

Kursus i algoritmiske grundprincipper og etik i AI: fra teori til praksis

Værktøjer til synkrone sessioner

Projektbaserede øvelser

Projektbaserede øvelser



Projektbaserede øvelser

	Beskrivelse
Beskrivelse af opgaven	I hold på 4-5 personer skal I arbejde med at evaluere en af seks AI-løsninger i løbet af fem kursusgange. Denne omfattende opgave indebærer en detaljeret analyse af AI-løsningen fra flere perspektiver i overensstemmelse med kursusstrukturen og de emner, der behandles i hver kursusgang.
Beskrivelse af, hvordan opgaven skal udføres	Instruktioner • Dit team skal arbejde på en AI-løsning fra listen på side 4. • Analysér den valgte løsning ud fra de fem perspektiver på side 6. • Dokumentér dine resultater i en logisk opbygget rapport, der behandler hvert evalueringskriterium i detaljer. Hvert punkt skal præsenteres på maks. 2 slides • Præsenter din analyse for klassen i den sidste synkrone session, og fremhæv de vigtigste læringspunkter og anbefalinger til forbedring af løsningen. • Din præsentation for teamet skal vare mellem 5-10 minutter
Anslået tid til opgaven	Ca. 1-2 timer / CU (5-10 timer i alt)
Forslag til kilder til at udføre opgaven	Interaktive hæfter og andet kursusmateriale + eventuelle andre kilder
Detaljeret beskrivelse af, hvordan opgaven skal udføres	Se de næste sider.
Information om deadline for aflevering af opgaven	Præsenteres i den sidste synkrone session.
Kontaktoplysninger eller hvordan man afklarer tvivlsspørgsmål	Læreren skal angive en form for kontaktmulighed (det kan være en e-mailadresse, et telefonnummer eller lign.).

Projektbaserede øvelser

Tidslinje og instruktioner

TIDSPLAN FOR PROJEKTØVELSEN

Præsentation af projektet og dannelse af teams

- Begynd så vidt muligt at arbejde på projektet efter den første synkrone session, og aftal med teamet, hvilke møder der skal afholdes for at arbejde på projektet.
- Alternativt kan du starte efter den 2. synkrone session, der dækker CU 1 og 2.

Gennemgå fremdriften

- På den 3. og 4. synkrone session er der mulighed for at gennemgå fremdriften og diskutere eventuelle udfordringer med læreren.

Endelig præsentation

- På den 5. synkrone session præsenteres projektøvelsen i teams.
 - Præsentation 5-10 minutter.
 - Du kan vælge PowerPoint- eller Canva-layout; husk at lave en indbydende præsentation.

BEMÆRK | Hvis en deltager ikke deltager i gruppearbejdet (eller ikke deltager aktivt i projektet), skal projektopgaven afleveres individuelt til læreren (10 slides).

Yderligere instruktioner

- Dit emne præsenteres på næste slide. Som team bestemmer I selv navnet, formålet og det mere detaljerede indhold af jeres service. Applikationen skal dog være inden for det felt, du tildeles på næste side.
- Fortsæt med at arbejde på projektet i det aftalte tempo med projektgruppen; indholdet er dog altid relateret til de enkelte kursusenheder (CU'er), så det er bedst at arbejde i gruppen, efter at gruppemedlemmerne har færdiggjort de enkelte e-learningsemner.
- Du kan selv bestemme, hvordan du vil opbygge din præsentation; spørgsmålene er blot vejledende, men du bør skabe et indhold, som engagerer tilhørerne.

