

Sluttvurdering – Anvendt maskinlæring

Emne: Introduksjon til maskinlæring og KI-modeller [10 stp.]

Temaer:

- Introduksjon KI og maskinlæring
- Introduksjon til nevrale nettverk

Vurderingen inkluderer pensum fra begge temaene og gir en felles karakter A–F for hele emnet.

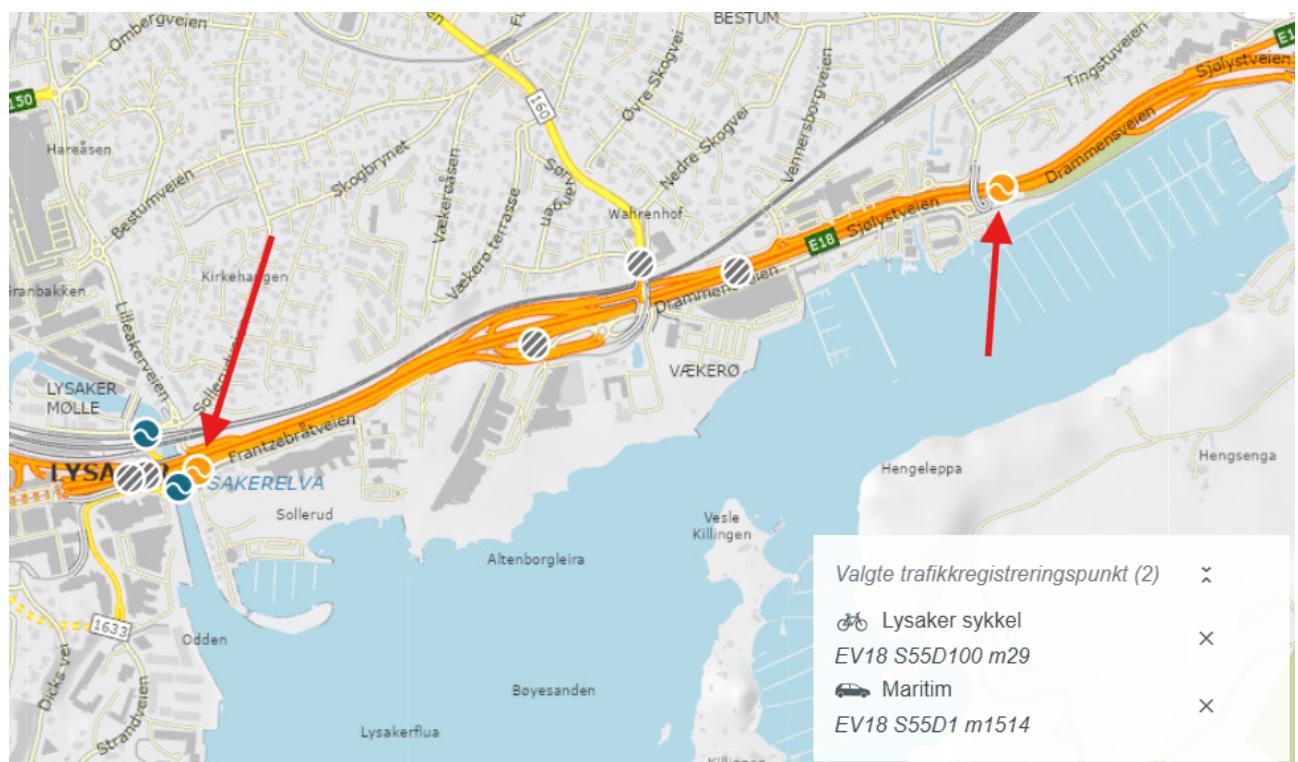
Individuell: Vurderingen er individuell. Samarbeid med medstudenter er ikke tillatt.

Innlevering: Oppgaven leveres på teams i valgfri format. Inkluder alle visualiseringer og kode.

