



KURS 4 · VURDERE

Skjeve data, skjeve svar

Hvem rammes når data er skjeve?

Lærerveiledning

9. trinn (forslag) · tre økter på 60 minutter · med fasit



KURS 4: SKJEVE DATA, SKJEVE SVAR**Innhold**

Om kurset	3
Økt 1: Hvem er med i dataene?	4
Økt 2: Hva lærte søknadssystemet?	7
Økt 3: Egne rettferdighetsregler	10
Vedlegg: Fasit	13
Kilder og kreditering	13

KURS 4

Om kurset

Hvem rammes når data er skjeve?

Kurset har tre økter. Hver økt er planlagt for rundt 40 minutter i en time på 60, og har fire deler: fakta, praktisk arbeid, diskusjon og skriving. Elevene starter med korte notater om hva de tror, og ser på dem igjen til slutt. Hver del er merket som kjerne, utdyping eller kan flyttes. Tidene er omtrentlige.

KJERNE

UTDYPING

KAN FLYTTES

Øktene

Økt	Tittel og mål	KI-bruk
Økt 1	Hvem er med i dataene? Elevene kan forklare hvorfor et system kan fungere dårligere for grupper som er lite med i treningsdataene, og viser at et gjennomsnitt kan skjule store forskjeller.	Uten KI
Økt 2	Hva lærte søknadssystemet? Elevene kan finne mønstre et system ville lært av tidligere beslutninger, forklare at egenskaper kan stå i stedet for noe annet, og vurdere hva som ikke bør telle.	Uten KI
Økt 3	Egne rettferdighetsregler Elevene kan skrive regler for et system som deler ut noe, se at ulike regler gir ulike vinnere, og forklare årsakene til skjeve resultater.	Uten KI

Materiell til utskrift

Materiell	Fil	Papir og antall
Datakort	datakort.pdf	A4, tre sider med 20 kort, ett sett per gruppe
Regelark	regelark.pdf	A4, ett per gruppe
Søkerkort	soekerkort.pdf	A4, 12 søkere og 4 nye søkere (holdes tilbake), ett sett per gruppe
Stipendkort	stipendkort.pdf	A4, seks kort, ett sett per gruppe
Stolpeark	stolpeark.pdf	A4, ett per gruppe
Testresultater	testresultater.pdf	A4, ett per gruppe
Vurderingsark	vurderingsark.pdf	A4, ett per gruppe
KI-logg	ki-logg.pdf	A4, ett ark per elev per økt

Materiellet og presentasjonene lastes ned fra kurssiden på nettsiden. Presentasjonene åpnes i nettleseren, og PDF av hver presentasjon ligger der som reserve. Fasiten ligger som vedlegg bakerst i heftet. Den er for læreren og skal ikke deles ut til elevene.

KURS 4 · ØKT 1

Hvem er med i dataene?

9. trinn

40 av 60 min

Uten KI

Mål for økta: elevene kan forklare hvorfor et system kan fungere dårligere for grupper som er lite med i treningsdataene, og viser at et gjennomsnitt kan skjule store forskjeller.

Kort fortalt: elevene får høre om forskning som viser at ansiktsgjenkjenning feiler oftere for noen grupper. Så teller de 60 datakort fra en oppdiktet ansiktslås til elevskap, tegner stolper og regner ut hvor ofte systemet feiler for hver gruppe. Elevene skal gå ut av timen med at grupper som er lite med i dataene, rammes oftest, og at et samlet tall på 82 prosent riktige skjuler det.

Før timen

rundt 25 min

- Skriv ut datakortene, ett sett per gruppe på tre eller fire: [datakort.pdf](#) (tre sider, 60 kort). Klipp ut kortene og legg hvert sett i en konvolutt.
- Skriv ut stolpearket, ett per gruppe: [stolpeark.pdf](#).
- Skriv ut testresultatene, ett per gruppe: [testresultater.pdf](#). Arket deles ut først i trinn 3.
- Elevene trenger KI-loggen sin ([ki-logg.pdf](#) hvis noen trenger nytt ark).
- Les «Det du trenger å vite» under, og fasiten i fasiten for kurset (egen PDF på øktsiden).
- Åpne presentasjonen: [økt-1.html](#). Trykk N for å se lærernotatene under lysbildet. PDF-en er reserve hvis nettleseren ikke virker.

