Python med Artificiell Intelligens

|  |  |
| --- | --- |
| **Utbildning och ort:** | Objektorienterad programmering med AI kompetens, Malmö |
| **Vår klassbeteckning:** | OPA23MA |
| **Utbildningsnummer och omgång:** | YH00947-2021-3 |
| **Omfattning i poäng:** | 30 yrkeshögskolepoäng |
| **Engelsk översättning**: | Python with Artificial Intelligence, 30 HVE credit points |
| **Språk:** | Svenska |
| **Förkunskapskrav:** | Inga |
| **Fastställd i ledningsgruppen:** | 2023-XX-XX |

**Kursens innehåll**

Kursen ger den studerande de teoretiska och praktiska grunderna av programmering i Python, där de praktiska elementen har den specifika inriktningen maskininlärning.

* Utvecklingsverktyg
* IDE och command line-verktyg
* Anaconda

Grundläggande syntax:

* Variabler, datatyper
* Styrstrukturer för selektioner och iterationer
* Datastrukturer
* Funktioner, moduler
* För kännedom: Objektorientering med Python
* klasser, objekt
* Arv, polymorfism
* Grundläggande Linjär Algebra grundläggande statistik
* Verktyg för Machine Learning: Jupyter notebooks
* Verktyg för Machine Learning: Biblioteken NumPy, Pandas och Matplotlib
* Filhantering
* Data cleaning & preparation

**Kursens mål**

Efter genomförd kurs med godkänt resultat ska den studerande kunna:

**Kunskaper**

1. Redogöra för grundläggande syntax, styrstrukturer och datastrukturer inom Python.
2. Hantera Objektorientering inom Python.
3. Grundläggande förståelse för Linjär Algebra såsom vektorer, matriser och operationer med dessa.
4. Grundläggande förståelse för Statistik

**Färdigheter**

1. Skriva funktioner och enklare program i Python.
2. Använda Pythons standardbibliotek
3. Använda Jupyter notebooks.
4. Använda sig av biblioteken Numpy, Pandas och Matplotlib.
5. Använda Statistik för att få ökad förståelse för data

**Kompetenser (ansvar och självständighet)**

1. Använda biblioteken NumPy, Pandas och Matplotlib.
2. Redovisa resultat på ett professionellt sätt.

**Former för kunskapskontroll**

Den studerandes kunskaper, färdigheter och kompetenser bedöms utifrån resultatet av följande kunskapskontroll/er:

Kursens läranderesultat 1, 2, 5, 6, 7: En skriftlig och muntlig gruppuppgift (IG/G).

Kursens läranderesultat 3, 4, 8, 9: En skriftlig individuell inlämningsuppgift (IG/G/VG).

Vid särskilda behov kan anpassning av kunskapskontrollerna göras. Efter ordinarie tillfälle har den studerande rätt till ytterligare två omprov eller kompletteringar inom två månader. Om den studerande kan styrka giltig frånvaro från kunskapskontroll enligt anordnarens anvisningar (t ex med sjukintyg) ersätts tillfället.

Kursbetyget baseras på en sammanvägning av samtliga bedömningsunderlag

**Principer för betygssättning**

Den studerandes prestation betygssätts efter genomförd kurs med betygen Icke Godkänt (IG), Godkänt (G) eller Väl Godkänt (VG).

**Icke godkänt (IG)**

Den studerande har fullföljt kursen men inte nått alla mål för kursen

**Godkänt (G)**

Den studerande har nått samtliga mål för kursen

**Väl Godkänt (VG)**

För att få betyget Väl Godkänt (VG) ska den studerande dels ha genomfört kursen och nått alla kursens läranderesultat, dels uppfylla kravet att kunna ta ansvar för och självständigt lösa problem i Python med hög säkerhet samt med högre kvalitet än för betyget G. Den studerande ska på ett reflekterande sätt kunna ta ett fullständigt och självständigt ansvar för genomförande, dess förutsättningar, risker, begränsningar, effekter, resultat och förbättringsmöjligheter.

