



Kursplan för:

Datateknik AV, Datamining, 6 hp

Computer Engineering MA, Data Mining, 6 credits

Allmänna data om kursen

Kurskod	DT047A
Ämne/huvudområde	Datateknik
Nivå	Avancerad
Inriktning (namn)	Datamining
Högskolepoäng	6.0
Fördjupning vs. Examen	A1N , Kursen ligger på avancerad nivå och har endast kurs(er) på grundnivå som förkunskapskrav.
Utbildningsområde	Teknik 100%
Ansvarig avdelning	Avdelningen för informationssystem och -teknologi
Inrättad	2018-10-01
Fastställd	2018-10-11
Senast reviderad	
Giltig fr.o.m	2019-01-01

Syfte

Studenten ska utveckla en grundläggande förståelse för aktuella maskininlärningsmetoder för informationsutvinning (datamining) ur stora mängder data. Studenten ska utveckla färdigheter i att hitta intressanta mönster och bygga prediktionsmodeller genom explorativ dataanalys med hjälp av dataanalysverktyg som R, Weka eller Orange och kunna förbereda data, tolka resultat och kritiskt utvärdera resultat. Studenten ska visa förmåga att tillämpa kunskaperna i ett verkligt projekt inom områden såsom big data-visualisering, affärsdata för omvärldsanalys, beslutsstödssystem, dataanalys av text, webbsidor, sensor-/positionsdata, kontextmedvetna applikationer, intelligenta agenter eller kognitiv radio.

Lärandemål

Studenten ska kunna:

- Diskutera vilka verkliga tillämpningar av datamining som är realistiska och etiska
- Utvinna information med hjälp av ett verktyg som till exempel skriptspråket R, Python-biblioteket Orange, Java-baserade verktyg Weka eller egna implementationer av algoritmer
- Förbereda data, tolka data och utvärdera resultat
- Identifiera variabler av betydelse i en multivariat datauppsättning
- Utvinna associationsregler och utvärdera deras reliabilitet
- Utveckla och validera prediktionsmodeller
- Följa en standardiserad arbetsprocess för tillförlitlig problemanalys, modellering och utvärdering
- Tillämpa datamining-tekniker på ett litet problem från verkliga livet

Innehåll

- Användningsområden för data mining
- Data och kunskapsrepresentation (relationer, attribut, glesa matriser, tabeller, beslutsträd, regler)
- Bayesiansk statistik
- Associativa och sekventiella mönster
- Grundläggande algoritmer
- Dataklustring
- Datakategorisering
- Datarengöring
- Datavisualisering
- Associationsregler
- Dataprediktering
- Laboration på verktyget R, Orange eller Weka
- Projekt

Behörighet

120 hp avslutade kurser, inkluderande följande:

Datateknik GR (AB), innefattande Databaser, modellering och implementering, 6 hp och Java, 6 hp. Matematik GR (A), 30 hp, innefattande Matematisk statistik, 6 hp.

Urvalsregler

Urval sker i enlighet med Högskoleförordningen och den lokala antagningsordningen.

Undervisning

Kursen kan komma att erbjudas som campuskurs eller som webbaserad distanskurs. Studieinsatsen uppskattas till cirka 160 timmar.

Examination

0.0 hp, I101: Projektval

Betyg: Godkänd eller underkänd

0.5 hp, L101: Laboration

Betyg: Godkänd eller underkänd

3.5 hp, T101: Tentamen

Betyg: A, B, C, D, E, Fx och F. A-E är Godkänt, F är Underkänt.

2.0 hp, P101: Projektpresentation

Betyg: Godkänd eller underkänd

Slutbetyget baseras på en sammanvägning av tentamens- och projektresultat.

Ämnets betygskriterier hittas på www.miun.se/betygskriterier.

Om en student har ett beslut från samordnaren vid Mittuniversitetet om pedagogiskt stöd vid funktionsnedsättning, har examinator rätt att ge anpassad examination för studenten.

Begränsning av examination

Studenter registrerade på denna version av kursplan har rätt att examineras 3 gånger inom loppet av 1 år enligt angivna examinationsformer. Därefter gäller examinationsform enligt senast gällande version av kursplan.

Betygsskala

På kursen ges något av betygen A, B, C, D, E, Fx och F. A - E är Godkänt, Fx och F är underkänt.

Litteratur

Obligatorisk litteratur

Författare/red: Witten, Frank, Hall

Titel: Datamining - Practical Machine Learning Tools and Techniques

Upplaga: Third edition 2011 or later

Förlag: Elsevier

Referenslitteratur

Författare/red: Ganguly et al

Titel: Knowledge discovery from sensor data

Upplaga: 2009 or later