Det du trenger å vite

Hovedideen: et system som lærer av data, blir best på det det har sett mest av. Er noen grupper lite med i treningsdataene, får systemet lite øvelse på dem, og feiler oftere for dem.

Forskningen elevene får høre om:

- **Gender Shades (2018):** to forskere testet tre kommersielle systemer som skulle avgjøre om et ansikt tilhørte en kvinne eller en mann. For menn med lys hud var feilprosenten høyst 0,8. For kvinner med mørk hud var den opptil 34,7. To mye brukte bildesamlinger hadde 79,6 og 86,2 prosent personer med lys hud.
- **NIST (2019 og 2025):** det amerikanske standardinstituttet testet 189 algoritmer fra 99 utviklere. De fleste feilet ulikt for ulike grupper, for noen typer feil mellom 10 og over 100 ganger så ofte. De mest nøyaktige algoritmene hadde minst forskjell. Underrepresentasjon i treningsdata kan være én årsak. Dårlig lys og bildekvalitet er en annen.

Øvelsen: 60 datakort, hvert merket med hudtone og kjønn. Kortene har ingen bilder. Tallene er oppdiktet, men mønsteret er det forskerne har funnet.

Gruppe	Hudtone, kjønn	Antall kort	Feil av 12 forsøk	Feilprosent
A	lys, gutt	22	0	0
B	lys, jente	20	1	8
C	middels, gutt	6	1	8
D	middels, jente	6	3	25
E	mørk, gutt	4	3	25
F	mørk, jente	2	5	42

Til sammen 59 av 72 riktige, altså 82 prosent. Det høres bra ut, men for gruppe F feiler systemet nesten annenhver gang.

Viktig: elevene skal ikke telle eller plassere seg selv eller hverandre i gruppene. Si det høyt før trinn 1.

I kurs 4 kommer faktadelen før det praktiske, fordi kurset er på 9. trinn.

Slik går timen

Del	Lysbilde	Tid
Start: hva tror du?	1 og 2	3 til 5 min
Fakta: forskningen og gjennomsnittet	3 til 6	6 til 10 min
Praktisk: sorter, tegn, les testresultater	7	14 til 18 min
Diskusjon	8	6 til 10 min
KI-logg	9	3 til 5 min

Lysbilde 10 er kilder, og trenger ikke vises i timen.

Start: hva tror du?

3 til 5 min **KJERNE**

Lysbilde 1 og 2. Del ut en lapp til hver elev.

Si: «Hvis et system skal kjenne igjen ansikter, hva må være med i bildene det lærer av, for at det skal virke for alle? Skriv det på lappen.» Lappen legges i konvolutten.

1. Fakta og info

6 til 10 min **KJERNE**

Lysbilde 3. Si: «Et system som lærer av data, vet bare det som finnes i eksemplene. Er noen lite med i dataene, får systemet lite øvelse på dem.»

Lysbilde 4. Si: «I 2018 testet to forskere tre systemer som skulle avgjøre om et ansikt tilhørte en kvinne eller en mann. For menn med lys hud tok systemene feil i under én prosent av tilfellene. For kvinner med mørk hud tok de feil i opptil 35 prosent. Bildesamlingene mange brukte, hadde rundt 80 prosent personer med lys hud.»

Lysbilde 5. Si: «Det amerikanske standardinstituttet NIST testet 189 algoritmer. De fleste feilet oftere for noen grupper enn for andre. De beste algoritmene hadde minst forskjell. En mulig årsak er at noen grupper er lite med i treningsdataene. Dårlig lys kan også gi feil.»