# Utbildare

Namn: Antonio Prgomet.   
LinkedIn: <https://www.linkedin.com/in/antonioprgomet/>

E-post: [antonio.ek@hotmail.se](mailto:antonio.ek@hotmail.se)

Telefon: 0720 – 16 08 81

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

## Kursmaterial

|  |  |
| --- | --- |
| **Typ av material** | **Kommentar** |
| **Kursens GitHub sida där material kommer laddas upp:** [**https://github.com/AntonioPrgomet/nbi\_python\_AI\_1**](https://github.com/AntonioPrgomet/nbi_python_AI_1) | Allt kursmaterial finns på GitHub länken.  Vill du lära dig Git & GitHub kan du t.ex. se följande video: <https://www.youtube.com/watch?v=SWYqp7iY_Tc&t=1293s> |
| Körner, S. & Wahlgren, L. (2015). Statistisk dataanalys. (5. uppl.). Lund: Studentlitteratur. | ISBN: 9789144108704  Kan köpas t.ex. här: <https://www.adlibris.com/se/bok/statistisk-dataanalys-9789144108704>   OBS: Införskaffas inför kursstart. |
| Körner, S. & Wahlgren, L. (2016). Tabeller och formler för statistiska beräkningar. (3. uppl.). Lund: Studentlitteratur. | ISBN: 9789144114545. Kan köpas t.ex. här: <https://www.adlibris.com/se/bok/tabeller-och-formler-for-statistiska-berakningar-9789144114545>  OBS: Införskaffas inför kursstart. |
| Förinspelade videos:   * Grundkurs i Python: <https://youtu.be/M7bnYJyCx0Q> * Statistik: <https://www.youtube.com/playlist?list=PLgzaMbMPEHEwkc-XVv3gpprOk7y2IHWLJ> * Om man vill lära sig Excel (inget som ingår i kursen) kan man göra det med denna videon: <https://youtu.be/4UMLFC1SoHM> |  |
| 8 videor om studieteknik, de kommer hjälpa dig mycket under dina studier och även senare i arbetslivet när du skall lära dig nya saker. <https://www.youtube.com/watch?v=gSbpRjxYq24&list=PLA09CC1B5671827AD> |  |
| Önskar man använda DataCamp så kan man använda länken nedan som ger tillgång fram till 12 Maj, 2024.  <https://www.datacamp.com/groups/shared_links/bbd0c6d3cbe430a45c6a6aa552daacda0adf67a5cf6c6b003b6ba3b69a4c323e>  Vi kommer I denna kursen inte använda DataCamp, men vill man bli väldigt duktig på Python så kan man t.ex. göra följande kurer vid tillfälle (71 timmar och 18 kurser så endast för den som har tid): <https://app.datacamp.com/learn/career-tracks/python-programmer> | DataCamp online kurser. Registrera dig i DataCamp med din skolmejl (t.ex. [malte.maltesson@utb.ecutbildning.se](mailto:malte.maltesson@utb.ecutbildning.se) ) **annars fungerar det inte. Skapa inbjudan för NBI mejladressen.** |

# Kunskapskontroll - Regler

**Resultat och betyg registreras senast inom 10 arbetsdagar från deadline.**

* **Har du inte lämnat in din examination i tid eller inte blivit godkänd så är *andra* examinationstillfället senast 7 dagar efter ordinarie deadline eller att du fått din kunskapskontroll rättad. Du kan lämna in din kunskapskontroll när du vill under dessa 7 dagarna.**
* **Har du inte lämnat in din examination i tid eller inte blivit godkänd så är *tredje* examinationstillfället senast 28 dagar efter ordinarie deadline eller att du fått din kunskapskontroll rättad. Du kan lämna in din kunskapskontroll när du vill under dessa 28 dagarna.**

**Har du missat samtliga examinationstillfällen så måste du kontakta din utbildningsledare snarast.**

# Kunskapskontroll – Information

Kursen består av två examinerande kunskapskontroller. Se veckoplaneringen längre ned för detaljer kring upplägg.

* Den första kunskapskontrollen har deadline första kursveckan.
* Den andra kunskapskontrollen består av 3 delar och vi kommer jobba med dem löpande.

Majoriteten av det vi arbetar med på lektionerna är direkt kopplat till kunskapskontrollen. Arbetet skall reflektera ens egna arbete, men precis som i verkligheten så uppmuntrar jag till att ni samarbetar och lär av varandra. Det som betygsätts är era kunskaper, färdigheter, kompetenser och det är just det vi skall försöka utveckla! Så samarbeta, dela med dig av din kunskap och erfarenhet och ta även del av andras.

* För att få VG i kursen så måste man få VG på kunskapskontroll 1 samt de 3 delarna i kunskapskontroll 2.

