Security and Privacy in Machine Learning

Säkerhet och integritet för maskininlärning
6 credits

Programme course

TDDE91

Valid from: 2027 Spring semester

Determined by	**Main field of study**	
Board of Studies for Computer Science and Media Technology	Information Technology, Computer Science and Engineering, Computer Science	
Date determined	**Course level**	**Progressive specialisation**
	Second cycle	A1N
Revised by	**Disciplinary domain**	
	Information missing	
Revision date	**Subject group**	
	Computer Technology	
Offered first time	**Offered for the last time**	
Spring semester 2027		
Department	**Replaced by**	
Institutionen för datavetenskap		

Course offered for

- Master of Science in Computer Science and Engineering
- Master of Science in Information Technology
- Master of Science in Computer Science and Software Engineering
- Master's Programme in Computer Science
- Master of Science in Applied Physics and Electrical Engineering
- Master of Science in Engineering Mathematics
- Master's Programme in Cybersecurity

Prerequisites

Knowledge equivalent to an introductory course in machine learning or artificial intelligence (e.g., Introduction to Machine Learning, Machine Learning, Neural Networks and Learning Systems, Artificial Intelligence (AI)). Programming skills for data science or machine learning applications e.g. in Python.

Basic knowledge of information security is useful, but not required.

Intended learning outcomes

The course aims to equip students with an understanding of security and privacy issues in modern machine learning systems and to introduce key concepts related to the trustworthiness of such systems. The course combines theoretical knowledge with practical skills for developing robust, privacy-preserving, and trustworthy machine learning systems. After completion of the course, the student shall be able to:

- Explain the major security and privacy risks in modern machine learning systems.
- Apply threat modeling and security analysis methods to identify and assess risks in machine learning pipelines, including evolving attacker models such as AI-assisted adversaries.
- Design and implement security and privacy attacks against machine learning systems, develop the corresponding defenses, and evaluate their effectiveness through empirical experiments.
- Explain core principles for building trustworthy machine learning systems and connect these concepts to real-world scenarios.
- Collaborate and communicate the results of a group project, including implementation, documentation, and demonstration.

Course content

- Vulnerability analysis and threat modeling for machine learning systems deployed across various domains (e.g., healthcare, network management, and industrial environments) and different training approaches including supervised learning, reinforcement learning, distributed learning, and generative models, taking into account emerging attacker capabilities such as AI-assisted attacks.
- Security: Adversarial examples, data and model poisoning, backdoor attacks, and defenses against such attacks, including robust training and detection.
- Privacy: Model extraction, membership inference, leakage of sensitive training data, unauthorized use of models and datasets, secure model deployment, and privacypreserving technologies for machine learning.
- Concepts and terminology related to regulatory compliance, governance, reliability, fairness, and bias mitigation in modern AI applications.

Teaching and working methods

The course consists of lectures, laboratory work, and a group project. In the lectures, students are introduced to the fundamental concepts of security and privacy challenges in machine learning systems, as well as governance and threat modeling. Laboratory sessions include hands-on exercises in which the student applies strategies and methods to identify vulnerabilities, conduct attacks, and implement defenses in machine learning systems. Project work includes group-based implementation, testing, and evaluation of a security or privacy challenge, concluded with a project report and demonstration.

Students are expected to take an active role in the learning process and prepare for project meetings outside of scheduled contact hours.

Examination

LAB1	Computer laboratory work	3 credits	U, G
PRA1	Project	3 credits	U, 3, 4, 5

Grades for examination modules are decided in accordance with the assessment criteria presented at the start of the course.

Grades

Four-grade scale, LiU, U, 3, 4, 5

Other information

About teaching and examination language

The teaching language is presented in the Overview tab for each course. The examination language relates to the teaching language as follows:

- If teaching language is “Swedish”, the course as a whole could be given in Swedish, or partly in English. Examination language is Swedish, but parts of the examination can be in English.
- If teaching language is “English”, the course as a whole is taught in English. Examination language is English.
- If teaching language is “Swedish/English”, the course as a whole will be taught in English if students without prior knowledge of the Swedish language participate. Examination language is Swedish or English depending on teaching language.

Other

The course is conducted in such a way that there are equal opportunities with regard to sex, transgender identity or expression, ethnicity, religion or other belief, disability, sexual orientation and age.

The planning and implementation of a course should correspond to the course syllabus. The course evaluation should therefore be conducted with the course syllabus as a starting point.

The course is campus-based at the location specified for the course, unless otherwise stated under “Teaching and working methods”. Please note, in a campus-based course occasional remote sessions could be included.

Common rules

Course syllabus

A syllabus must be established for each course. The syllabus specifies the aim and contents of the course, and the prior knowledge that a student must have in order to be able to benefit from the course.

Timetabling

Program courses are timetabled after a decision has been made for this course concerning its assignment to a timetable module. Single subject courses can be timetabled at other times.