Lysbilde 6. Si: «Et system kan ha 82 av 100 riktige til sammen, og likevel feile nesten annenhver gang for én gruppe. Det skal dere se selv nå.»

2. Praktisk

14 til 18 min **KJERNE**

Lysbilde 7. Grupper på tre eller fire. Del ut datakortene og stolpearket.

Si først: «Dere får 60 datakort fra en oppdiktet ansiktslås til elevskap. Tallene er oppdiktet, men mønsteret er det forskerne har funnet. Kortene har ingen bilder. Ingen skal telle eller plassere seg selv eller andre i gruppene.»

Trinn 1: sorter og tell (5 til 6 min). **Si:** «Sorter kortene i seks bunker etter det som står på kortet. Skriv antallet i hver bunke øverst på stolpearket.»

Trinn 2: tegn stolpene (3 til 4 min). **Si:** «Fargelegg én rute per kort, så dere får seks stolper. Svar på spørsmålet under: hvilken stolpe er høyest, hvilken er lavest, og hva tror dere det betyr for systemet?»

Trinn 3: les testresultatene (6 til 8 min). Del ut testresultatene.

Si: «Hver rad er 12 forsøk der systemet skulle åpne et skap. Et kryss betyr at systemet ikke kjente igjen eleven. Tell kryssene i hver rad, regn ut prosent ved å dele på 12, og skriv det under stolpen for samme gruppe. Til slutt: hvor mange av alle 72 forsøkene gikk riktig?»

Se etter, og spør:

- Grupper som regner prosent ut fra 60 i stedet for 12. Pek på at hver rad har 12 forsøk.
- Grupper som sier at systemet «bare er dårlig på mørk hud». Be dem se på gruppene A til F hver for seg. (D, middels hud og jente, har like mange feil som E.)
- Grupper som ser sammenhengen mellom lav stolpe og høy feilprosent. Spør: «Hvorfor tror dere det er slik?»

3. Diskusjon

6 til 10 min

Lysbilde 8. Still ett spørsmål om gangen. La elevene tenke alene i et halvt minutt, snakke med den som sitter ved siden av i ett minutt, og si det til hele klassen etterpå.

- **Kjernespørsmål:** «Hvem merker feilen først, og hvem får vite om den?» (Eleven som ikke får åpnet skapet merker det med en gang. De som laget systemet, får kanskje aldri vite det.)
- **Utdyping:** «82 prosent riktige til sammen. Hvorfor kan det være en dårlig målestokk alene?» (Det skjuler at gruppe F feiler 42 prosent av gangene.)
- **Utdyping:** «Hvordan kunne de som laget systemet oppdaget feilen før det ble brukt?» (Telle hvem som var med i dataene, og teste hver gruppe for seg.)
- **Utdyping:** «Hvem har ansvaret hvis eleven i gruppe F ikke får åpnet skapet: de som samlet bildene, de som laget systemet, eller skolen som tok det i bruk?» (Ikke ett riktig svar. Kurs 9 handler om dette.)

4. Skrivning og refleksjon

3 til 5 min **KAN FLYTTES**

Lysbilde 9. Elevene ser på lappen fra starten. Så skriver de i KI-loggen:

- Dagens spørsmål: «Hva måtte vært annerledes i dataene for at systemet skulle virke for alle?» (Godt svar: flere kort fra gruppene som var lite med, og test for hver gruppe.)
- «Nå tenker jeg ...»
- «Jeg lurte på ...»

Har dere dårlig tid, flytter du denne delen til starten av økt 2.

Misforståelser å følge med på

- At skjeve resultater alltid skyldes at noen har laget systemet med vilje. Her skyldes det hvem som var med i dataene.
- At mer data løser det. Det hjelper bare hvis dataene inneholder de som mangler.
- At et høyt gjennomsnitt betyr at systemet virker for alle.
- At forskjellene bare gjelder ansikter. Det samme kan skje med stemmer, dialekter og språk.

