

Er det hund eller katt?

1: BESKRIV PROBLEMET

OMFANG / SCOPE

Modellen skal kunne skille mellom hunder og katter. Det er nyttig for å definere om det er en hund eller katt på bildet. Det finnes modeller som kan klassifisere bilder av hunder og katter i dag og nettsider med veiledning på hvordan en slik modell kan utvikles. Maskinlæring er en god løsning fordi at det er raskere å klassifisere med modellen enn å selv bestemme. Dette kan hjelpe brukeren å lære å skille mellom hunder og katter.

METRIKKER

Noen mål som brukes for å vurdere hvor godt modellen fungerer er nøyaktighet, presisjon, gjentakelse, F1-score og confusion matrix.

Nøyaktighet (accuracy) er et mål på om modellen gir bildet rett merkelapp. Høy nøyaktighet betyr at modellen er god på å predikere. Nøyaktighet er et godt mål når datasettet er balansert.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Bilde 1: Formel for accuracy.

Presisjon (precision) er et mål på hvor mange positive prediksjoner som er rett klassifisert.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Bilde 2: Formel for precision.

Gjentakelse (recall) er et mål på hvor mange av de positive dataene som ble klassifisert som positive.

Gruppe 40

Ingelin Lovise Vikre

Dato: 29.10.25

$$\text{Recall} = \frac{TP}{TP + FN}$$

Bilde 3: Formel for recall.

Presisjon og gjentakelse går på bekostning av hverandre. Det vil si at med høy presisjon blir det lavere gjentakelse og vice versa. F1 er det harmoniske gjennomsnittet av presisjon og gjentakelse. Ved å regne ut F1 finner man ut om parameterene er vektet likt. Er en av verdiene lave får man en lav F1-verdi. Verdien kan være mellom 0-1(0-100%), og en verdi over 0,5 kan være en bra F1 score. (Kundu, Rohit, 16.12.22)

$$F_1 = \frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}}$$

Bilde 4: Formel for F1-score.

Confusion matrix er en tabell for å måle hvor bra modellen presterer. I tabellen illustreres hva modellen predikerer og om det stemmer med fasiten. Dette hjelper oss å forstå modellen bedre.

TP er når modellen predikerer positiv og merkelapp er positiv.

FP er når modellen predikerer positiv og merkelapp er negativ

TN er når modellen predikerer negativ og merkelapp er negativ.

FN er når modellen predikerer negativ og merkelapp er positiv.

(GeeksForGeeks 30.05.25)

		Predicted label	
		Positive	Negative
True label	Positive	True positive (TP)	False negative (FN)
	Negative	False positive (FP)	True negative (TN)

Bilde 5: Illustrasjon av confusion matrix.

For at modellen skal være vellykket er det laget noen krav:

- Modellen må ha en nøyaktighet på 80%
- Presisjon og gjentakelse bør være over 0,8
- Modellen skal være raskere enn manuell klassifisering

Dette vil jeg beregne etter testing. Etter launch vil jeg gjennomføre brukerundersøkelse for å sjekke brukertilfredsheten.

2: DATA

Dataen som benyttes er bilder av hunder og katter. Modellen skal kunne gjenkjenne om det er hund eller katt på bildet og sortere dem etter merkelappene "hund" og "katt". Det finnes flere datasett på [knaggle.com](https://www.kaggle.com/). Først tok jeg i bruk datasettet "[anthonytherrien/dog-vs-cat](https://www.kaggle.com/anthonytherrien/dog-vs-cat)", men oppdaget at det var AI-genererte bilder. Noe man kan tenke seg påvirker modellens evne til å klassifisere bilder av ekte hunder og katter. Derfor gikk jeg over til å trene på "[bhavikjidadara/dog-and-cat-classification-dataset](https://www.kaggle.com/bhavikjidadara/dog-and-cat-classification-dataset)" som er bilder av ekte hunder og katter.

Dato: 29.10.25

Hvor mye data man trenger avhenger av flere faktorer som kompleksitet til modellen, antall features, støy, hvor stor feilmargin som aksepteres og balanse mellom klassene. Det er en rekke artikler på hvor mye data man bør ha og det nevnes at det bør være mellom 500-1000 per klasse (Zantvoort K., Nacke B., Görlich D., Hornstein S., Jacobi C. og Funk B., 18.12.24). Datasett "[anthonytherrien/dog-vs-cat](#)" har totalt 1000 bilder og har ikke optimal størrelse, mens "[bhavikjidadara/dog-and-cat-classification-dataset](#)" er et mye større datasett på 24 998 bilder. Dette datasettet er en del større enn det som er nødvendig.

