Introduction to Machine Learning Safety

Sommer

(in German: Introduction to Machine Learning Safety)

Module-ID: FIN-INF-120532

Abbreviation	Credit Points	Semester	Term	Duration	Language	Level
MLS	6	Sommer	ab 1.	1 Semester	english	Master

Link:	LSF
Limited accessability:	
Responsibility:	Prof. Dr. Frank Ortmeier
Lecturer:	Prof. Dr. Frank Ortmeier, Konstantin Kirchheim
Classes:	Vorlesung *Introduction to Machine Learning Safety* Übung *Introduction to Machine Learning Safety*
Applicability in curriculum:	- M.Sc. INF: Informatik - M.Sc. INGINF: Informatik - M.Sc. WIF: Informatik - M.Sc. DKE: Learning Methods and Models for Data Science - M.Sc. DE: Methoden der Informatik - M.Sc. VC: Computer Science

Intended learning outcomes:

- Verständnis grundlegender Unterschiede zwischen klassischer Softwareentwicklung und datengetriebener Entwicklung von ML-Systemen sowie der daraus resultierenden Sicherheitsimplikationen.

- Fähigkeit, typische Fehlermodi von ML-Systemen systematisch zu analysieren und deren Ursachen (z. B. Datenverteilung, Modellunsicherheit, Verteilungsverschiebung) zu identifizieren.
- Kompetenz, geeignete Test- und Evaluationsverfahren für ML-Modelle in sicherheitskritischen Kontexten auszuwählen und anzuwenden.
- Einordnung und Bewertung von Unsicherheitsabschätzungen, Anomalieerkennung und Robustheitsmechanismen im Hinblick auf ihre sicherheitstechnische Relevanz.
- Verständnis adversarieller Bedrohungsmodelle sowie grundlegender Verteidigungsstrategien.
- Reflexion regulatorischer, ethischer und gesellschaftlicher Rahmenbedingungen beim Einsatz lernbasierter Systeme.

Content:

- Grundlagen und Sicherheitskontext
- Datenqualität und Modelltestverfahren
- Robustheit und Angriffsmodelle
- Unsicherheit und Anomalieerkennung
- Alignment und gesellschaftliche Aspekte

Workload:

- 20h Übungsteilnahme,
- 20h Vorlesung,
- 40h Vorlesungsvorbereitung,
- 40h Übung,
- 60h Prüfungsvorbereitung

Pre-examination requirements:	Type of examination:	Teaching method / lecture hours per week (SWS):
Teilnahme an der Übung	Hausarbeit	Übung (2 SWS) Vorlesung (2 SWS)

Prerequisites according to examination regulations:	Recommended prerequisites:
keine	Introduction to Deep Learning

Media:	Literature:

Comments:

Revision #1
Created 2026-10-20 17:07:47 UTC by Fakultät für Informatik

Updated 2026-10-20 17:07:47 UTC by Fakultät für Informatik