Tilpasning

- **Kort tid:** hopp over trinn 2. Gruppene skriver antallene på stolpearket og går rett til testresultatene.
- **Mer tid:** gruppene foreslår hvor mange kort av hver gruppe de ville hatt i et nytt datasett på 60, og begrunner.
- **Knytt til dialekter:** kortet «Tale til tekst» i kurs 1 kan hentes fram. Spør hvilke dialekter som er lite med i et norsk datasett.

KURS 4 · ØKT 2

Hva lærte søknadssystemet?

9. trinn

40 av 60 min

Uten KI

Mål for økta: elevene kan finne mønstre et system ville lært av tidligere beslutninger, forklare at egenskaper kan stå i stedet for noe annet, og vurdere hva som ikke bør telle.

Kort fortalt: elevene får tolv oppdiktete jobbsøkere og ser hvem et firma tidligere kalte inn til intervju. De finner mønsteret et system ville lært: unge søkere uten hull i CV-en og uten kvinnelag kom videre, mens erfaring ikke betydde noe. Elevene skal gå ut av timen med at et system som lærer av tidligere beslutninger, også lærer urettferdigheten i dem, og at kjønn og alder kan snike seg inn gjennom andre opplysninger.

Før timen

rundt 20 min

- Skriv ut søkerkortene, ett sett per gruppe på tre eller fire: [søkerkort.pdf](#) . Side 1 er 12 søkere med tidligere beslutning, side 2 er fire nye søkere. Klipp ut kortene.
- Legg side 2 (nye søkere) til side. Den deles ut først i trinn 3.
- Skriv ut vurderingsarket, ett per gruppe: [vurderingsark.pdf](#) .
- Elevene trenger KI-loggen sin ([ki-logg.pdf](#) hvis noen trenger nytt ark).
- Les «Det du trenger å vite» under, og fasiten i fasiten for kurset (egen PDF på øktsiden).
- Åpne presentasjonen: [okt-2.html](#) . Trykk N for å se lærernotatene under lysbildet. PDF-en er reserve hvis nettleseren ikke virker.

Ikke si til elevene hva mønsteret er. Det skal de finne selv.

Det du trenger å vite

Hovedideen: et system som lærer av tidligere beslutninger, finner det som skiller dem som fikk ja, fra dem som fikk nei. Det spør ikke om beslutningene var rettferdige.

Stedfortredere: systemet trenger ikke å få vite alder eller kjønn. Andre opplysninger kan stå i stedet: hull i CV-en (ofte omsorgspermisjon), hvilket lag man spiller på, årstallet man tok eksamen. Systemet finner dem likevel.

Mønsteret i søkerkortene: alle som fikk intervju, er under 35 år, har ikke hull i CV-en og har ikke kvinnelag som interesse. Erfaring og utdanning skiller ikke bunkene. Hull i CV og kvinnelag står i stedet for kjønn. Alder står på kortet og teller direkte.

Ny søker	Systemet ville svart	Hvorfor
N1	Intervju	Ung, ingen hull, bare 1 års erfaring
N2	Ikke intervju	49 år. 10 års erfaring teller ikke
N3	Ikke intervju	Hull i CV-en (omsorgspermisjon)
N4	Ikke intervju	Kvinnelag. 5 års erfaring teller ikke

Workday-saken: i mai 2025 lot en dommer i California et gruppesøksmål mot selskapet Workday gå videre. Workday lager et system mange firmaer bruker til å sortere jobbsøkere. Saksøkeren mener at søkere over 40 år vurderes dårligere. Workday avviser det. Saken er ikke avgjort, og ingen domstol har sagt at systemet er skjevt. Dommeren sa bare at saken kunne gå videre. Vær nøye med dette hvis elevene spør.

I kurs 4 kommer faktadelen før det praktiske, fordi kurset er på 9. trinn.

