



EP234U Grunderna i tillämpad maskininlärning 5,0 hp

Fundamentals of Applied Machine Learning

Fastställande

Kursplan för EP234U gäller från och med VT22

Betygsskala

P, F

Utbildningsnivå

Avancerad nivå

Huvudområden

Datalogi och datateknik

Särskild behörighet

- Kunskaper i envariabelanalys motsvarande IX1304 Matematik 7,5 hp
- Kunskaper i linjär algebra motsvarande SF1672 Linjär algebra 7,5 hp
- Kunskaper i sannolikhetssteori motsvarande SF2940 Sannolikhetssteori 7,5 hp
- Kunskaper i programmering motsvarande DD1315 Programmering och Matlab 7,5 hp
- Gymnasiekursen Engelska B/6

Undervisningsspråk

Undervisningsspråk anges i kurstillfällesinformationen i kurs- och programkatalogen.

Lärandemål

Efter godkänd kurs ska studenten kunna

- sammanfatta maskininlärning i ett grafikrenderingssystem, motivera dess komponenter och diskutera de problem som kan uppstå
- tillämpa olika befintliga övervakade och oövervakade maskininlärningsmetoder för givna datamängder och bedöma och granska deras resultat
- förklara olika maskininlärningsmetoder och kontrastera deras positiva och negativa egenskaper
- tolka befintliga realiseringar av olika maskininlärningsmetoder och anpassa dem för specifika situationer
- diskutera etiska dimensioner av maskininlärningsforskning, utveckling och tillämpning.

Kursinnehåll

Introduktion och motivation

Översikt över motiverande tillämpningar, bra och dåliga

Kursplanering och hemuppgiftsstruktur

Genomgång av lärande och modellering: Maskininlärning i ett grafikrenderingssystem

Exempel: Near Neighbor-klassificering

Parametrar och hyperparametrar

Träning, validering och testning

Partitionering av data: Exklusion, bootstrap, K-fold CV, LOOCV, etc.

Resultatmetrik: Förvirringssystem, korrekthet, precision och täckning

Övervakad inlärning 1

Probabilistisk klassificering och regression

Inkludera begrepp om risker vid klassificering och regression

Bayesiansk klassificering: Linjär diskriminantanalys

Bayesiansk klassificering: Kvadratisk diskriminantanalys

Bayesiansk klassificering: Naiv Bayes

Övervakad inlärning 2

Parameterskattning

Minstakvadratregression

Normering: LASSO ridge regression,

Bayesiansk regression

Logistisk regression

Oövervakat lärande 1

Vad är oövervakat lärande?

Dimensionens förbannelse

Principalkomponentanalys

Multidimensionell skalning

Översikt över K-means

Oövervakat lärande 2

Hierarkisk klustring
Densitetbaserad klustring
Avvikelsedetektion, avvikare (isoleringsskog)
Gaussiska blandade modeller
Deterministisk eller sannolikhetsteoretisk klustring

Arbeta med tidsserier

Motiverande exempel
Transformation mellan tids- och frekvensplan
Autoregressiv modellering
Autoregressiv modellering av rörligt medelvärde

Datarepresentation och utformning av särdrag

Utveckling av särdrag
Urval av särdrag
Samordnad optimering av särdragsutveckling och -klassificering

Pipeline för maskininlärning

AutoML-verktyg
Fallgropar med standardmetoder
Utökning och andra datatrick
Ansvar för utvecklaren och användaren
Tolka modeller för att motivera beslut
Korrelation och orsakssamband: maskininlärning är inte magiskt

Fördjupning: Förstärkningsinlärning (RL)

Översikt av tillämpningar inom förstärkningsinlärning
Grunderna för förstärkningsinlärning
Q-inlärning

Examination

- LAB1 - Laborationer, 4,0 hp, betygsskala: P, F
- PRO1 - Projekt, 1,0 hp, betygsskala: P, F

Examinator beslutar, baserat på rekommendation från KTH:s handläggare av stöd till studenter med funktionsnedsättning, om eventuell anpassad examination för studenter med dokumenterad, varaktig funktionsnedsättning.

Examinator får medge annan examinationsform vid omexamination av enstaka studenter.

När kurs inte längre ges har student möjlighet att examineras under ytterligare två läsår.

Etiskt förhållningssätt

- Vid grupparbete har alla i gruppen ansvar för gruppens arbete.
- Vid examination ska varje student ärligt redovisa hjälp som erhållits och källor som använts.

- Vid muntlig examination ska varje student kunna redogöra för hela uppgiften och hela lösningen.