Siden brukeren laster opp bildet sitt på nettsiden er det viktig å tenke på personvern. Det kan være elementer i bildet av folk, adresser og osv, som er underlagt GDPR. Derfor bør man legge inn at brukeren må samtykke før modellen tas i bruk. I samtykket skal det stå om hva bildet brukes til, hvor det lagres, hvor lenge det lagres, hvem som har tilgang og mulighet for å be om sletting. Dette er ikke gjort denne gangen, men bør implementeres i fremtiden.

Modellen kommuniserer med nettleseren gjennom en enkel klient-server-struktur. Når brukeren laster opp et bilde via nettsiden, sendes en forespørsel fra nettleseren til serveren hvor modellen ligger. Serveren prosesserer bildet og bruker den trenede modellen til å bestemme om bildet viser en hund eller en katt. Resultatet sendes deretter tilbake til nettleseren og vises til brukeren. Etter klassifiseringen blir bildet slettet fra serveren for å unngå lagring av persondata over tid, noe som bidrar til bedre personvern og etterlevelse av GDPR-regelverket.

I det valgte datasettet er det ikke behov for rensing av data da alle bildene var av hund eller katt. Når modellen kjøres blir det sikret at alle bildene er av samme størrelse. Jeg begynte først å trene med 128x128 piksler, men dette tok for lang tid med det andre datasettet. Derfor er valgt størrelse er 64x64 piksler, som gjør at bildene er små nok til at modellen kjører raskt, men har fortsatt nok detaljer.

Pikslene i bildene blir skalert slik at det har en verdi mellom 0-1. Dette gjøres for at alle input-verdier har samme skala som fører til at modellen blir mer robust og lærer fortere.

3: MODELLERING

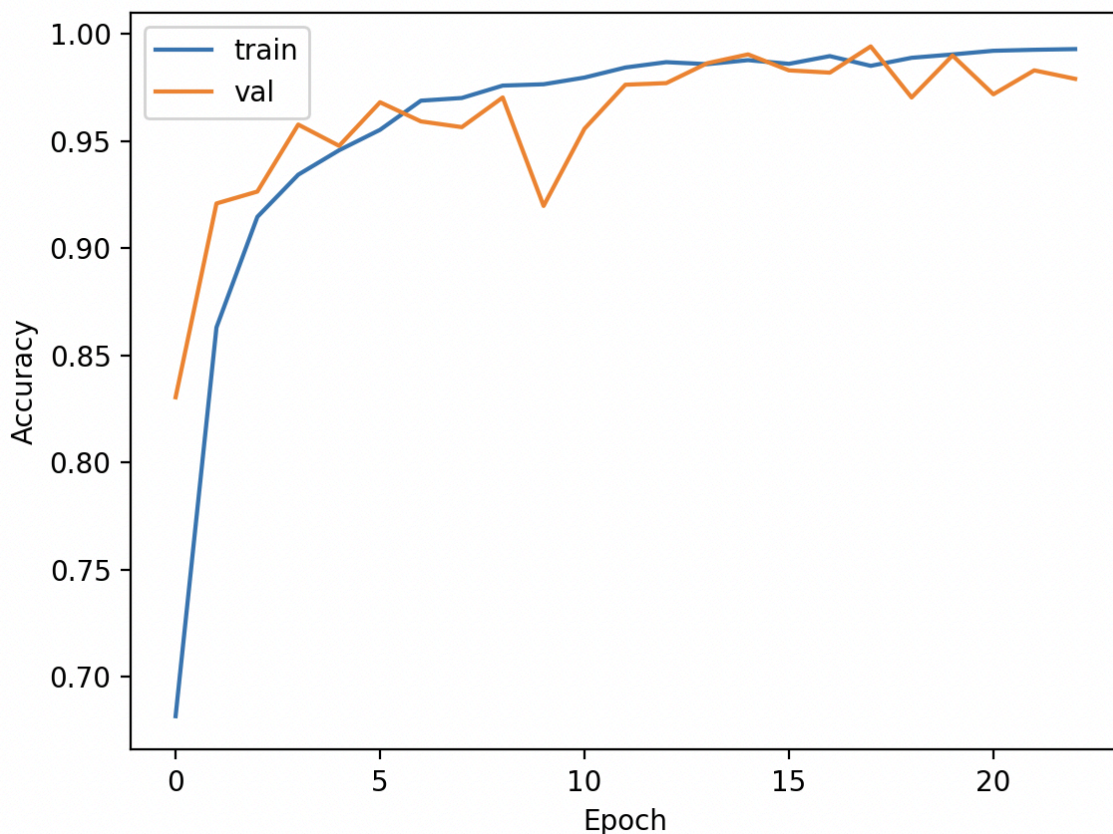
Modellen som er brukt for å binært klassifisere hunder og katter er Convolutional Neural Network (CNN). Convolutional Neural Network er spesielt god til å behandle bilder og visuell data. Binær klassifisering brukes når det skal klassifiseres i to ulike kategorier. Modellen lærer ved å undersøke bilder og gjenkjenner mønstre, som fagre, form, størrelse og osv, i de ulike bildene (GeeksForGeeks 23.07.25). Denne delen av læringen skjer i konvolusjonslaget. Modellen jeg har utviklet har tre konvolusjonslag. Etter konvolusjonslagene blir dataen redusert og lagret i poolinglaget. Til slutt brukes dense-laget for å bestemme hva som er på bildet. I dense-laget lagres dataen som 1 eller 0. Denne