Interruption in and deregistration from a course

The LiU decision, Guidelines concerning confirmation of participation in education, Dnr LiU-2020-02256 (https://styrdokument.liu.se/Regelsamling/VisaBeslut/764582), states that interruptions in study are to be recorded in Ladok. Thus, all students who do not participate in a course for which they have registered are therefore obliged to report the interruption so that this can be noted in Ladok. Deregistration from or interrupting a course is carried out using a Web-based form.

Cancelled courses and changes to the course syllabus

Courses with few participants (fewer than 10) may be cancelled or organised in a manner that differs from that stated in the course syllabus. The Dean is to deliberate and decide whether a course is to be cancelled or changed from the course syllabus. For single subject courses, the cancellation must be done before students are admitted to the course (in accordance with LiUs regulation Dnr LiU-2022-01200, https://styrdokument.liu.se/Regelsamling/VisaBeslut/622645).

Guidelines relating to examinations and examiners

For details, see Guidelines for education and examination for first-cycle and second-cycle education at Linköping University, Dnr LiU-2023-00379 (http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592).

An examiner must be employed as a teacher at LiU according to the LiU Regulations for Appointments, Dnr LiU-2022-04445 (https://styrdokument.liu.se/Regelsamling/VisaBeslut/622784). For courses in second-cycle, the following teachers can be appointed as examiner: Professor (including Adjunct and Visiting Professor), Associate Professor (including Adjunct), Senior Lecturer (including Adjunct and Visiting Senior Lecturer), Research Fellow, or Postdoc. For courses in first-cycle, Assistant Lecturer (including Adjunct and Visiting Assistant Lecturer) can also be appointed as examiner in addition to those listed for second-cycle courses. In exceptional

cases, a Part-time Lecturer can also be appointed as an examiner at both first- and second cycle, see Delegation of authority for the Board of Faculty of Science and Engineering.

Forms of examination

Principles for examination

Written and oral examinations and digital and computer-based examinations are held at least three times a year: once immediately after the end of the course, once in August, and once (usually) in one of the re-examination periods. Examinations held at other times are to follow a decision of the faculty programme board.

Principles for examination scheduling for courses that follow the study periods:

- courses given in VT1 are examined for the first time in March, with re-examination in June and August
- courses given in VT2 are examined for the first time in May, with re-examination in August and January
- courses given in HT1 are examined for the first time in October, with re-examination in January and August
- courses given in HT2 are examined for the first time in January, with re-examination in March and in August.

The examination schedule is based on the structure of timetable modules, but there may be deviations from this, mainly in the case of courses that are studied and examined for several programmes and in lower grades (i.e. 1 and 2).

Examinations for courses that the faculty programme board has decided are to be held in alternate years are held three times during the school year in which the course is given according to the principles stated above.

Examinations for courses that are cancelled or rescheduled such that they are not given in one or several years are held three times during the year that immediately follows the course, with examination scheduling that corresponds to the scheduling that was in force before the course was cancelled or rescheduled.

When a course, or a written or oral examination (TEN, DIT, DAT, MUN), is given for the last time, the regular examination and two re-examinations will be offered. Thereafter, examinations are phased out by offering three examinations during the following academic year at the same times as the examinations in any substitute course. The exception is courses given in the period HT1, where the three examination occasions are January, March and August. If there is no substitute course, three examinations will be offered during re-examination periods during the following academic year. Other examination times are decided by the faculty programme board. In all cases above, the examination is also offered one more time during the academic year after the following, unless the faculty programme board decides otherwise. In total, 6 re-examinations are offered, of which 2 are regular re-examinations. In the examination registration system, the examinations given for the penultimate time and the last time are denoted.

If a course is given during several periods of the year (for programmes, or on different occasions for different programmes) the faculty programme board or boards determine together the scheduling and frequency of re-examination occasions.

For single subject courses, written and oral examinations can be held at other times.

Retakes of other forms of examination

Regulations concerning retakes of other forms of examination than written examinations and digital and computer-based examinations are given in the LiU guidelines for examinations and examiners, Dnr LiU-2023-00379 (http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592).

In principle, other examination forms should be handled in the same way as a written examination when they are given for the last time. However, the times for the examination may vary based on the nature of the element compared to the times for the written examinations.

Course closure

For Decision on Routines for Administration of the Discontinuation of Educational Programs, Freestanding Courses and Courses in Programs, see Dnr LiU-2021-04782 (https://styrdokument.liu.se/Regelsamling/VisaBeslut/1156410). After a decision on closure and after the end of the discontinuation period, the students are referred to a replacement course (or similar) according to information in the course syllabus or programme syllabus. If a student has passed some part/parts of a closed program course but not all, and there is an at least partially replacing course, an assessment of crediting can be made. For questions about the crediting of course components, contact the Study councellors.

Registration for examination

In order to take an written, digital or computer-based examination, registration in advance is mandatory, see decision in the university's rule book Dnr LiU-2020-04559 (https://styrdokument.liu.se/Regelsamling/VisaBeslut/622682). An unregistered student can thus not be offered a place. The registration is done by the student at the Student Portal or in the LiU-app during the registration period. The registration period opens 30 days before the date of the examination and closes 10 days before the date of the examination. Candidates are informed of the location of the examination by email, four days in advance.

