EXAMINERS AND TEACHERS

Lecturer and Examiner: Mattias Villani
E-mail: mattias.villani@stat.su.se
Reception hours: on agreement.

Lecturer and Examiner: Frank Miller
E-mail: frank.miller@stat.su.se
Reception hours: on agreement.

Computer lab assistant: Karl Sigfrid
E-mail: karl.sigfrid@stat.su.se
Reception hours: on agreement.

The Department of Statistics is located on the 7th floor in the B-building (Södra Husen). General information about the department (office hours, phone numbers, schedules etc.) is posted on statistics.su.se. Information about the course is posted on the course site at Athena.

COURSE EVALUATION

A course evaluation is made after the course is completed. The course evaluation is used as a basis for course quality work and as part of the student influence procedure. The evaluation is carried out by a questionnaire sent out via e-mail to all registered course participants. The course participants' responses to the questionnaire are compiled and posted at Athena together with the course examiner´s final course evaluation.

TEACHING FORMAT

Teaching consists of 12 lectures (L1-L12) and 3 computer sessions (C1-C3) according to schedule, see link to TimeEdit on the course home page. A description of the course contents and reading instructions are given on the course web page at `https://github.com/mattiasvillani/MLcourse`.

MANDATORY ATTENDANCE

There is no mandatory attendance at the course, however it is strongly recommended to participate in all lectures and computer sessions.

EXAMINATION AND GRADES

Students are assessed by examination of the expected learning outcomes through two written examinations:

- Module 1 - Individual exam, is graded according to a criterion-referenced seven-grade scale.
- Module 2 – Hand-in assignment, is graded according to a two-grade scale with the grades Pass (G) and Fail (U)

Grading of Module 1 is done according to a seven-point scale related to the specified learning outcomes:

Grade	Description
A	Excellent
B	Very Good
C	Good
D	Satisfactory
E	Adequate
Fx	Inadequate
F	Totally inadequate

In order to pass the course, the grade E or higher is required on Module 1 and Pass on course Module 2. Grading for the whole course is then equivalent with the grading of Module 1.

- When obtaining a failing grade F or Fx in Module 1, no extra exercises or extra assignments are given to obtain a passing grade.
- Students who have received the grade Fx or F on Module 1 are entitled to at least four additional examinations to achieve at least grade E as long as the course is given.
- Students who have received the grade E on Module 1 may not retake this examination in order to attempt to achieve a higher grade.
- Students who have received the grade Fx or F on Module 1 on two occasions by the same examiner have the right to request that a different examiner be appointed to set the grade of the next examination. The request must be in writing and sent to the head of the department.
- Every time the course is given, there should be two examination opportunities for each Module during the current semester.
- When the course syllabus has been withdrawn, the student has the right to request examination once per semester during a period of three semesters in accordance with this syllabus. The request must be in writing and sent to the head of department.

MODULE 1 - Machine learning exam, 4.5 ECTS credits

During the spring semester 2021, the exam and re-exam will be given as a home exam. The writing time will be extended one hour i.e. a total of 6 hours, including time for electronical submission.

- Module 1 is examined through a written computer-based exam.
- The writing time is 5 hours. The written exam is an individual exam and during the exam, collaboration is prohibited, as are other aids than those allowed by the examiner.
- Special support, if required, may be allowed after request at the department's student counsellor and upon permission of the examiner. Contact the student counsellor well in advance before the exam, preferably no later than three weeks.
- Applicable rules for examinations at Stockholm University can be found at:

https://www.su.se/medarbetare/organisation-styrning/styrdokument-regelboken/utbildning/regler-f%C3%B6r-salstentamen-1.513161

Note! You must register for the exam no later than 8 days before the exam date. You register for the exam via Ladok. If you have problems with the registration, contact expedition@stat.su.se. If you have not registered correctly you cannot take the exam.

Dates for Module 1

- Exam: Home exam, June 4, 9-15.
- Re-exam: Home exam, TBA.

MODULE 2 - Hand-in assignment, 3 ECTS credits

- Module 2 is examined through three hand-in assignments and should be completed in groups of 2 students. The three hand-in assignments must be presented in the form of three written reports. Instructions for the hand-in assignments can be found at Athena.
- Module 2 is graded as Pass (all three hand-in assignments are approved) or Fail (at least one hand-in assignments is not approved). If one or more hand-in assignments are graded as Fail, the students will have a chance to re-submit the incorrect hand-in assignment(s).
- Collaboration within the group is of course allowed, however individual testing and grading within the group may occur. Please note that all group members are responsible for and should be able to answer to all parts of the report. Cooperation between groups is allowed, but each group must submit its unique report. Any type of plagiarism is prohibited and text-matching tools may be used.
- **Note!** It is forbidden to post the solution of the assignment publically in any form, for example by uploading the solution in a public repository such as GitHub.
- **Note!** If you miss the first submission occasion and submit for the first time at the re-submission occasion and then fail, you have no more chances to resubmit during the current semester.
- **Note!** All hand-in assignments must be approved no later than after the grading of the final submission on August 20 (see below) for the entire Module 2 to be graded as "Pass". Results from hand-in assignments are not saved and cannot be transferred to future semesters.

Dates for Module 2

- See the deadline for the submission of each hand-in assignment in Athena.
- Re-submission for one or more incorrect hand-in assignments: August 20, no later than 18.00 via Athena.

GRADING CRITERIA

Module 1 - Machine learning exam, 4.5 ECTS credits

Module 1 is an individual written computer-based exam. The written exam covers the course content material. The following grading criteria apply for Module 1:

A: (Excellent): The student can in an excellent way formulate and solve machine learning problems that have not necessarily been directly addressed in the course. The student provides exhaustive and clear explanations of machine learning concepts, models and methods. Corresponds to at least 90% of the total written test score.
B: (Very good): The student can in a very good way formulate and solve all types of machine learning problems that have been considered in the course. The student provides clear explanations of machine learning concepts, models and methods. Corresponds to 80-89% of the total written test score.
C: (Good): The student can in a well-structured way formulate and solve most types of machine learning problems that have been considered in the course. The student provides correct explanations of machine learning concepts, models and methods. Corresponds to 70-79% of the total written test score.
D: (Satisfactory): The student can in a satisfactory way formulate and solve most types of machine learning problems that have been considered in the course. The student provides satisfactory explanations of machine learning concepts, models and methods. Corresponds to 60-69% of the total written test score.
E: (Adequate): The student can in an adequate way formulate and solve most types of machine learning problems that have been considered in the course. The student can in an adequate way explain machine learning concepts, models and methods. Corresponds to 50-59% of the total written test score.
Fx: (Inadequate): The student's achievements are inadequate with respect to at least one of the criteria for the grade E. Corresponds to 40-49% of the total written test score.
F: (Totally Inadequate): The student shows significant inadequacies in relation to the assessment criteria. Corresponds to 0-39% of the total written test score.

Module 2 comprises of three hand-in assignments carried out in a group. Module 2 is graded as Pass or Fail. The following grading criteria apply for Module 2:

Module 2 - Hand-in assignment, 3 ECTS credits

Pass: The student is be able to analyze machine learning problems in the R language in accordance with the assignment instructions. The student can summarize his/her work in well-written reports.
Fail: The student's achievements are inadequate with respect to at least one of the criteria for the grade Pass.