

# Prosjektrapportering

Dette dokumentet er en sjekkliste over forskjellige ting vi ser på når vi vurderer prosjektene, og som forhåpentligvis avklarer forventninger. Denne sjekklisten er basert på sjekklisten for INF264 laget av Natacha Galmiche og Pekka Parviainen.

## Rapport

### Tilnærming og designvalg

- Hva var din tilnærming til å analysere dataene du fikk? Hvorfor? Er det noen utfordringer du bør ta hensyn til med dette datasettet?
- Samsvarer alle dine valg med dataene og oppgaven du fikk? Rettferdiggjøre.
- Har du prøvd andre tilnærminger som ikke var vellykket og derfor ikke ble brukt i den endelige modellen? Vet du hva som gikk galt? Ellers, hva ville være din beste forklaring?
- Hva var din tilnærming til å analysere resultatene dine? Hvorfor?

### Resultater

- Fremstår den valgte modellen sammen med ytelsen klart og tydelig i rapporten din? De skal ikke skjules et sted midt i et avsnitt.
- For hver rapporterte ytelse, har du tydelig spesifisert beregningen som brukes og på hvilket datasett? Og har du spesifisert hvordan du deler datasettet ditt i utgangspunktet?
- Er resultatene dine overraskende? Med andre ord, i forhold til din generelle kunnskap innen data science, din forståelse av dataene og problemet, og i forhold til tilliten du legger i designvalgene og implementeringen, forventet du slik ytelse?
- Er resultatene dine tilfredsstillende? Med andre ord, er modellen din bedre enn noen grunnlinjemodell? Tror du at du kan bruke modellen din i det virkelige liv? I så fall, under hvilke forhold?
- Hvis du fikk ubegrenset tid, hva ville du prøve for å forbedre ytelsen din?
- I hvilke tilfeller sliter den valgte modellen? I hvilke tilfeller gir modellen gode resultater? Hvordan vil du forklare det?

### Figurer

- Har du inkludert alle relevante figurene i rapporten din?
- Er figurene kommenterte (aksebeskrivelse, forklaring)?
- Vet vi nøyaktig på hvilket datasett figurene dine ble generert?
- Forklarer du hva som kan konkluderes fra figurene?

### Klarhet i rapporten

- Er rapporten din godt strukturert eller er det ett stort avsnitt?
- Kan leseren forstå rapporten din fullt ut uten å lese koden?

## Kode

### Forståelighet

- Er det lett å lese koden din (bruk av relevante variabel-/funksjons-/klassenavn? nyttige kommentarer? hensiktsmessig oppdeling inn i flere kodebiter)?
  1. "Ikke egentlig. For å forstå koden din måtte leseren være kjent med prosjektbeskrivelsen, vite hvordan oppgaven skulle løses og ha gode programmerings- og pythonferdigheter."
  2. "Til en viss grad. For å følge med på hva du gjør, må leseren ha et raskt blikk på prosjektbeskrivelsen først."
  3. "Ja, alle med noen grunnleggende python-ferdigheter ville forstå koden din, selv om de ikke fikk prosjektbeskrivelsen."
- Er koden din klar og konsistent (variabel navnekonvensjoner er konsistente gjennom hele prosjektet, og koden din ser ikke ut som en copy-paste fra forskjellige kilder)?

### Reproduserbarhet

- Er all koden som du brukte for å lage resultatene, figurene, osv. med i de innsendte filene?
- Hvis vi kjører koden på nytt selv, vil vi få nøyaktig de samme resultatene? Er alle tilfeldige prosesser seedet med et fast seed?
- Gir kjøring av cellene i Jupyter-notebooken fra topp til bunn de forventede resultatene (for eksempel skal en celle ikke kalle en funksjon som er definert senere)?