Dato: 29.10.25

modellen bruker supervised learning som er at datasettet har merkelapp og modellen trenes med klassifiserte bilder.

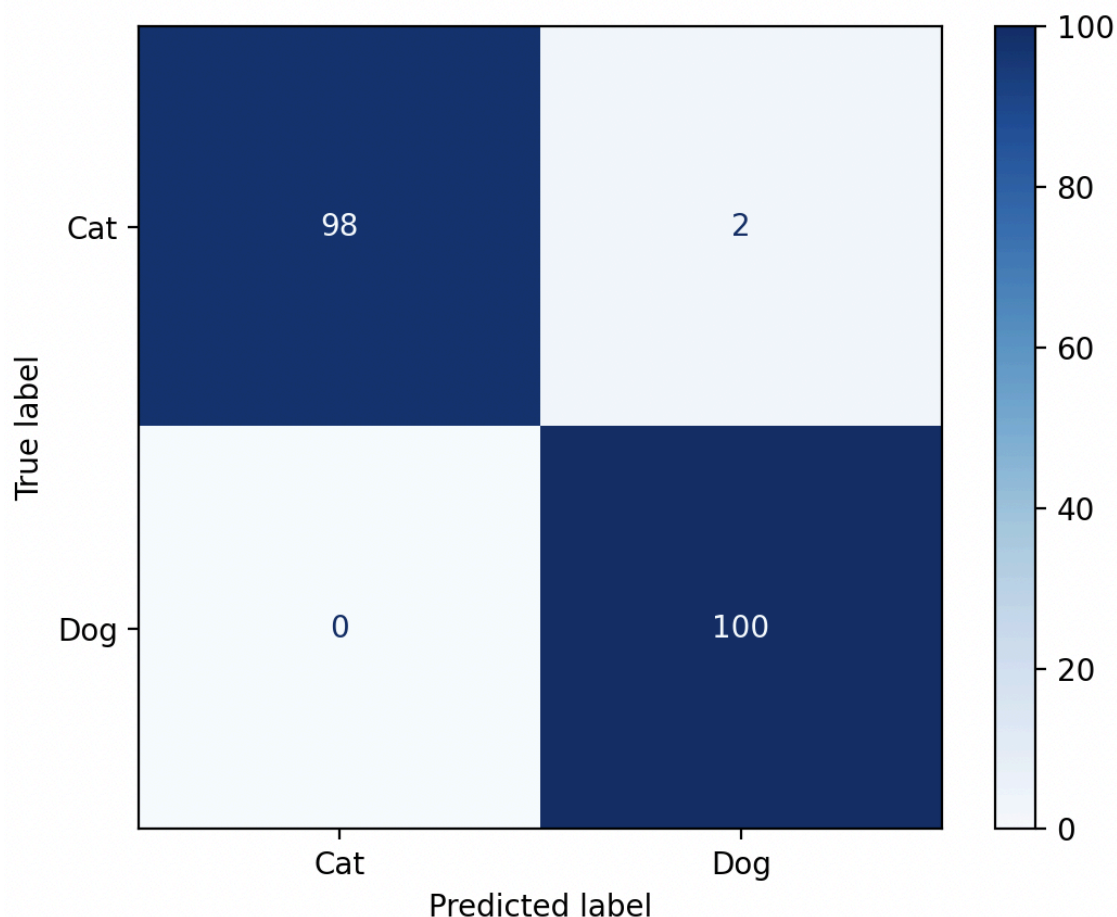
Videre vil jeg se på resultatene fra modellen. Først ble det testet på datasettet med AI bilder av dyr ("[anthonytherrien/dog-vs-cat](#)") og deretter datasettet med ekte bilder av dyr ("[bhavikjikadara/dog-and-cat-classification-dataset](#)").