Slik går timen

Del	Lysbilde	Tid
Start: hva tror du?	1 og 2	3 til 5 min
Fakta: tidligere beslutninger, stedfortredere, Workday	3 til 5	6 til 10 min

Del	Lysbilde	Tid
Praktisk: finn mønsteret, bør det telle, nye søkere	6	14 til 18 min
Diskusjon	7	6 til 10 min
KI-logg	8	3 til 5 min

Lysbilde 9 er kilder, og trenger ikke vises i timen.

Start: hva tror du?

3 til 5 min **KJERNE**

Lysbilde 1 og 2. Ble KI-loggen fra økt 1 flyttet, gjør elevene den nå.

Ellers deler du ut en lapp til hver elev. **Si:** «Et firma lar en datamaskin velge hvem som får komme på intervju. Hva bør maskinen se på, og hva bør den ikke se på? Skriv det på lappen.» Lappen legges i konvolutten.

1. Fakta og info

6 til 10 min **KJERNE**

Lysbilde 3. Si: «Et firma har hatt mange søkere. Noen kom på intervju, andre ikke. Et system som lærer av dette, finner det som skiller de to gruppene. Det spør ikke om beslutningene var rettferdige.»

Lysbilde 4. Si: «Systemet trenger ikke å få vite alder eller kjønn. Andre opplysninger kan stå i stedet, for eksempel hull i CV-en, hvilket lag man spiller på, eller året man tok eksamen. Det kaller vi stedfortredere.»

Lysbilde 5. Si: «Dette er ikke bare en tanke. I 2025 lot en dommer i USA et søksmål mot selskapet Workday gå videre. Saksøkeren mener at systemet deres vurderer søkere over 40 år dårligere. Workday sier at det ikke stemmer. Saken er ikke avgjort.»

2. Praktisk

14 til 18 min **KJERNE**

Lysbilde 6. Grupper på tre eller fire. Del ut de tolv søkerkortene og vurderingsarket.

Trinn 1: finn mønsteret (5 til 7 min). Si: «Søkerne er oppdiktet. Legg de tolv kortene i to bunker: intervju og ikke intervju. Let etter egenskaper som skiller bunkene, og skriv ned alle dere finner.»

Trinn 2: bør det telle? (4 til 5 min). Si: «For hver egenskap dere fant: sett kryss for bør telle, bør ikke telle eller usikker. Skriv én setning om hvorfor.»

Trinn 3: la systemet gjette (5 til 6 min). Del ut de fire nye søkerkortene.

Si: «Hva ville et system som har lært av de tolv, svart for hver av disse fire? Skriv det. Skriv så hva en rettferdig vurdering ville sagt.»

Se etter, og spør:

- Grupper som bare ser på erfaring og utdanning. Be dem sammenligne S12 (1 års erfaring, intervju) med S10 (12 års erfaring, ikke intervju).
- Grupper som oppdager at kvinnelag følger beslutningene. Spør: «Kjønn står ikke på kortet. Hvordan kan systemet likevel finne det?»
- Grupper som mener alle uvanlige trekk bør telle. Spør: «Hva med hull som skyldes omsorgspermisjon?»

3. Diskusjon

6 til 10 min

Lysbilde 7. Still ett spørsmål om gangen. La elevene tenke alene i et halvt minutt, snakke med den som sitter ved siden av i ett minutt, og si det til hele klassen etterpå.

- **Kjernespørsmål:** «Hva er rettferdig, og for hvem?» (Ikke ett riktig svar. Få elevene til å begrunne.)

- **Utdyping:** «Systemet fikk aldri vite kjønn. Er det likevel urettferdig mot kvinner? Hvorfor?» (Ja. Kvinnelag og hull i CV-en sto i stedet for kjønn.)
- **Utdyping:** «Hvem tjener på at et firma bruker et slikt system, og hvem taper?» (Firmaet sparer tid. Eldre søkere og kvinner taper.)
- **Utdyping:** «Hva ville dere gjort med dataene for å rette det? Er det nok å fjerne hull i CV-en?» (Nei. Kvinnelag og alder ville fortsatt telt.)
- **Utdyping:** «Bør et firma si fra til søkerne at en maskin sorterer?»