# 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. **Förbered på att detta kommer bli en krävande kurs men som tidigare klasser har tyckt är väldigt användbar och utvecklande.**

# Schema:

|  |  |
| --- | --- |
| **Dag** | **Form** |
| Tisdagar Lektion kl: 09.00 – 16.00. | På plats i Halmstad. |
| Torsdagar Lektion kl: 08.00 – 16.00. | På distans via Teams. |
| Fredagar ”Öppen frågestund” kl: 09.00 – 12.00 | På distans via Teams.  1.5h gemensam repetition på sådant ni önskar repetera. Sista 1.5h kan vi boka in samtal individuellt eller i grupp om flera vill diskutera liknande saker. (mejl Sandra 2023-11-08) |

# 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 och kom ihåg att ingen förstår 100 % första gången man gör något. Vissa saker från vecka 1 kommer man t.ex. förstå djupare först under vecka 2. Så är det för alla.**

|  |  |
| --- | --- |
|  | **Kursvecka 1 (v.1) - Introduktion till Python** |
| **Mån** | Kolla på inspelad föreläsning och experimentera med tillhörande kod: <https://www.youtube.com/watch?v=M7bnYJyCx0Q&t=4478s>  från tiden 0:00 till 1:02:20.   * Innan du kollar på videon om mängder (”Sets” på engelska) vid tidpunkt [39:38] så läs igenom texten och kolla videorna i följande länk om Mängdlära: <https://www.matteboken.se/lektioner/matte-5/mangdlara> .   Kolla på inspelad föreläsning: <https://www.youtube.com/watch?v=M7bnYJyCx0Q&t=4478s>  från 1:02:20 till 1:14:59. |
| **Tis** | Lektion kl: 09.00 - 16.00 på plats i Halmstad.  - Vi kommer arbeta med "kunskapskontroll\_1". |
| **Ons** | Arbeta med "kunskapskontroll\_1". |
| **Tor** | Lektion kl: 08.00 - 16.00 på distans via Teams.  - Vi kommer arbeta med "kunskapskontroll\_1" och ha muntliga presentationer på vad som gjorts. |
| **Fre** | Repetition & Frågestund kl: 09.00 - 12.00 via Teams. **Inlämning av kunskapskontroll 1 i LearnPoint. Deadline senast kl: 23.59.**   * Skumma igenom dokumenten: ”summation” och " linear\_algebra". |

|  |  |
| --- | --- |
|  | **Kursvecka 2 (v.2) - NumPy, Matplotlib och Pandas** |
| **Mån** | **Läs dokumenten ”summation” och "linear\_algebra".**   Kolla på inspelad föreläsning och experimentera med tillhörande kod: <https://www.youtube.com/watch?v=M7bnYJyCx0Q&t=4478s>  från tiden 1:14:59 till slutet.  Läs länkarna nedan och experimentera med koden:   * <https://numpy.org/doc/stable/user/quickstart.html> * <https://matplotlib.org/stable/tutorials/introductory/usage.html#sphx-glr-tutorials-introductory-usage-py> * <https://matplotlib.org/stable/tutorials/introductory/pyplot.html#sphx-glr-tutorials-introductory-pyplot-py> * <https://pandas.pydata.org/docs/user_guide/10min.html> |
| **Tis** | Lektion kl: 09.00 - 16.00 på plats i Halmstad.  **- Genomgång av ”summation” och "linear\_algebra".**  - Arbete med Kunskapskontroll 2 - Arbetsblad 1. |
| **Ons** | - Arbete med Kunskapskontroll 2 - Arbetsblad 1. |
| **Tor** | Lektion kl: 08.00 - 16.00 på distans via Teams.  - Arbete med Kunskapskontroll 2 - Arbetsblad 1. |
| **Fre** | Repetition & Frågestund kl: 09.00 - 12.00 via Teams. **- Inlämning av kunskapskontroll 2 - Arbetsblad 1 i LearnPoint, deadline senast kl: 23.59.** |

|  |  |
| --- | --- |
|  | **Kursvecka 3 (v.3) - Sannolikhetsteori** |
| **Mån** | Spellistan där kursvideos finns: <https://www.youtube.com/playlist?list=PLgzaMbMPEHEwkc-XVv3gpprOk7y2IHWLJ>   - Kolla video 1: Introduktion - Sannolikhetsteori & Statistisk Inferens  - Kolla video 2: Introduktion - Sannolikhetsteori   Läs kapitel 1 i kursboken.  - Avsnitt 1.4 "Odds" är mindre viktigt, läs det översiktligt.   Arbeta med uppgifterna:  - 101-104 - 106-107 - 109-114 - 115-121 |
| **Tis** | Lektion kl: 09.00 - 16.00 på plats i Halmstad. - Arbete med Kunskapskontroll 2 - Arbetsblad 2. |
| **Ons** | Kolla video 3: Sannolikhetsbegreppet.   Läs kapitel 2 i kursboken.   Arbeta med uppgifterna:  - 201-203 - 204-207 - 208-213 - 214-217 - 218-221 - 222-227 |
| **Tor** | Lektion kl: 08.00 - 16.00 på distans via Teams.  - Arbete med Kunskapskontroll 2 - Arbetsblad 2. |
| **Fre** | Repetition & Frågestund kl: 09.00 - 12.00 via Teams. |

