

Kursplan – Data scientist, Utbildnings nr YH01458-2024 1/ YH01076-2024-1		
Kurs: Machine Learning	Poäng: 30 Yhp	Utgåva: Utgåva 1
Framtagen av UL granskad av RUC: Monika Sandhaven/Molly Tagesson/ Stefan Olsson	Språk: Svenska / Engelska	Datum: 2024-05-23
Förkunskaper: Inga förkunskaper	Granskad/Fastställd av: Ledningsgruppen	Sidan 1 (2)

Kursens huvudsakliga innehåll

Kursen syftar till att ge den studerande specialiserade kunskaper i att förstå och använda maskininlärningstekniker för att kunna fatta datadrivna beslut. Under kursen kommer den studerande att lära sig att göra modellval, modellanpassning samt modellutvärdering i övervakade och oövervakade inlärningsproblem. Utöver detta kan den studerande nyttja AI (Artificiell Intelligens) såsom generativ AI i arbetsprocessen.

Kursen omfattar följande moment:

- Statistik och maskininlärning
- Modellval, modellanpassning och modellutvärdering
- Övervakade och oövervakade inlärningsproblem
- Regression
- Klassificering
- Beslutsträd
- Ensemble learning
- Klustering
- Förståelse av datans betydelse

Kursens mål/läranderesultat

Kunskaper:

1. Redogöra för övervakad inlärningsmodellering såsom regression-och klassificeringsmodeller
2. Redogöra för oövervakad inlärnings modellering
3. Förklara hur en kan utvärdera modeller

Färdigheter:

4. Implementera metodik och modeller från maskininlärning med syfte att prediktera och förstå data
5. Diskutera motivera och argumentera för olika modelleringsval utifrån data

Former för undervisning

Kursen kommer att genomföras med blended learning med inspelningar, aktiva lektioner samt ett grupparbete.

Former för kunskapskontroll

Examination kommer att ske genom:

- 1 individuell inlämningsuppgift (IG/G/VG)
- 1 gruppuppgift vilken ska redovisas muntligt (IG/G)

Betygsskala

Följande betygsskala tillämpas:

VG = Väl Godkänd, G = Godkänd, IG = Icke Godkänd

Läranderesultat	Inlämningsuppgift (VG)	Gruppuppgift (G)
1	X	
2	X	
3	X	
4	X	X
5	X	X

Principer för betygssättning

För betyget Godkänd ska den studerande

- Ha nått samtliga läranderesultat för kursen

För betyget Väl Godkänd ska den studerande:

- Uppnått kraven för betyget Godkänd
- I en skriftlig rapport lösa ett problem genom att implementera metoder och modeller från maskininlärning på ett fördjupat sätt med hög säkerhet.
- Redogöra för och kritiskt diskutera modellval, modellanpassning och modellutvärdering

Icke Godkänd ges till studerande som har fullföljt kursen men inte nått alla mål för kursen.

Utbildare

Utbildare: Linus Rundberg Streuli, Antonio Prgomet.

E-post: Omniway.

Tillgänglighet: Möjligheten att ställa frågor och diskutera med utbildaren sker på lektionstid och vid behov via mejl funktionaliteten på Omniway. Försök att nyttja lektionstiden framför mejl för kommunikation.

Kursmaterial

Typ av material	Kommentar
Kursens GitHub sida: https://github.com/AntonioPrgomet/ds24_ml.git	Allt kursmaterial finns på GitHub länken.
Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems av Aurelien Geron (andra eller tredje upplagan). ISBN: 9781098125974 Boken finns i pdf format och har skickats till er på Omniway. Ni rekommenderas att ha den fysiska boken för effektiv studieteknik. Bokens finns att köpa t.ex. här: https://www.bokus.com/bok/9781098125974/hands-on-machine-learning-with-scikit-learn-keras-and-tensorflow-3e/	Finns att beställa på t.ex. Bokus eller Adlibris. Se vart den är billigast.

Kunskapskontroll - Regler

Resultat och betyg registreras senast inom 10 arbetsdagar från deadline. Missar du ordinarie examination kontakta din Utbildningsledare och informera varför. Missar man ordinarie deadline för kunskapskontrollen eller inte klarar den så är andra chansen 3 veckor efter ordinarie deadline. Du kan lämna in när du vill under denna period. Missar man deadline för andra chansen eller inte klarar den så är tredje chansen 7 veckor efter ordinarie deadline. Du kan lämna in när du vill under denna period. Har du missat samtliga examinationstillfällen så måste du kontakta din utbildningsledare snarast.