I den første grafen er det illustrert hvor god nøyaktighet modellen har etter å ha trent på treningsdata og valideringsdata. Treningsdata vises med blå linje og valideringsdataen vises med oransje linje. Modellen har kjørt gjennom dataen 20 ganger (epoch). Etter de første rundene ser man at nøyaktigheten har blitt betydelig bedre enn fra start, og nøyaktigheten flater ut fra runde fem og utover. Valideringsdataen viser omtrentlig lik nøyaktighet som treningsdataen og linjen er relativt flat hvis man ser bort fra duppen på tiende runde. Dette tyder på at det verken er underfittig eller overfittig.



Bilde 6: Graf om accuracy for datasettet "[anthonytherrien/dog-vs-cat](#)".

I confusion matrix kan man se at det var ingen hunder som ble klassifisert som katt, mens det var to katter som ble klassifisert som hund. Dette er en liten feilklassifisering, noe som er forventet fordi at AI-bildene var nok så like innenfor katt eller hund kategori.



Bilde 7: Confusion matrix for datasettet "[anthonytherrien/dog-vs-cat](#)".

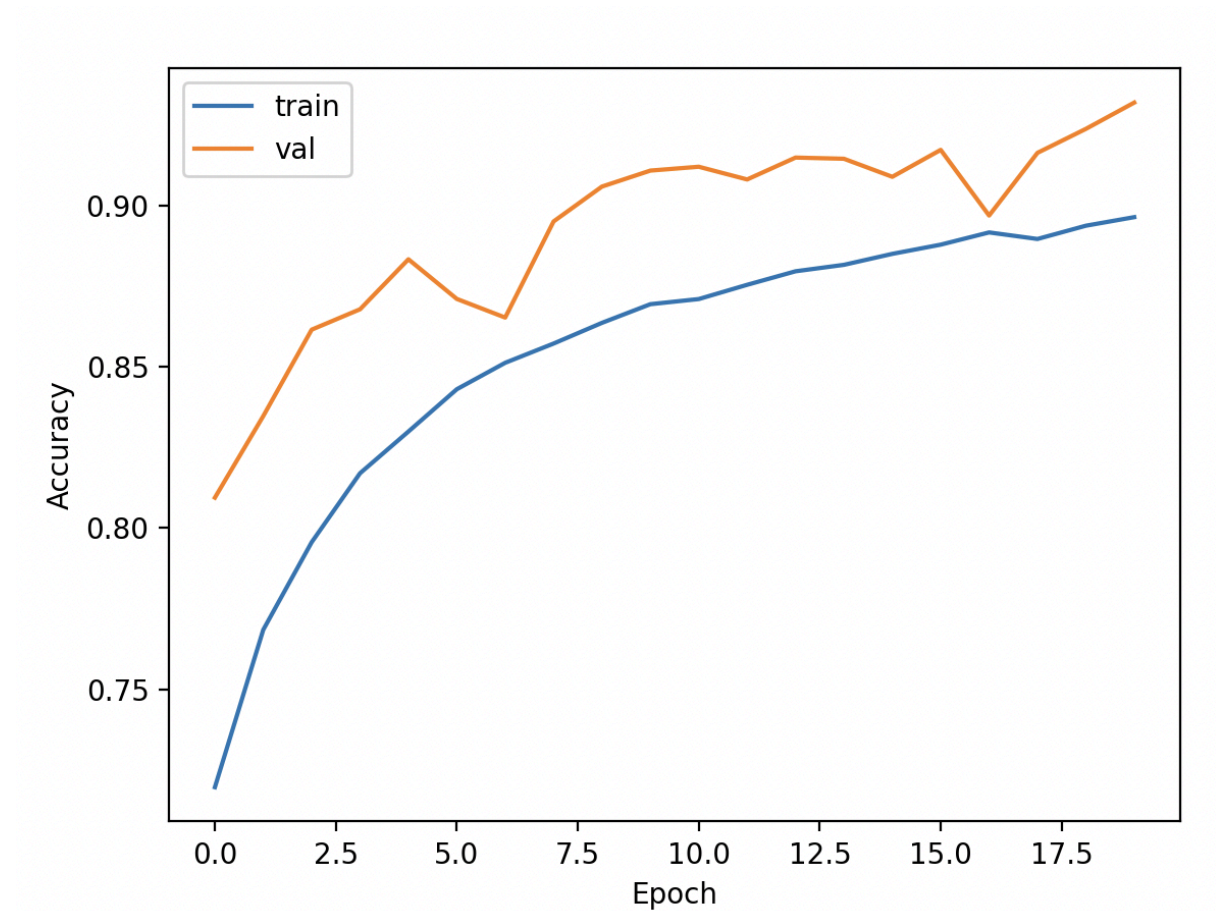
Modellen oppnådde gode resultater på datasettet "[anthonytherrien/dog-vs-cat](#)". Med en presisjon på 0,98, gjentakelse på 1,00 og en F1-score på 0,99 viser modellen høy nøyaktighet og god evne til å skille mellom bilder av hunder og katter. Dette indikerer at modellen generaliserer godt og gjør få feil i klassifiseringen.