|  |  |
| --- | --- |
|  | **Kursvecka 4 (v.4) - Sannolikhetsteori** |
| **Mån** | Kolla video 4 "Diskret slumpvariabel". Läs kapitel 3 i kursboken.  Arbeta med uppgifterna: - 301 - 304, 305 - 306 - 307-313 - 314-316 - 317 - 318-320 |
| **Tis** | Lektion kl: 09.00 - 16.00 på plats i Halmstad.  - Arbete med Kunskapskontroll 2 - Arbetsblad 2. |
| **Ons** | Repetition. |
| **Tor** | Lektion kl: 08.00 - 16.00 på distans via Teams.  - Arbete med Kunskapskontroll 2 - Arbetsblad 2. |
| **Fre** | Repetition & Frågestund kl: 09.00 - 12.00 via Teams. **- Inlämning av kunskapskontroll 2 - Arbetsblad 2 i LearnPoint, deadline senast kl: 23.59.** |

|  |  |
| --- | --- |
|  | **Kursvecka 5 (v.5) - Sannolikhetsteori** |
| **Mån** | Kolla video 5: "Tvådimensionell Slumpvariabel". Läs kapitel 4 i kursboken.   Arbeta med uppgifterna: - 401-404, 407-412, 413 (I uppgift 413 härleder ni väntevärdet och variansen för Binomialfödelningen som är en summa av Tvåpunktsfördelade/Bernoulli slumpvariabler) 414-416. |
| **Tis** | Lektion kl: 09.00 - 16.00 på plats i Halmstad.  - Arbete med Kunskapskontroll 2 - Arbetsblad 3. |
| **Ons** | Kolla video 6: "Normalfördelningen".  Läs kapitel 5 i kursboken.   Arbeta med uppgifterna: - 501-508 - 509-512  - 513-517 |
| **Tor** | Lektion kl: 08.00 - 16.00 på distans via Teams.  - Arbete med Kunskapskontroll 2 - Arbetsblad 3. |
| **Fre** | Repetition & Frågestund kl: 09.00 - 12.00 via Teams. |

|  |  |
| --- | --- |
|  | **Kursvecka 6 (v.6) - Statistisk Inferens & Sista Kursveckan** |
| **Mån** | Kolla video 7 "Slumpmässigt urval och punktskattning".  Läs kapitel 6 i kursboken.   Arbeta med uppgifterna: - 601-609 |
| **Tis** | Lektion kl: 09.00 - 16.00 på plats i Halmstad.  - Vi kommer diskutera Kunskapskontroll 2 - Arbetsblad 3. - Vi kommer även diskutera kapitel 6 i kursboken (inget som examineras men användbart kapitel). |
| **Ons** | Kolla video 8 "Konfidensintervall" och läs kapitel 7 i kursboken.   Arbeta med uppgifterna: - 701, 702 - 703,  - 704, 705 - 706 - 710 |
| **Tor** | Lektion kl: 08.00 - 16.00 på distans via Teams.  - Vi kommer diskutera Kunskapskontroll 2 - Arbetsblad 3. - Vi kommer även diskutera kapitel 7 i kursboken (inget som examineras men användbart kapitel). |
| **Fre** | Repetition & Frågestund kl: 09.00 - 12.00 via Teams. **Inlämning av kunskapskontroll 2 - Arbetsblad 3 i LearnPoint, deadline senast kl: 23.59.** |

# Videos Kopplat till Statistisk Dataanalys

I veckoplaneringen längre ned så är videos jag spelat in inkluderade. Här är fler videos som ni kan kolla på för att få bredare och djupare förståelse. Kapitel hänvisningarna nedan refererar till boken ”Statistisk Dataanalys” av Körner & Wahlgren (2015).

Jag rekommenderar att ni kollar på videorna som ett komplement till de jag gjort så får ni två förklaringar på många av koncepten.

