

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

Værktøjer til synkrone sessioner

CU4 | Data-fairness og bias i AI
Casestudie

Casestudie



Casestudie

Beskrivelse til undervisere

	Beskrivelse	Kommentarer
Beskrivelse af opgaven	Denne session har til formål at hjælpe eleverne med at forstå det grundlæggende i privatlivets fred og bekvemmelighed, især i dagligdags aktiviteter som smart home-enheder, mobilapps og online shopping. Den vil involvere interaktive diskussioner, refleksioner og aktiviteter for at forstærke læringen.	
Beskrivelse af, hvordan du udfører opgaven	Præsenter casestudiet.	Se instruktionerne og retningslinjerne på de næste sider.
Anslået tid til opgaven	30-40 minutter.	
Forslag til kilder til udførelse af opgaven	Websøgning på hændelser relateret til brud på privatlivets fred eller datalækager.	
Detaljeret beskrivelse af, hvordan opgaven skal løses	Gruppediskussion + Google-formularer.	Se de foreslåede spørgsmål til Google-formularer i dokumentet: CU4_CaseStudy_Tentative_questions_for_google_forms.docx
Information om deadline for aflevering af opgaven	Under den 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.).	

Casestudie

Fiktiv case for studerende

Casestudie: AI-retfærdighed og bias i uddannelsessystemer

Et stort offentligt skoledistrikt implementerede et AI-drevet værktøj til at hjælpe med vurdering af elever, som omfattede forudsigelse af elevernes præstationer, identifikation af elever, der kunne have brug for ekstra hjælp, og endda bidrage til beslutninger om optagelse på college. Værktøjet analyserede en lang række data, herunder tidligere præstationer under uddannelsen, fremmøde, socioøkonomisk baggrund og fritidsaktiviteter.

Et år efter implementeringen blev det opdaget, at AI-systemet i uforholdsmæssig høj grad udvalgte elever fra visse demografiske baggrunde (f.eks. elever med lavere socioøkonomisk status eller visse racemæssige og etniske minoriteter) som værende i fare for at underpræstere, selv når deres faktiske præstationer var på niveau med eller bedre end deres jævnaldrende. Disse elever blev ofte anbefalet at deltage i specialtilbud eller blev nægtet adgang til avancerede kurser og fritidsaktiviteter, hvilket havde langsigtede konsekvenser for deres akademiske udvikling og muligheder for at blive optaget på universitetet.

Diskussionspunkter

- Bør fordelene ved AI i uddannelse opveje risikoen for bias? Hvordan finder vi den rette balance?
- Hvordan påvirker data-bias AI-resultaterne for eleverne? Hvilke skridt kan skoler tage for at reducere bias?
- Hvorfor er det afgørende, at AI-beslutninger i undervisningen er gennemsigtige? Hvordan kan skoler gøre AI mere forståelig?
- Bør rigtige/menneskelige undervisere altid have det sidste ord i AI-drevne beslutninger? Hvad er risikoen ved fuld automatisering?



BILLEDKILDE

<https://plat.ai/blog/bias-and-fairness-in-ai-algorithms/>

Casestudie

Tips til undervisere

AI har i stigende grad gennemsyret uddannelsessektoren, fra personaliserede læringsplatforme til optagelsessystemer og elevvurderinger. Mens AI har potentiale til at forbedre uddannelsesresultaterne, er der voksende bekymring for retfærdighed og bias, især i forhold til, hvordan disse teknologier påvirker unge og studerende. Dette casestudie vil undersøge et scenarie, hvor et AI-drevet system i en uddannelsessammenhæng viste sig at være biased, konsekvenserne af denne bias og de skridt, der blev taget for at løse problemerne.

1

Vigtige spørgsmål

- **Data-bias:** AI-systemet blev trænet på historiske data, der afspejlede eksisterende bias i uddannelsessystemet, såsom overrepræsentationen af visse demografiske grupper i specialklasser.
- **Manglende gennemsigtighed:** Lærere og administratorer kunne ikke forstå eller udfordre AI'ens anbefalinger, fordi beslutningsprocessen var uigennemsigtig.
- **Konsekvenser for de studerende:** Studerende eller elever, der fejlagtigt blev vurderet som underpræsterende, blev stigmatiseret og gik glip af muligheder for at udvikle sig, hvilket påvirkede deres selvtillid og fremtidsudsigter.
- **Ethiske bekymringer:** Brugen af biased AI-systemer rejste etiske spørgsmål om retfærdighed, lige muligheder og uddannelsesinstitutionernes ansvar.