Projektbaserede øvelser

AI-løsninger opdelt pr. team

Den AI-løsning, dit team bør arbejde på:

AI-drevet diagnoseværktøj til sundhedssektoren

- Et værktøj, der hjælper læger med at stille foreløbige diagnoser baseret på patientens symptomer og sygehistorie.
 - XX grupper, der arbejder med emnet (tilføj her navnene på de studerende, som læreren har inddelt i grupper på 4-5)

System til anbefaling af streamingtjenester

- En applikation, der foreslår film og serier til brugerne baseret på deres seerhistorik og præferencer.
 - XX grupper, der arbejder med emnet (tilføj her navnene på de studerende, som læreren har inddelt i grupper på 4-5)

Smart home-automatiseringssystem

- En løsning, der automatiserer smart home-enheder såsom lys, termostater og sikkerhedssystemer baseret på brugernes rutiner og præferencer.
 - XX grupper, der arbejder med emnet (tilføj her navnene på de studerende, som læreren har inddelt i grupper på 4-5)

AI-drevet finansiel investeringsrådgiver

- En rådgiver, der giver investeringsanbefalinger baseret på markedsdata og brugerens økonomiske mål.
 - XX grupper, der arbejder med emnet (tilføj her navnene på de studerende, som læreren har inddelt i grupper på 4-5)

Applikation til online shopping

- En app, der anbefaler produkter til brugerne baseret på deres browser- og købshistorik.
 - XX grupper, der arbejder med emnet (tilføj her navnene på de studerende, som læreren har inddelt i grupper på 4-5)

Personaliseret e-læringsplatform

En platform, der tilpasser læringsforløb og indhold til studerende baseret på deres læringsstile, fremskridt og præferencer.

- XX grupper, der arbejder med emnet (tilføj her navnene på de studerende, som læreren har inddelt i grupper på 4-5)

Projektbaserede øvelser

CU 1-relateret indhold

Start med at definere følgende (denne opgave bør ikke fylde mere end 2 præsentationsslides)

Problemformulering: Hvilket problem løser din applikation?

Mål: Hvad er målet med din AI-applikation? Hvad hedder applikationen? Hvad er succeskriterierne for din løsning?

Lav en kortlægning af interessenter, der som minimum skal vise bruger-interessenter, eksterne interessenter og udviklingsinteressenter (brug CU1-materialet som vejledning).

Overvej, hvor godt din applikation opfylder nedenstående etiske principper; hovedfokus bør være på de samfundsmæssige og miljømæssige aspekter:

Gavnlighed

Sørg for, at AI-løsningen fremmer den enkeltes og samfundets ve og vel. Vigtige områder til overvejelse:

- Beskyttelse af grundlæggende rettigheder: Værne om individuelle rettigheder styrker etisk integritet.
- Miljømæssig bæredygtighed: Minimere miljøpåvirkningen gennem en bæredygtig praksis.
- Begrundelse: Definer klart og vis sammenhængen mellem systemets formål og dets konkrete samfundsmæssige fordele.

Ikke-skadelighed (*non-maleficence*)

- Sørg for, at problemformuleringen og den foreslåede løsning ikke skader enkeltpersoner eller samfundsgrupper.
- Overvej og overvåg nøje systemets indvirkning på menneskers fysiske og mentale velbefindende.
- Vurder potentielle sikkerhedsrisici og strategier for afhjælpning.

Retfærdighed

Evaluer systemets indvirkning på institutioner/instanser, demokrati og samfund.

Projektbaserede øvelser

CU 1 (hjælpe slide)

Brug denne tabel til nærmere at vurdere hvert etiske princip (her i en version udarbejdet af Morley et al., 2020 og oversat).

Tabel 5 viser forbindelsen mellem vigtige, etiske principper og konkrete systemkrav; en bearbejdet version af den metode, der er skitseret i kapitel II i Europa-Kommissionens "Ethics Guidelines for Trustworthy AI".