Videre testet jeg på datasettet "[bhavikjikadara/dog-and-cat-classification-dataset](#)". Her ser man at valideringsdataen ligger konsekvent høyere enn treningsdataen. Dette kan skyldes bruk av dropout under trening, som reduserer ytelsen midlertidig for å hindre overtilpasning, samt naturlige variasjoner mellom trenings- og valideringsdatasettet. Resultatet tyder på at modellen generaliserer godt til nye data.

Gruppe 40

Ingelin Lovise Vikre

Dato: 29.10.25



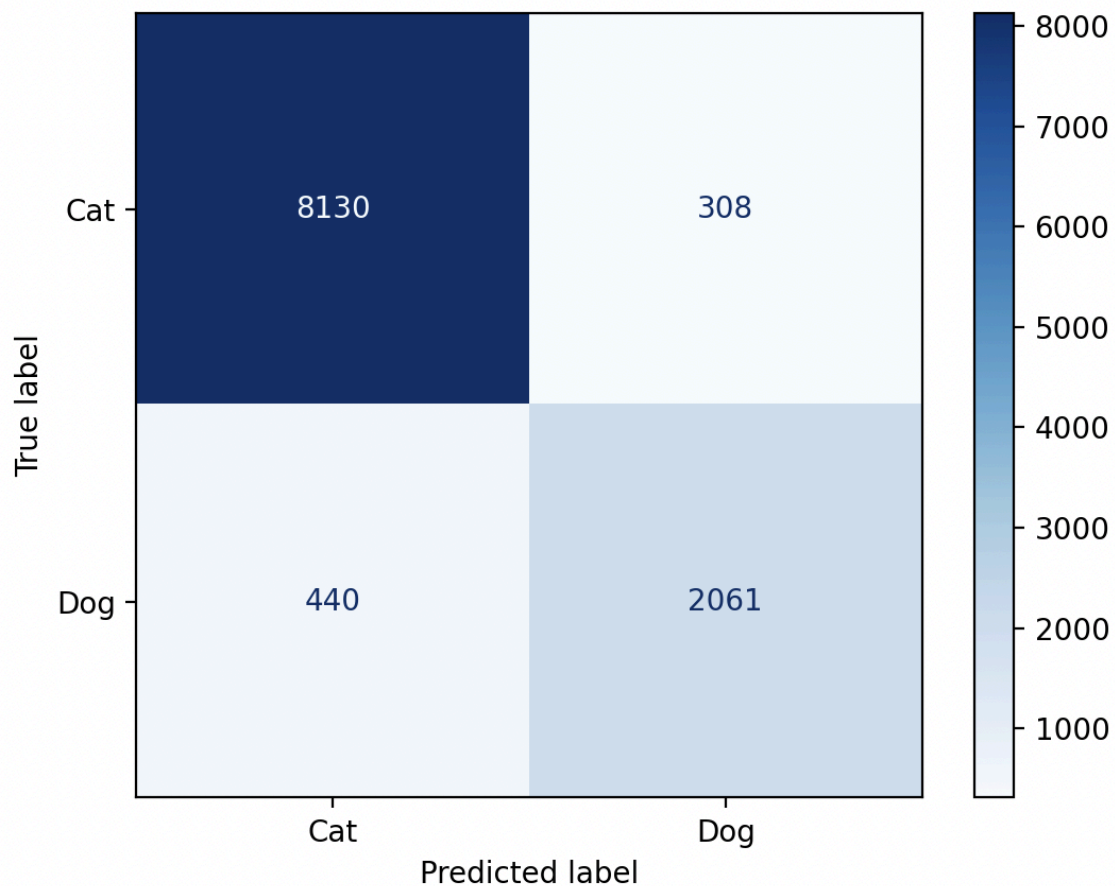
Bilde 8: Graf om accuracy for datasettet "[bhavikjickadara/dog-and-cat-classification-dataset](#)".

