

Kurs PM

Data engineering och agila metoder 45 p

Beskrivning av kursen/innehåll:

I verkligheten är data sällan städad och lätt att tolka. Denna kurs behandlar hur man skapar så goda förutsättningar som möjligt för maskininlärning i praktiken och hur en modell går från idé till driftsatt värdeskapande lösning. Kursen behandlar olika **koncept inom data engineering för att bygga en infrastruktur för datan** och förstå hur den på smidigt sätt kan användas och struktureras för ett maskininlärningsprojekt. Kursen har även ett avsnitt med etik och regler inom området. I kursen görs ett **mindre projekt i agil form** där modellen driftsätts i molnet med exempelvis Amazon Web Services (AWS), **Microsoft Azure** eller Google Cloud Platform.

Kursen syftar också till att ge den studerande kunskap och färdigheter inom projektmetodik, projektutformning och agila metoder för att kunna delta i eller leda ett agilt utvecklingsprojekt. Kursens mål är att studerande ska lära sig navigera problem kring data i verkligheten, kunna driftsätta och underhålla en AI-modell i molnet samt kunna jämföra olika lösningar. Kursens syfte är att ge studerande praktiska kunskaper och färdigheter att kunna arbeta med att utveckla, nyttja och underhålla AI- lösningar under verkliga förhållanden

Teknologier/verktyg som ska användas:

T.ex

- Visual Studio Code
- `venv`
- Git/GitHub
- Github actions
- Pre-commit
- Data version control (DVC)
- Docker
- Linux och bash
- Windows subsystem for linux (WSL)
- Python och några paket: Numpy, Pandas, Scikit-learn, Airflow
- Visualiseringsverktyg – matplotlib, seaborn, plotly
- Azure

Utbildare

- Kokchun Giang
- Olof Harrysson

Datum

Kursen är uppdelad i två delar, en del 1 före sommar och del 2 efter sommar. Del 1 undervisas av Kokchun och del 2 av Olof.

Del 1: 2023-05-22 – 2023-06-09

Del 2: 2023-08-14 – 2023-09-22

Kurslitteratur:

Bloggar, handouts, kompendier och videoklipp som utbildaren tillhandahåller:

- Finns samlat i kursens [Github repo](#)

Examinationsformer:

- Projekt i grupp med muntlig presentation och skriftlig rapport

Betygssättning:

Studerande ska få sitt kursbetyg senast 3 veckor efter det att kursen avslutats

Betygskriterier:

Kursen bedöms med betygen Icke Godkänd, Godkänd eller Väl godkänd. Grunden för betyg avgörs av hur väl studenten har uppfyllt betygskriterierna.

Studenten ska få sitt kursbetyg senast 3 veckor efter det att kursen avslutats.

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

1. Projektplanering
2. Agila projektmetoder, t.ex. Scrum, Kanban och andra koncept
3. Data engineeringkoncept som data pipelines, data processing
4. Databehandling inom tex spark och databricks
5. Packetering av mjukvara tex via Docker containers och Kubernetes
6. Grundläggande kunskaper och förståelse om molntjänster exempelvis AWS, Azure, Google Cloud
7. Kodgranskning
8. Kännedom om de etiska frågeställningar som uppkommer vid behandling av data.
9. Förmåga att använda agila koncept och ett agilt arbetssätt
10. Planera, utföra, dokumentera och utvärdera projekt enligt agila projektmetoder
11. Använda data engineering i molnet för ett projekt
12. Driftsätta, monitorera och underhålla en AI-modell i molnet
13. Applicera AI för att lösa verkliga problem
14. Diskutera och resonera kring etiska förhållningssätt och regler inom området
15. Nyttja tekniker för att anpassa verklig data för maskininlärningsmodeller
16. I team kunna bygga upp data engineering i molnet för ett projekt
17. Kunna designa, implementera och driftsätta en data engineering lösning under verkliga förhållanden

Väl godkänt (VG)

Den studerande har nått samtliga mål för kursen. Den studerande kan dessutom:

18. Med säkerhet utföra de olika stegen i arbetsflödet för att lösa verkliga problem med hjälp av maskininläring och data engineering
19. Med säkerhet kunna utnyttja en molntjänst i arbetsflödet för data engineering lösningar på ett kostnadseffektivt sätt

Examinationer:

Innehåll	Introduktion	Deadline	Täcker upp betygskriterier
Projekt	2023-08-21	2023-09-15	1-19