4. Skrivning og refleksjon

3 til 5 min

KAN FLYTTES

Lysbilde 8. Elevene ser på lappen fra starten. Så skriver de i KI-loggen:

- Dagens spørsmål: «Hva ville du endret i dataene?»
- «Nå tenker jeg ...»
- «Jeg lurar på ...»

Har dere dårlig tid, flytter du denne delen til starten av økt 3.

Misforståelser å følge med på

- At det holder å fjerne kjønn og alder fra dataene. Stedfortredere kan bringe dem tilbake.
- At et system som lærer av tidligere beslutninger, er objektivt fordi det er en maskin.
- At mennesker alltid er mer rettferdige enn maskiner. Her lærte maskinen av menneskers beslutninger.
- At Workday-saken beviser at systemet er skjevt. Den er ikke avgjort.

Tilpasning

- **Kort tid:** hopp over trinn 2 og la gruppene gå rett til de nye søkerne.
- **Mer tid:** gruppene lager et nytt tolvkorts datasett som er rettferdig, og tester om mønsteret forsvinner.

KURS 4 · ØKT 3

Egne rettferdighetsregler

9. trinn

40 av 60 min

Uten KI

Mål for økta: elevene kan skrive regler for et system som deler ut noe, se at ulike regler gir ulike vinnere, og forklare årsakene til skjeve resultater.

Kort fortalt: skolen skal dele ut to stipend til seks oppdiktete søkere. Gruppene velger først selv, skriver så tre regler et system kunne fulgt, og ser at hver regel gir nye vinnere. Elevene skal gå ut av timen med at en maskin ikke kan avgjøre hva som er rettferdig. Det må mennesker bestemme, og hver regel har noen som taper.

Før timen

rundt 15 min

- Skriv ut stipendkortene, ett sett per gruppe på tre eller fire: [stipendkort.pdf](#) (seks søkere). Klipp ut kortene.
- Skriv ut regelarket, ett per gruppe: [regelark.pdf](#).
- Finn fram KI-kubene fra kurs 1 hvis gruppene har dem (valgfritt).
- Elevene trenger KI-loggen sin ([ki-logg.pdf](#) hvis noen trenger nytt ark).
- Les «Det du trenger å vite» under, og fasiten i fasiten for kurset (egen PDF på øktsiden).
- Åpne presentasjonen: [okt-3.html](#). Trykk N for å se lærernotatene under lysbildet. PDF-en er reserve hvis nettleseren ikke virker.

Søkerne er oppdiktet, og elevene oppgir ingen egne opplysninger.

Det du trenger å vite

Fem veier til skjeve svar (oppsummering av kurset):

1. Hvem som er med i dataene (økt 1, ansiktslåsen).
2. Hva tidligere beslutninger sier (økt 2, søknadssystemet).
3. Stedfortredere for noe som ikke skulle telle (økt 2, kvinnelag og hull i CV-en).
4. Hva systemet er bedt om å gjøre, for eksempel «velg dem som ligner på dem vi har hatt før».
5. Hvor systemet brukes. Et system som virker i én gruppe eller ett land, kan virke dårlig et annet sted.

De seks søkerne:

Søker	Snitt	Familieinntekt	Bosted	Frivillig innsats per uke
K1	5,6	høy	by	2 timer
K2	5,2	lav	distrikt	0
K3	4,8	lav	by	6 timer
K4	5,0	middels	distrikt	4 timer
K5	4,3	lav	distrikt	8 timer
K6	5,5	middels	by	1 time

Hva ulike regler gir:

Regel	Gir stipend til
Høyest snitt	K1 og K6
Lav familieinntekt, så høyest snitt	K2 og K3
Mest frivillig innsats	K5 og K3
Lotteri blant dem med snitt 4,5 eller mer	To av K1, K2, K3, K4 og K6. K5 er ute

K3 er med i to regler. K5 er bare med i innsatsregelen. Det finnes ingen regel som er rettferdig for alle seks. Valget gruppene gjør selv i trinn 1, er sjelden likt noen av reglene.