Code of conduct for students during examinations

Details are given in a decision in the university's rule book, Dnr LiU-2020-04559 (http://styrdokument.liu.se/Regelsamling/VisaBeslut/622682).

Retakes for higher grade

Students at the Faculty of Science and Engineering at LiU have the right to retake written examinations and digital and computer-based examinations in an attempt

to achieve a higher grade. This is valid for all examination components with code “TEN”, “DIT” and "DAT". The same right may not be exercised for other examination components, unless otherwise specified in the course syllabus.

A retake is not possible on courses that are included in an issued degree diploma.

Grades

The grades that are preferably to be used are Fail (U), Pass (3), Pass not without distinction (4) and Pass with distinction (5).

- Grades U, 3, 4, 5 are to be awarded for courses that have written or digital examinations.
- Grades Fail (U) and Pass (G) may be awarded for courses with a large degree of practical components such as laboratory work, project work and group work.
- Grades Fail (U) and Pass (G) are to be used for degree projects and other independent work.

Examination components

The following examination components and associated module codes are used at the Faculty of Science and Engineering:

- Grades U, 3, 4, 5 are to be awarded for written examinations (TEN) and digital examinations (DIT).
- Examination components for which the grades Fail (U) and Pass (G) may be awarded are laboratory work (LAB), project work (PRA), preparatory written examination (KTR), digital preparatory written examination (DIK), oral examination (MUN), computer-based examination in a computer lab (DAT), digital preparatory written examination in a computer lab (DAK), home assignment (HEM), and assignment (UPG).
- Students receive grades either Fail (U) or Pass (G) for other examination components in which the examination criteria are satisfied principally through active attendance such as tutorial group (BAS) or examination item (MOM).
- Grades Fail (U) and Pass (G) are to be used for the examination components Opposition (OPPO) and Attendance at thesis presentation (AUSK) (i.e. part of the degree project).

In general, the following applies:

- Mandatory course components must be scored and given a module code.
- Examination components that are not scored, cannot be mandatory. Hence, it is voluntary to participate in these examinations, and the voluntariness must be clearly stated. Additionally, if there are any associated conditions to the examination component, these must be clearly stated as well.
- For courses with more than one examination component with grades U,3,4,5, it shall be clearly stated how the final grade is weighted.

For mandatory components, the following applies (in accordance with the LiU Guidelines for education and examination for first-cycle and second-cycle

education at Linköping University, Dnr LiU-2023-00379 http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592):

- If special circumstances prevail, and if it is possible with consideration of the nature of the compulsory component, the examiner may decide to replace the compulsory component with another equivalent component.

For possibilities to alternative forms of examinations, the following applies (in accordance with the LiU Guidelines for education and examination for first-cycle and second-cycle education at Linköping University, Dnr LiU-2023-00379 http://styrdokument.liu.se/Regelsamling/VisaBeslut/917592):

- If the LiU coordinator for students with disabilities has granted a student the right to an adapted examination for a written examination in an examination hall, the student has the right to it.
- If the coordinator has recommended for the student an adapted examination or alternative form of examination, the examiner may grant this if the examiner assesses that it is possible, based on consideration of the course objectives.
- An examiner may also decide that an adapted examination or alternative form of examination if the examiner assessed that special circumstances prevail, and the examiner assesses that it is possible while maintaing the objectives of the course.

Reporting of examination results

The examination results for a student are reported at the relevant department.

Plagiarism

For examinations that involve the writing of reports, in cases in which it can be assumed that the student has had access to other sources (such as during project work, writing essays, etc.), the material submitted must be prepared in accordance with principles for acceptable practice when referring to sources when the text, images, ideas, data, etc. of other people are used. This is done by using references or quotations for which the source is specified. It is also to be made clear whether the author has reused his or her own text, images, ideas, data, etc. from previous examinations, such as degree projects, project reports, etc. (this is sometimes known as “self-plagiarism”).

A failure to specify such sources may be regarded as attempted deception during examination.

Attempts to cheat

In the event of a suspected attempt by a student to cheat during an examination, or when study performance is to be assessed as specified in Chapter 10 of the Higher Education Ordinance, the examiner is to report this to the disciplinary board of the university. Possible consequences for the student are suspension from study and a formal warning. More information is available at Cheating, deception and plagiarism.

Linköping University has also produced a guide for teachers and students' use of generative AI in education (Dnr LiU-2023-02660). As a student, you are always expected to gain knowledge of what applies to each course (including the degree project). In general, clarity to where and how generative AI has been used is important.

Regulations (apply to LiU in its entirety)

The university is a government agency whose operations are regulated by legislation and ordinances, which include the Higher Education Act and the Higher Education Ordinance. In addition to legislation and ordinances, operations are subject to several policy documents. The Linköping University rule book collects currently valid decisions of a regulatory nature taken by the university board, the vice-chancellor and faculty/department boards.

LiU's rule book for education at first-cycle and second-cycle levels is available at https://styrdokument.liu.se/Regelsamling/Innehall.