

# Övningstenta maskininlärning AI21

---

Inga hjälpmedel, endast penna och sudd. Samtliga lösningar görs på separat provskrivningspapper. I samtliga uppgifter krävs fullständig lösning, endast svar ger 0 poäng. För att få godkänt krävs 8 poäng. För att få VG krävs 15 poäng.

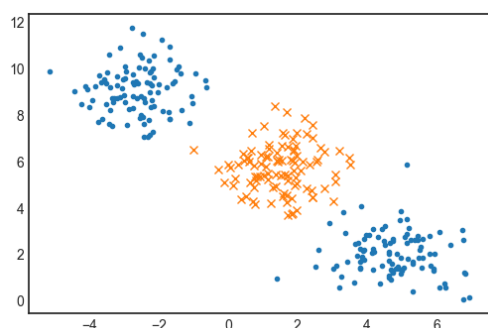
---

1. Beskriv vad  $k$ -folded cross-validation är för något och ge exempel på hur det kan användas.

(2p)

---

2. Datasetet du ser här är inte linjärt separerbart, det har två olika klasser. Beskriv hur du skulle gå tillväga för att klassificera punkterna i datasetet med SVM.



(2p)

---

3. Vad är stora skillnaden mellan regression och klassificering?

(2p)

---

4. Beskriv begreppen input layer, hidden layer, outputlayer i ett multilayered perceptronnätverk.

(2p)

---

5. Beskriv fördelarna att använda regulariseringsmetoder som  $\ell_1$  och  $\ell_2$  regularisering för linjär regression. Beskriv även hur man kan gå tillväga för att välja penalty parametern  $\lambda$ .

(3p)

---

6. Du har ett dataset som ser ut som följande:

x	y	klass
0	1	1
1	0	1
3	2	1

x	y	klass
2	2	0
2	1	0

Använd KNN med  $k = 2$  och Euklidisk distansmått för att klassificera testpunkten  $(1, 1)$ .

(3p)

7. Använd min-max normalization för att skala  $X_{train}$  och  $X_{test}$ :

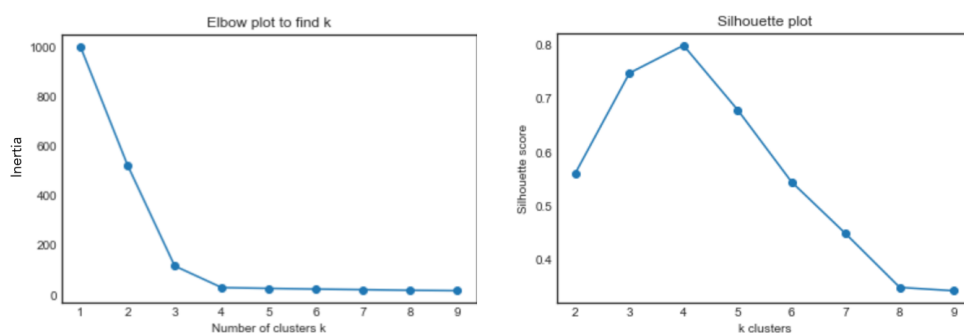
$$X_{train} = \begin{pmatrix} 1 & 1 \\ 2 & 4 \\ -2 & 2 \end{pmatrix}, X_{test} = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$$

Tips: formeln för min-max normalization för en feature  $\mathbf{x}$ :

$$\mathbf{x}' = \frac{\mathbf{x} - \mathbf{x}_{\min}}{\mathbf{x}_{\max} - \mathbf{x}_{\min}}$$

(3p)

8. Företaget Xhopper driver en webshop som säljer saxar av olika slag och har samlat in data på kunders shoppinghistorik, features som datum, typer av produkter och en del andra beteenderelaterade features. De har hört att AI är coolt, och att k-means skulle kunna hjälpa dem att segmentera kunderna. Med några paket och typexempel har de lyckats skala datan, använda k-means på datasetet och fått fram följande plots:



Du gör praktik på företaget och de frågar dig som expert vad det är de fått fram, vad k-means är och vad de kan göra med resultatet.

(4p)