I kurs 4 kommer faktadelen før det praktiske, fordi kurset er på 9. trinn.

Slik går timen

Del	Lysbilde	Tid
Start: hva tror du?	1 og 2	3 til 5 min
Fakta: fem veier, tre regler	3 til 5	6 til 10 min
Praktisk: velg, skriv regler, kjør reglene	6	14 til 18 min
Diskusjon	7	6 til 10 min
KI-logg	8	3 til 5 min

Lysbilde 9 er kilder, og trenger ikke vises i timen.

Start: hva tror du?

3 til 5 min **KJERNE**

Lysbilde 1 og 2. Ble KI-loggen fra økt 2 flyttet, gjør elevene den nå.

Ellers deler du ut en lapp til hver elev. **Si:** «Skolen deler ut to stipend. Hvem bør få dem? Skriv det på lappen.» Lappen legges i konvolutt. Ikke si noe om reglene ennå.

1. Fakta og info

6 til 10 min **KJERNE**

Lysbilde 3. Si: «Vi har sett flere veier til skjeve svar. Hvem som er med i dataene, som ansiktslåsen. Hva tidligere beslutninger sier, som søknadssystemet. Stedfortredere, som kvinnelaget. I tillegg kan det handle om hva systemet er bedt om å gjøre, og hvor det brukes.»

Lysbilde 4. Si: «En maskin kan ikke avgjøre hva som er rettferdig. Noen må skrive regelen.»

Lysbilde 5. Si: «Tre vanlige regler er å gi til dem som har best resultat, til dem som trenger det mest, eller å gi alle som oppfyller et minstekrav lik sjanse. Reglene gir ofte ulike vinnere.»

2. Praktisk

14 til 18 min **KJERNE**

Lysbilde 6. Grupper på tre eller fire. Del ut stipendkortene og regelarket.

Trinn 1: velg selv (3 til 4 min). **Si:** «Les de seks kortene. Bli enige om to som skal få stipend, og skriv hvorfor.»

Trinn 2: skriv tre regler (5 til 6 min). **Si:** «Skriv tre regler et system kunne fulgt for å velge to av seks. Regelen må kunne brukes uten å spørre et menneske. For eksempel: den med høyest snitt.»

Trinn 3: kjør reglene (5 til 6 min). **Si:** «Bruk hver regel på de seks kortene, og skriv hvem som får stipend. Sammenlign med det dere valgte selv.»

Valgfritt (3 min): gruppen skriver en lapp til KI-kuben fra kurs 1: «Hva er rettferdig her, og for hvem?» og limer den på kubens.

Se etter, og spør:

- Regler som bare passer for én av de seks. Be gruppen skrive en regel som også kunne brukes på seks andre søkere.
- Regler som bruker bosted eller inntekt uten å si hvorfor. Spør: «Hvorfor skal det telle?»
- Grupper som er uenige. La dem skrive begge reglene.

3. Diskusjon

6 til 10 min

Lysbilde 7. Still ett spørsmål om gangen. La elevene tenke alene i et halvt minutt, snakke med den som sitter ved siden av i ett minutt, og si det til hele klassen etterpå.

- **Kjernespørsmål:** «Hvor kolliderer reglene, og hvem bestemmer hvilken som gjelder?» (Snitt og behov gir helt ulike vinnere. Noen må velge, for eksempel skolen eller de som deler ut stipendet.)
- **Utdyping:** «K5 har lavt snitt, lav inntekt og mye frivillig arbeid. Hvilken regel hjelper K5, og er det rettferdig?» (Bare innsatsregelen. I lotteriet er K5 ute.)
- **Utdyping:** «Hvis stipendet ble delt ut av et system som lærte av tidligere vinnere, hvilke regler ville det lært uten at noen skrev dem?» (Det ville kopiert det som kjennetegnet tidligere vinnere, som i økt 2.)
- **Utdyping:** «Bør en maskin kunne avgjøre dette alene? Hva bør et menneske alltid få se?»