Upplägg på Föreläsningar / Lektioner

Lektionerna kommer fokusera på genomgångar och lösningar av uppgifter. Uppstår det frågor under självstudierna så skriv ned dem så tar vi upp dem på lektionstid. Viktigt att man *inför* lektionerna arbetat med det material som förväntas enligt planeringen.

Schema:

	Förmiddag: 08.15 – 12.00.	Eftermiddag: 13.15 – 17.00
Måndag	Egenstudier för båda orter	Egenstudier för båda orter
Tisdag	Helsingborg / Malmö	Göteborg
Onsdag	Helsingborg / Malmö	Göteborg
Torsdag	Egenstudier för båda orter	Egenstudier för båda orter
Fredag	Egenstudier för båda orter	Egenstudier för båda orter

Veckoplanering – Vad skall jag göra varje arbetsdag?

I detta avsnitt så framgår i detalj vad som skall göras varje dag. Viktigt att du följer schemat.

	Kursvecka 1 (v.7)
Mån	<p>Kolla video "Introduktion till Maskininläring": https://www.youtube.com/watch?v=zORv5vrxwok&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=2&t=1719s</p> <p>Kolla följande videos "Klassificering": https://www.youtube.com/watch?v=-OzFLifZgNw&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=2</p> <p>"Linjär Regression & Logistisk Regression": https://www.youtube.com/watch?v=0f1tWgarTkQ&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=4&t=1912s</p> <p>Läs och lek med koden "Regression - Kod Exempel": https://www.youtube.com/watch?v=T24GpzKHSI4&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=4&t=101s&pp=gAQBIAQB</p> <p>Läs och lek med koden "Klassificering - Kod Exempel": https://www.youtube.com/watch?v=MNELnQx7L8M&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=6&t=2s</p> <p>Kolla igenom "instuderingsfrågor" från "I1" mappen på GitHub.</p> <p>Videorna är kopplade till kapitel 1-4 i kursboken. Det kommer ni läsa under veckan.</p>
Tis	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <p>EM: Arbeta med "instuderingsfrågor" från lektionen. Finns på Github under "I1" mappen.</p>
Ons	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <p>- Diskussion av "instuderingsfrågor".</p>
Tor	<p>Notera, du kommer inte förstå allt när du läser. Häng inte upp dig på detaljer då det inte är en bok gjord för att man ska förstå alla detaljer.</p> <p>Läs kapitel 1 i kursboken. - Vill du arbeta med uppgifter så prioritera dessa: 1, 2, 3, 4, 7, 12, 15, 16, 17, 19.</p> <p>Läs kapitel 2 i kursboken. - Vill du göra några uppgifter så prioritera dessa: 1, 4.</p> <p>Läs kapitel 3 i kursboken. - Om du har tid över jobba med uppgift 3 "Titanic Dataset".</p> <p>Läs kapitel 4 i kursboken.</p>

Fre	<p>Läs kapitel 1 i kursboken.</p> <ul style="list-style-type: none"> - Vill du arbeta med uppgifter så prioritera dessa: 1, 2, 3, 4, 7, 12, 15, 16, 17, 19. <p>Läs kapitel 2 i kursboken.</p> <ul style="list-style-type: none"> - Vill du göra några uppgifter så prioritera dessa: 1, 4. <p>Läs kapitel 3 i kursboken.</p> <ul style="list-style-type: none"> - Om du har tid över jobba med uppgift 3 "Titanic Dataset". <p>Läs kapitel 4 i kursboken.</p>
------------	--

