

Gruppeoppgaver MIEVU 4020 våren 2023

Dag 1

Ønsket læringsutbytte

- Du skal kjenne til noe av bakgrunnen for at dyp læring har blitt en stor suksess i maskinsyn, naturlig språkprosessering og spillintelligens.
- Du skal kjenne til hvordan man ofte måler suksess i dyp læring og utfordringer knyttet til stabilitet/hallusinasjoner.
- Du skal kjenne til modellen for lineær regresjon.
- Du skal kjenne til hvordan gradient descent fungerer.
- Du kan trene din egen lineære regresjonsmodell.

Oppgaver

Oppgave 1. Har din bedrift/organisasjon data eller problemer hvor dyp læring kan være et nyttig verktøy? Hvilke eventuelle utfordringer ser du ved å bruke dyp læring på dine data/problemer? Diskuter.

Oppgave 2. Stabilitet er en utfordring i dyp læring. Har du ideer til hvordan vi kan lage robuste systemer?

Oppgave 3. I dag har vi sett på lineær regresjon.

- a) Skriv ned hvilken funksjon vi ønsker å lære når vi jobber med lineær regresjon.
- b) Hvilke fordeler/ulempes har slike funksjoner?
- c) Vi har sett på loss funksjonen

$$L(\hat{y}, y) = (y - \hat{y})^2.$$

Finnes det andre loss funksjoner du kan bruke?

Oppgave 4. Regn ut gradienten til funksjonen.

- a) $\Gamma(\theta_1, \theta_2, \theta_3) = \theta_1 \theta_2 + \theta_3^2$
- b) $\Gamma(\theta_1, \theta_2) = \sqrt{\theta_1 + \theta_2}$
- c) $\Gamma(\theta_1, \theta_2) = 2\sqrt{\theta_1^2 + \theta_2^2}$
- d) $\Gamma(\theta_1) = \max\{0, \theta_1\}$
- e) $\Gamma(\theta_1) = |\theta_1|$

Oppgave 5.

Oppgave 6. La

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}, \quad \text{og} \quad v = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

og regn ut følgende uttrykk, enten for hånd eller i Python.

- a) $A + B$
- b) $5A$
- c) $(A + B)v$
- d) $A^\top$

Oppgave 7. a) Finn en ligning for linja som går gjennom punktene $(3, 3)$ og $(2, 1)$.

b) Tegn linja.

Oppgave 8. Du er gitt datasettet i Figur 1. Legg merke til at dette datasettet består av vektorer i $\mathbb{R}^2$.

- a) Kan du finne en funksjon $\Psi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, slik at punktene kan separeres lineært i $\mathbb{R}^3$, dvs. at

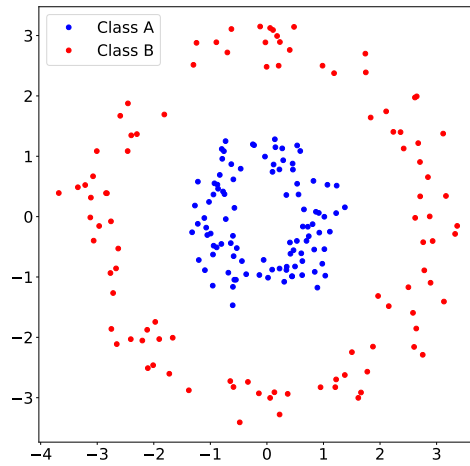
$$\langle w, \Psi(x) \rangle + b$$

vil kunne separere punktene. (Hint: Tenk på avstand fra origo).

- b) Kan du gjøre det samme med en funksjon $\Psi: \mathbb{R}^2 \rightarrow \mathbb{R}$?

Oppgave 9. Bytt ut datasettet i filen `1_2_Ames_housing_dataset_LR.ipynb` og bruk en lineær model for å tilpasse de nye dataene. Kan du visualisere den nye modellen?

Oppgave 10. a) Beskriv med ord hva ideen bak «gradient descent» er.



Figur 1. Datasett med to klasser.

- b) I treningprosessen for dyp læring søker vi etter parametere θ , slik at

$$\Gamma(\theta) := \sum_{i=1}^K L(\phi(x^{(i)}, \theta), y^{(i)})$$

er liten. Gradient descent søker etter parametere θ ved å bruke følgende iterasjon

$$\theta^{(k)} = \theta^{(k-1)} - \eta_k \nabla \Gamma(\theta^{k-1})$$

Steglengden $\eta_k > 0$ velges før man starter iterasjonene. Kan du tenke deg andre måter denne hyperparameteren kan velges på?