4. Skrivning og refleksjon

3 til 5 min

KAN FLYTTES

Lysbilde 8. Elevene ser på lappen fra starten. Så skriver de i KI-loggen:

- Dagens spørsmål: «Hvilken regel ville du ha valgt, og hvem taper på den?»
- «Nå tenker jeg ...»
- «Jeg lurar på ...»

Har dere dårlig tid, flytter du denne delen til starten av neste kurs.

Misforståelser å følge med på

- At det finnes én riktig definisjon av rettferdig som bare må programmeres.
- At en regel som behandler alle likt, alltid er rettferdig.
- At skjeve data bare er et teknisk problem. Det er også et spørsmål om verdier.

Tilpasning

- **Kort tid:** hopp over trinn 1 og la gruppene skrive reglene rett.
- **Mer tid:** gruppene bytter regler og søkerkort med en annen gruppe, og finner en søker reglene behandler urimelig.

VEDLEGG

Fasit

For lærere. Ikke del fasiten med elevene før øktene er gjennomført.

Økt 1: ansiktslåsen

Antall datakort per gruppe: A (lys, gutt) 22, B (lys, jente) 20, C (middels, gutt) 6, D (middels, jente) 6, E (mørk, gutt) 4, F (mørk, jente) 2. Til sammen 60.

Gruppe	Feil av 12	Feilprosent
A	0	0
B	1	8
C	1	8
D	3	25
E	3	25
F	5	42

Til sammen 13 feil av 72, altså 59 riktige av 72 (82 prosent). Feilprosenten øker jevnt når antallet kort synker. Tallene er oppdiktet.

Økt 2: søkerkortene

Systemet ville lært at alle som fikk intervju, er under 35 år, har ikke hull i CV og har ikke kvinnelag som interesse. Erfaring og utdanning skiller ikke bunkene.

Nytt kort	Systemet ville svart	Hvorfor
N1	Intervju	Ung, ingen hull, 1 års erfaring
N2	Ikke intervju	49 år, og 10 års erfaring teller ikke
N3	Ikke intervju	Hull i CV (omsorgspermisjon)
N4	Ikke intervju	Kvinnelag, og 5 års erfaring teller ikke

Hull i CV og kvinnelag står i stedet for kjønn. Alder står på kortet og teller direkte.

Økt 3: stipendkortene

Regel	Gir stipend til
Høyest snitt	K1 (5,6) og K6 (5,5)
Lav familieinntekt, så høyest snitt	K2 (5,2) og K3 (4,8)
Mest frivillig innsats	K5 (8 timer) og K3 (6 timer)
Lotteri blant dem med snitt 4,5 eller mer	Blant K1, K2, K3, K4 og K6. K5 er ute

K3 er med i to regler. K5 er med bare i innsatsregelen og ute av lotteriet. Det finnes ingen regel som er rettferdig for alle seks.

Kilder og kreditering

Alle kildene materielt bygger på, står under Kilder på nettsiden. Kildene sjekkes før trykk, og deretter hvert år.

Økt 1: Hvem er med i dataene?

- Buolamwini og Gebru, «Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification», Proceedings of Machine Learning Research 81 (2018).
- Grother, Ngan og Hanaoka, NISTIR 8280, «Face Recognition Vendor Test Part 3: Demographic Effects» (2019), og NISTs «FRTE Demographics» (tabeller oppdatert 5. mars 2025).
- Tallene i øvelsen er oppdiktet.

Økt 2: Hva lærte søknadssystemet?

- Workday-saken: Mobley v. Workday, Inc., US District Court, Northern District of California, sak 3:23-cv-00770. Om vedtaket i mai 2025: Fisher Phillips og Davis Wright Tremaine (begge 2025).
- Søkerne og tallene er oppdiktet.

Økt 3: Egne rettferdighetsregler

- De fem veiene til skjeve svar bygger på det som er vist i økt 1 og 2 og kildene der.
- Søkerne er oppdiktet.