2

Handlinger

Da bias var identificeret, tog skoledistriktet følgende skridt:

- **Revision og omskoling af modellen:** AI-systemet blev revideret for at identificere og fjerne forudindtagne data og omskole modellen med et mere varieret datasæt, der bedre afspejlede den mangfoldige elevpopulation.
- **Menneskeligt tilsyn:** Rigtige/menneskelige undervisere blev geninddraget i beslutningsprocessen, hvor AI-systemet gav anbefalinger snarere end afgørelser. Det gav lærerne mulighed for at bruge deres dømmekraft og viden om eleverne til at træffe de endelige beslutninger.
- **Gennemsigtighed og forklarlighed:** Skoledistriktet arbejdede sammen med AI-eksperter for at forbedre gennemsigtigheden i AI-systemet og sikre, at lærere og administratorer kunne forstå, hvordan beslutninger blev truffet, og give elever og forældre forklaringer på vurderinger og anbefalinger.
- **Løbende overvågning:** Der blev etableret et system til løbende overvågning af AI'ens resultater for at sikre, at der ikke igen opstod bias, og at systemet fortsat behandlede alle elever retfærdigt.

Konklusion

Det omskolede AI-system kombineret med menneskeligt tilsyn reducerede antallet af studerende, der blev uretfærdigt kategoriseret, betydeligt. Med tiden førte det til mere retfærdige uddannelsesresultater, hvor flere elever fik mulighed for at udmærke sig. Sagen udløste også en bredere diskussion i skoledistriktet om ansvarlig brug af AI i undervisningen og vigtigheden af retfærdighed og gennemsigtighed i algoritmisk beslutningstagning.

Casestudie

Lærervejledning til undervisning i casestudiet

Målsætning

Denne session har til formål at hjælpe deltagerne med at forstå det grundlæggende i privatlivets fred og bekvemmelighed, især i dagligdags aktiviteter som smart home-enheder, mobilapps og online shopping. Der kan indgå interaktive diskussioner, refleksioner og aktiviteter for at styrke læringen.

Generelle tips til undervisere

- **Engager deltagerne:** Hold sessionen interaktiv og tilskynd deltagerne til deltagelse.
- **Relater til det virkelige liv:** Brug eksempler og scenarier fra det virkelige liv, som deltagerne kan relatere til.
- **Vær fleksibel:** Juster sessionen ud fra deltagernes svar og engagement.
- **Tilskynd deltagerne til kritisk tænkning:** Få deltagerne til at tænke kritisk over kompromiserne mellem privatlivets fred og bekvemmelighed.
- **Hold dig opdateret:** Hold dig orienteret om de seneste tendenser og problemer inden for privatlivets fred og teknologi (seneste hændelse(r)).

Underviservejledning til undervisning i casestudie

1

Forberedelse

- Nødvendige materialer:
 - Præsentationsslides med casestudiet
 - Handouts med scenarier fra casestudiet
 - Google Forms-link med spørgsmål til aktivitet/evaluering
- Setup:
 - Indret rummets siddepladser hensigtsmæssigt for at lette gruppediskussioner.
 - Del linket til Google Forms med deltagerne før/under sessionen.

2

Oversigt over sessionen

- Forklar sessionens formål: at undersøge, hvordan teknologi påvirker vores privatliv, og hvordan vi kan afbalancere bekvemmelighed med beskyttelse af privatlivets fred.
- Icebreaker: Bed deltagerne om at fortælle om en teknologi, de bruger hver dag, og hvordan den gør deres liv lettere.
- Præsenter casestudiet, og diskuter bekvemmeligheden og problemerne med privatlivets fred.
- Forklar Johns oplevelse med personlig shopping og målrettede annoncer og de bekymringer om privatlivets fred, det gav anledning til.
- Del deltagerne op i små grupper på 3-4 personer, og bed hver gruppe om at identificere de største problemer med privatlivets fred i det tildelte casestudie.
 - Få dem til at diskutere potentielle løsninger eller handlinger, der kan mindske risici for privatlivets fred.
 - Opfordr dem til at tænke over balancen mellem bekvemmelighed og privatliv.
- Del linket til Google Forms med deltagerne.
 - Få deltagerne til at udfylde spørgsmålene individuelt eller i par.
 - Spørgsmålene vil omfatte udfyldning af tomme felter, multiple choice, flere korrekte svar og opfordringer til selvrefleksion.
- Scenariebaserede diskussioner:
 - Præsenter korte scenarier, og bed eleverne om at reflektere over, hvad de ville gøre.
 - Eksempler: "Hvad ville du gøre, hvis du fandt ud af, at din smart-enhed lyttede til private samtaler?" eller "Hvordan ville du håndtere en meddelelse om databrud fra en fitness-app, du bruger?"

3

Refleksion og konklusion

- **Selvrefleksion**

- Bed deltagerne om at skrive et kort afsnit om, hvordan de kan balancere privatliv og bekvemmelighed i deres hverdag.

- **Vigtige pointer**

- Opsummer de vigtigste punkter, der blev diskuteret i løbet af sessionen.
- Understreg vigtigheden af at beskytte privatlivets fred, mens man nyder godt af teknologiske bekvemmeligheder.

TAK

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



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