For dette datasettet er det flere feilklassifiseringer av katter og hunder. Det er 308 katter som er feil klassifisert som hunder, mens det er 8130 som er rett klassifisert. Det er 440 hunder som er klassifisert som katt og 2061 hunder som er klassifisert som hund. Det viser at det er flere hunder som blir feilklassifisert. Noe som kan være fordi at det er mer variasjon på hundenes størrelse enn hos katter.

Gruppe 40

Ingelin Lovise Vikre

Dato: 29.10.25



Bilde 9: Confusion matrix for datasettet "[bhavikjadara/dog-and-cat-classification-dataset](#)".

Videre var presisjonen 0,87, gjentakelsen 0,82 og F1-score 0,85. Dette er et annet mål som viser at modellen gjør en del feilklassifiseringer, men likevel oppnår en forholdsvis god ytelse. Den noe lavere presisjonen og gjentakelsen bygger opp under argumentet at modellen enkelte ganger forveksler hunder og katter, noe som kan skyldes variasjon i bildekvalitet, lysforhold eller likheter mellom artene.

Ut ifra disse datasettene kan man se at når det er mer variasjon innad i datasettet er det større sjanse for feilklassifiseringer, men man kan likevel lage en god klassifiseringsmodell.

4: DEPLOYMENT

Modellen kan benyttes via nettleser for å klassifisere bilder av hunder og katter. Det er nyttig for brukeren når man lur på om det er hund eller katt på bildet. Dette kan være nyttig for å lære forskjellen på hunder og katter. Siden hunder og katter endrer seg lite over tid er dette en modell som krever lite vedlikehold. Derfor bestemmer jeg at dette skal være en statisk modell som kan brukes til den er utdatert uten vedlikehold.

Om jeg skulle ha jobbet videre med denne modellen ville jeg ha lagt inn GDPR regulering. Det ville vært interessant å utvide modellen til å kunne gjenkjenne flere dyrearter, og implementere den i SoMe miljøer slik at modellen kan lære seg å kjenne hvilke dyr brukeren liker å se.

5: REFERANSER

- Zantvoort K., Nacke B., Görlich D., Hornstein S., Jacobi C. og Funk B. (18.12.24) Estimation of minimal data sets sizes for machine learning predictions in digital mental health interventions. Tilgjengelig fra: https://www.nature.com/articles/s41746-024-01360-w?utm_source=chatgpt.com (Hentet: 24.10.25)
- Kundu, Rohit (16.12.22) F1 Score in Machine Learning: Intro & Calculation Tilgjengelig fra: <https://www.v7labs.com/blog/f1-score-guide> (Hentet: 25.10.25)
- GeeksForGeeks (23.07.25) Getting started with Classification. Tilgjengelig fra: <https://www.geeksforgeeks.org/machine-learning/getting-started-with-classification/> (Hentet: 25.10.25)
- GeeksForGeeks (30.05.25) Understanding the Confusion Matrix in Machine Learning. Tilgjengelig fra: <https://www.geeksforgeeks.org/machine-learning/confusion-matrix-machine-learning/> (Hentet: 25.10.25)
- Punyakeerthi BL (21.06.24) Understanding Feature Scaling in Machine Learning. Tilgjengelig fra: https://medium.com/@punya8147_26846/understanding-feature-scaling-in-machine-learning-fe2ea8933b66 (Hentet: 26.10.25)
- ChatGPT (2025) *Samtale med ChatGPT (GPT-5-modell)*, 30. oktober. Tilgjengelig fra: <https://chat.openai.com/> (Tilgang: 30. oktober 2025).