Princip	Gavnlighed	Ikke-skadelighed (non-maleficence)	Autonomi	Retfærdighed	Forklarlighed
Krav	Deltagelse af interessenter: For at udvikle systemer, der er troværdige og understøtter menneskelig fremgang, skal de, der vil blive påvirket af systemet, høres.	Modstandsdygtighed over for angreb og sikkerhed: AI-systemer skal beskyttes mod huller, der kan gøre det muligt for modstandere at udnytte dem.	Menneskelig udførelse: Brugere skal kunne træffe informerede, selvstændige beslutninger om AI-systemer.	Undgå uretfærdig bias	Sporbarhed: Datasæt og de processer, der fører til AI-systemets beslutninger, skal kunne dokumenteres.
	Beskyttelse af grundlæggende rettigheder	Nødplan og generel sikkerhed: AI-systemer bør have sikkerhedsforanstaltninger, der indeholder en nødplan, hvis der opstår problemer.	Menneskelig kontrol: Kan opnås gennem styringsmekanismer som <i>human-on-the-loop</i> , <i>human-in-the-loop</i> , <i>human-in-command</i>	Tilgængelighed og universelt design	Forklarlighed: Evnen til at forklare både de tekniske processer i et AI-system og de tilhørende menneskelige beslutninger.
	Bæredygtig og miljøvenlig AI: Systemets forsyningskæde bør vurderes med hensyn til ressource- og energiforbrug.	Nøjagtighed: F.eks. dokumentation, der viser, om systemet klassificerer resultaterne korrekt.		Samfund og demokrati: Systemets indvirkning på institutioner, demokrati og samfundet som helhed bør overvejes.	Fortolkningsmuligheder
	Begrundelse: Formålet med at konstruere systemet skal være klart og have klare fordele - systemet skal ikke bygges for systemets skyld.	Beskyttelse af personlige oplysninger og personlige data: AI-systemer bør garantere privatlivets fred og databeskyttelse gennem hele systemets livscyklus.		Revision: Mulighed for revision af algoritmer, data og designprocesser.	
		Pålidelighed og reproducerbarhed: Fungerer systemet på samme måde i en række forskellige scenarier?		Minimering og rapportering af negative indvirkninger: Der bør træffes foranstaltninger til at identificere, vurdere, dokumentere, minimere og reagere på potentielle negative indvirkninger af AI-systemer.	
		Datakvalitet og -integritet: Når data indsamles, kan de indeholde socialt konstruerede bias, unøjagtigheder, fejl og mangler - det skal der tages hånd om.		Trade-offs: Når trade-offs mellem krav er nødvendige, bør der indføres en proces, der eksplicit viser den nødvendige trade-off, og hvordan den vil blive håndteret.	
		Social indvirkning: Systemets indvirkning på menneskers fysiske og mentale velbefindende bør overvejes nøje og overvåges.		Klageadgang: Der skal være en mekanisme på plads til at reagere, når noget går galt.	

KILDE | Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Science and Engineering Ethics, 26(4), 2141–2168. <https://doi.org/10.1007/s11948-019-00165-5>

Projektbaserede øvelser

CU-relateret indhold

CU 2

Bekymringer for privatlivets fred

- Identificer de bekymringer for privatlivets fred, der er forbundet med løsningen.
- Overvej, hvordan data indsamles, opbevares og bruges.
- Foreslå foranstaltninger til at beskytte brugernes privatliv og opretholde datasikkerhed.

CU 3

Anvendt algoritme

- Beskriv den algoritme, der ligger bag løsningen.
- Forklar, hvordan algoritmen er opbygget og trænet på baggrund af de indsamlede data (CU2).
- Diskuter eventuelle specifikke tekniske detaljer, der er vigtige for at forstå løsningen.

CU 4

Mulige bias

- Identificer potentielle bias i systemet.
- Diskuter, hvad der kan forårsage disse bias (f.eks. træningsdata, algoritmedesign).
- Foreslå måder at forebygge eller afbøde disse bias på.

CU 5

- TU Delft har lavet en tjekliste til etisk AI-udvikling. Tjek værktøjet via dette [link](#), og brug AI APP-projektets tjekliste på side 12-13, og udfyld tjeklisten på side 13. Vis din tjekliste i slutningen af din projektpresentation.
- Det sidste slide skal vise, hvad I som team har lært, og hvilken ny læring I har opnået ved at lave denne præsentation.

TAK

Projektnummer: 2022-1-ES01-KA220-HED-000085257



Europa-Kommissionens støtte til produktionen af denne publikation udgør ikke en del af indholdet, som kun afspejler forfatterens synspunkter, og Kommissionen kan ikke holdes ansvarlig for nogen brug, der måtte blive gjort af oplysningerne heri.