	Kursvecka 2 (v.8)
Mån	Arbeta med materialet "kategorisk_data" som finns på GitHub.
Tis	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <ul style="list-style-type: none"> - Genomgång av hur man hanterar kategorisk data i ML modeller. - Utdelning av kunskapskontroll 1. <p>EM: Arbeta med kunskapskontroll 1.</p>
Ons	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <ul style="list-style-type: none"> - Diskussion av kunskapskontroll 1.
Tor	Deadline kunskapskontroll 1 kl: 17.00. Det är en god vana att följa vanliga arbetstider. För de som önskar är det ok att lämna in senast kl: 23.59.
Fre	<p>Kolla på videon "Design Principles of Scikit-Learn: Theory and Code Demonstrations": https://youtu.be/Fy9o4pXRVzs och jobba igenom de olika skripten tillgängliga på GitHub, länk till GitHub finns i videobeskrivningen.</p>

	Kursvecka 3 (v.9)
Mån	<p>Kolla video "Support Vector Machines": https://www.youtube.com/watch?v=NzGiMYinmog&list=PLgzaMbMPEHEEx9Als3F3sKKXexWnyEKH45&index=7</p> <p>Läs kapitel 5 i kursboken.</p> <ul style="list-style-type: none"> - Avsnitten under "Under the Hood" kan du läsa översiktligt för att se helhetsbilden, men detaljerna är överkurs och ingår inte i kursen. - Jobba med följande uppgifter: 1, 2, 3, 6 från kapitel 5.
Tis	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <p>EM: Arbeta med uppgiften "mnist_svm", finns tillgänglig på kursens GitHub. Kunskapskontroll 2 kommer vara baserad på denna uppgiften så bra att ni arbetar med den.</p>
Ons	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <p>- Diskussion av uppgiften "mnist_svm".</p>
Tor	<p>Kolla video om "Beslutsträd": https://www.youtube.com/watch?v=X-uXL4x2Ap0&list=PLgzaMbMPEHEEx9Als3F3sKKXexWnyEKH45&index=8</p> <p>Läs kapitel 6 i kursboken.</p> <ul style="list-style-type: none"> - Vill du göra några uppgifter så prioritera uppgifterna: 3, 4.
Fre	<p>Arbeta med uppgiften "moon_tree", finns tillgänglig på kursens GitHub.</p>

	Kursvecka 4 (v.10)
Mån	<p>Kolla video "Ensemble Learning & Random Forest": https://www.youtube.com/watch?v=-kErwtsQRKk&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=9</p> <p>Läs kapitel 7 i kursboken (allt under avsnittet "Boosting" kan du skippa då det inte ingår i kursen, alternativt så läser du det kort och översiktligt för att ha "hört det" någon gång).</p> <p>- Vill du göra några uppgifter så prioritera uppgifterna: 1, 2.</p> <p>Arbeta med uppgiften "mnist_ensemble" som finns på kursens GitHub.</p>
Tis	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <p>- Diskussion av uppgifterna "moon_tree" och "mnist_ensemble".</p>
Ons	Lektion via Teams kl: 08.15 - 12.00.
Tor	Repetition.
Fre	<p>Kolla video "Dimensionsreducering - PCA": https://www.youtube.com/watch?v=VHZRRXEZijA&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=10</p> <p>Läs kapitel 8 i kursboken.</p> <p>- Jobba med följande uppgifter: 1, 2, 4.</p>

	Kursvecka 5 (v.11)
Mån	<p>Kolla video "Unsupervised Learning Techniques" (OBS: knappt 3 h lång): https://www.youtube.com/watch?v=Pj-eAU1hrGk&list=PLgzaMbMPEHEX9Als3F3sKKXexWnyEKH45&index=11</p> <p>Läs kapitel 9 i kursboken.</p> <p>- Allt efter "Gaussian Mixtures" ingår inte i kursen, skippa det.</p> <p>- Jobba med följande uppgifter: 1, 2.</p>
Tis	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <p>- Utdelning av kunskapskontroll 2.</p>
Ons	Lektion via Teams kl: 08.15 - 12.00.
Tor	Arbeta med kunskapskontroll 2.
Fre	Arbeta med kunskapskontroll 2.

	Kursvecka 6 (v.12)
Mån	Arbeta med kunskapskontroll 2.
Tis	Lektion via Teams kl: 08.15 - 12.00.
Ons	Lektion via Teams kl: 08.15 - 12.00.
Tor	Arbeta med kunskapskontroll 2.
Fre	Deadline kunskapskontroll 1 kl: 17.00. Det är en god vana att följa vanliga arbetstider. För de som önskar är det ok att lämna in senast kl: 23.59.