# Kapitel 1-2

För kapitel 1-2, kolla följande spellista som en ”introduktion”: <https://www.youtube.com/watch?v=B1v9OeCTlu0&list=PLvxOuBpazmsOGOursPoofaHyz_1NpxbhA&index=1>

1. Video 5: *What Does Independence Look Like on a Venn Diagram?* Kan du skippa.
2. En extra video om Kombinatorik (kap 1.6): <https://www.youtube.com/watch?v=XJnIdRXUi7A&t=79s>

# Kapitel 3:

För kapitel 3 kolla igenom följande spellista innehållande 13 videos: <https://www.youtube.com/watch?v=oHcrna8Fk18&list=PLvxOuBpazmsNIHP5cz37oOPZx0JKyNszN>

1. Video 3 som handlar om Bernoulli fördelningen är det boken kallar för ”Tvåpunktsfördelad variabel”.
2. På s.91 i boken så framgår det att man kan approximera binomialfördelningen med Poissionfördelningen, för detta se video 9 och 10, video 10 bevisar det och är”överkurs” men det är ett vackert bevis så kolla gärna på det.
3. Video 6 handlar om ”Geometric Distribution” som boken kallar för ffg-fördelning.
4. Video 11 kopplat till ”Negative Binomial Distribution” ingår inte i kursen och kan skippas.
5. Video 12 kopplat till ”Multinomial Distribution”, är likt trinomial fördelningen som dyker upp först i kapitel 4.6 i boken. Förstår man den ena så förstår man den andra.

Hur man använder Tabell 1: Binomialfördelningen i formelsamlingen kan ses här: [**https://www.youtube.com/watch?v=gfDWDujLtfM&t=71s**](https://www.youtube.com/watch?v=gfDWDujLtfM&t=71s)

# Kapitel 4:

**Kapitel 4.4**

* <https://www.youtube.com/watch?v=KDw3hC2YNFc>
* <https://www.youtube.com/watch?v=85Ilb-89sjk>

**Kapitel 4.5**

* Härledning av räkneregler för varians: <https://www.youtube.com/watch?v=zdhkXWyy0K0> , i videon används den generella egenskapen av förväntans värdet att:   
  E[aX + bY + c] = aE[X] + bE[Y] + c, dvs. förväntans värdet är en linjär operation.

**Kapitel 4.6**

Multinomial fördelningen är väldigt lik trinomial fördelningen, förstår man den ena så förstår man den andra. I följande video förklaras multinomial fördelningen: <https://www.youtube.com/watch?v=syVW7DgvUaY&t=382s>

# Kapitel 5:

Kolla först på video 1-7 i följande spellista om kontinuerliga sannolikhetsfördelningar:

<https://www.youtube.com/watch?v=OWSOhpS00_s&list=PLvxOuBpazmsPDZGwqhhjE3KkLWnTD34R0>

* Vi är hittills vana vid diskreta fördelningar och vi vet t.ex. att eftersom summan av alla sannolikheter är 1. I det kontinuerliga fallet så använder man integraler istället för summor, motsvarande formel blir: där vi integrerar från minus oändligheten till oändligheten. I det diskreta hade vi en ”Probability mass function” medan vi i det kontinuerliga fallet hade en ”probability density function” .
* På liknande sett vet vi att väntevärdet för diskreta fördelningar kan räknas ut enligt följande formel: E[X] = (d.v.s. utfall multiplicerat med sannolikhet). För kontinuerliga fördelningar får vi: E[X] = .   
  Notera att vi kan använda t, x eller någon annan integrationsvariabel, det spelar ingen roll: = .
* Kontinuitetskorrektion (se sid. 137 i boken) gås igenom i video 7 vid tiden: 5:42. Bra förklarat på något som kan upplevas förvirrande i början.

Hur man använder normalfördelningstabellen (tabell 3a i formelsamlingen): <https://www.youtube.com/watch?v=p_KApjpyBHE>.

För kapitel 5.3, se följande spellista: <https://www.youtube.com/watch?v=Zbw-YvELsaM&list=PLvxOuBpazmsP7UN00cNZX64N1o_8635ds>

# Kapitel 6:

6.3: <https://www.youtube.com/watch?v=xJlwSkyeP0k>

# Kapitel 7

Se hela spellistan som är kopplad till kapitel 7:

<https://www.youtube.com/watch?v=27iSnzss2wM&list=PLvxOuBpazmsMdPBRxBTvwLv5Lhuk0tuXh>