Datasett

I denne oppgaven er det lagt til to datasett, begge for 2025:

1. **lysaker_tellinger.csv:** tellinger av biler og syklister.
2. **lysaker_vær.csv:** værdata for Lysaker-området.



Figur 1: Sykkel og bilteller stasjoner

Sykkel- og biltellinger – Lysaker

Besvar følgende oppgaver basert på vedlagte datasett.

Oppgave 1 – Statistikk

Finn gjennomsnitt, median og standardavvik for:

1. Daglig sykkeltelling
2. Daglig biltelling
3. Temperatur og nedbør

Visualiser og gi en kort beskrivelse for sammenhengen mellom:

1. Sykkeltrafikk og temperatur
2. Sykkeltrafikk og nedbør

Oppgave 2 – Regressjon

Prediksjon

Foreta en prediksjon av antall syklister på grunnlag av:

1. Temperatur
2. Nedbør

Visualisering

- Visualiser resultater
- Vurder hvilket grunnlag (temperatur eller nedbør) som gir best resultat
- Argumenter for valg av regresjonsmodell

Oppgave 3 – Nevrale nettverk: Bildebehandling i trafikktelling

Statens vegvesen bruker i dag induktive sensorer for trafikktelling i Oslo. Sensorene består av ledninger lagt i flere lag i asfalten, som danner et magnetfelt når det sendes strøm gjennom dem. Et trafikktellingsapparat tolker deretter endringer i magnetfeltet når et metallobjekt passerer.

Signalet fra en sykkel har en karakteristisk form, og sensoren klarer derfor å skille sykler fra for eksempel barnevogner. Likevel finnes det flere typer objekter som ikke registreres korrekt, eller som ikke kan skilles fra hverandre med denne teknologien.

Oslo kommune ønsker derfor å forbedre tellingene ved å ta i bruk bildebehandling (Computer Vision) for å kunne klassifisere flere typer objekter, blant annet:

- ulike typer sykler (vanlige, med vogn, med barnesete, el-sykler),
- barnevogner,
- sparkesykler,
- elektriske mopeder,
- segways.

Formålet er å kunne få mer nøyaktig og detaljert statistikk over trafikk på sykkelveier.

3.1 Forklar formålet

Forklar med egne ord hvordan Computer Vision-modeller kan brukes til å klassifisere ulike objekter på sykkelveier.

Krav:

- Beskriv hva et nevratlt nettverk “ser” i et bilde.
- Forklar forskjellen på objektgjenkjenning og bildeklassifikasjon.
- Nevn én fordel og én utfordring ved å bruke bildebaseret klassifisering i trafikk.

3.2 Datasett

Beskriv hvilket datasett som er nødvendig for å kunne trenere en god modell.

Krav:

- Hvilke typer bilder trengs? (lysforhold, vinkler etc.)
- Hvor mange eksempler omtrent bør hver kategori ha?
- Hvorfor må datasettet være balansert?

3.3 Arbeidsflyt – Fra problem til modell

Beskriv de viktigste stegene du ville fulgt for å gå fra problemforståelse til ferdig modell.

Krav:

a) Innsamling av data

- Hvordan ville du skaffet bilder av de ulike objektene?
- Hva må man passe på når man samler inn data i offentlig rom?

b) Datamerking (labeling)

- Hvordan markerer man objekter i et bilde? (bounding boxes, klasser)
- Hvem bør gjøre merkingen og hvorfor?

c) Valg av modell

- Nevn én passende modelltype (f.eks. CNN, MobileNet, YOLO, Vision Transformer).
- Forklar kort hvorfor den er egnet til dette formålet.

d) Trening og testing

- Hva menes med treningsdata og testdata?
- Hvorfor er evalueringsetrikker (accuracy, precision/recall) viktige?

e) Implementering og drift

- Hvordan kunne modellen brukes i praksis i Oslo kommune?
- Hvilke utfordringer kan oppstå i virkelige situasjoner?