

Forberedelser til muntlig eksamen - SOK3023

Fagansvarlig: Markus J. Aase
Universitetslektor i matematikk og statistikk

March 10, 2025

Dette dokumentet inneholder mulige spørsmål studenter i SOK3023 *Maskinlæring for Økonomer* **kan** få på muntlig eksamen. Spørsmålene er hentet fra pensum og dekker tematikk vi har diskutert i forelesningene, kompendiet eller som dere har jobbet med i prosjektarbeidet. Alle kandidater **vil** få noen spørsmål fra listen nedenfor.

Spørsmål

1. Hva er forskjellen på **regresjon** og **klassifikasjon**?
2. Hva er en **tapsfunksjon**? Gi to eksempler, og forklar hva de gjør? Hvorfor er disse nyttige i maskinlæring, og når er de nyttig?
3. Hvordan fungerer **beslutningstrær**, og hvorfor er disse så nyttige?
4. Hvorfor er konseptet derivasjon viktig i nevrale nettverk?
5. Forklar strukturen til et **nevralt nettverk**.
6. Forklar hva **batch**, **batch-størrelse** og **epochs** er, og hvordan disse relaterer til trening av nevrale nettverk.
7. Hva er **konvolusjonelle nevrale nettverk** (CNN), og hva skiller dem fra mer tradisjonelle nevrale nettverk?
8. Hva er **mean square error** (MSE) og **mean absolute error** (MAE), og hva kan de brukes til? Fordeler/ulempene ved disse?
9. I kurset har vi diskutert **Whisper-modellen**, hva er det og hva kan den brukes til? Kan du gi noen bruksområder det kan være nyttig?
10. Hva bruker vi **valideringssett** til, i maskinlæring?
11. Hva er en **loss-funksjon**? Kan du gi noen eksempler?
12. Hva er **beslutningstrær**? Hvordan *lages* dem? Kan du si noe om fordelene ved beslutningstrær?

13. Hva er **random forests**, og hva kan disse brukes til? Hva er fordelene med random forests sammenlignet med et nevralt nettverk?
14. Hva er **gradient descent**? Forklar og illustrer hva det gjør?
15. Hva er **lineær regresjon**? Når egner det seg?
16. Hva er **logistisk regresjon**? Når egner det seg?
17. Hva er **bias-variance trade-off**?
18. I kurset har vi diskutert **reducible** og **irreducible error**, ved å se på forventningsverdien av den kvadrerte forskjellen mellom Y og \hat{Y} , hva handler dette om?
19. Hva er **kryssvalidering**, og hva kan det brukes til?
20. Kan du gi noen eksempler hvor maskinlæring kan brukes innenfor økonomi og/eller samfunnsøkonomi?
21. I kurset har vi diskutert hva **læring** betyr i maskinlæring. Kan du snakke litt om det?
22. Hvordan relaterer L2 normen til MSE?
23. Hva er hensikten med **backpropagation** (norsk: **tilbakepropagering**)?
24. Hvilke metoder ville du brukt for å evaluere en klassifikasjonsmodell? Hva med en regresjonsmodell?
25. Hva er **regularisering** i maskinlæring? Vi har diskutert noen forskjellige metoder, kan du forklare noen av dem?
26. Hvorfor benytter vi oss av **max pooling** i konvolusjonelle nevralt nettverk? Hva med **konvolusjon**?
27. En binær klassifikasjonsmodell har følgende **confusion matrix**:

$$\begin{bmatrix} 980 & 10 \\ 8 & 2 \end{bmatrix}$$

Der:

- 980 er sanne negative (TN)
- 10 er falske positive (FP)
- 8 er falske negative (FN)
- 2 er sanne positive (TP)

Hva sier denne matrisen oss om modellen?

28. Hva er forskjellen på **veiledet læring** (supervised) og **ikke-veiledet læring**?
29. Forklar begrepene **accuracy**, **precision**, **recall** og **F1-score**. Hvordan kan disse brukes til å evaluere en klassifikasjonsmodell?
30. Hva er fordelene med **LSTM**, og hvorfor er disse nyttige for sekvensielle data? Har du et eksempel på en økonomisk anvendelse?
31. Hvilke problemer kan man møte på med dyplæringsteknikker som nevrale nettverk?
32. Du har et klassifikasjonsproblem, og ønsker å bruke et nevralt nettverk, hva er viktig at du tenker på med hensyn til utgangslaget av nettverket?
33. I forelesningene har vi brukt det klassiske eksempelet om håndskrevne tall, hvordan kan man klassifisere håndskrevne tall med nevrale nettverk? Forklar.
34. I kurset har vi diskutert **tensorer**, hva kan det brukes til? Hva er forskjellen på rank-1, rank-2 og rank-3 tensorer?
35. Hva er **overfitting**? Hvordan kan vi unngå det? Hva med **underfitting**?

Lykke til. Kan dere alle spørsmålene her, har dere en god oversikt over pensum. Likevel, **vil** dere få spørsmål som ikke står her, men som kommer fra pensum. Husk, at det også er viktig at dere har god oversikt over prosjektarbeidet deres, da dere kan bli utfordret på hvorfor dere har brukt de metodene dere har.

Muntlig eksamen varer 30 minutter, hvor dere har 10 minutter til å presentere hva dere har gjort i prosjektet deres. Resterende tid blir en dialog med sensor, hvor dere blant annet får spørsmål fra listen over.