

i Forside

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Skriftlig eksamen i IN2110

2023, vår

Varighet: 13. juni, 15:00 til 19:00 (4 timer)

Tillatte hjelpemidler: Ingen

Det er viktig at du leser denne forsiden nøye før du starter. Vi anbefaler også at du ser igjennom hele oppgavesettet før du starter.

1 Vektor-typer

I forbindelse med modellering av betydningen til ord har vi i forelesningene nevnt både tradisjonelle vektor-representasjoner og såkalte *word embeddings* – diskuter kort både likheter og ulikheter mellom disse representasjonene.

Skriv ditt svar her

Format

B

I

U

x_2

x^2

I_x

Words: 0

Maks poeng: 5

2 P R 1

I vedlagte tabell (pdf) finner du resultatet av en tenkt modell som har predikert om våre ti samples 01 – 10 er i kategorien "Blue" eller "Green". Tabellen viser både gullstandarden og prediksjonene. Vi skal nå evaluere hvor godt modellen fungerer i forhold til kategori "Green". Angi for hver prediksjon om den er en True Positive (TP), True Negative (TN), False Positive (FP), eller False Negative (FN).

Angi rett kategori for Predicted, med hensyn til "Green" kategori.

	TP	TN	FP	FN
Sample-01	☐	☐	☐ ✓	☐
Sample-02	☐	☐ ✓	☐	☐
Sample-03	☐ ✓	☐	☐	☐
Sample-04	☐	☐ ✓	☐	☐
Sample-05	☐ ✓	☐	☐	☐
Sample-06	☐	☐ ✓	☐	☐
Sample-07	☐	☐ ✓	☐	☐
Sample-08	☐	☐ ✓	☐	☐
Sample-09	☐	☐ ✓	☐	☐
Sample-10	☐	☐ ✓	☐	☐

Maks poeng: 5

Sample-ID	Gold	Predicted
Sample-01	Blue	Green
Sample-02	Blue	Blue
Sample-03	Green	Green
Sample-04	Blue	Blue
Sample-05	Green	Green
Sample-06	Blue	Blue
Sample-07	Blue	Blue
Sample-08	Blue	Blue
Sample-09	Blue	Blue
Sample-10	Blue	Blue

3 P R 2

I forrige oppgave fant du et visst antall av TP, TN, FP og FN. Legg inn disse fire verdiene i feltet under. (Eksempel: TP:0, TN:0, FP:14, FN:14)

Skriv ditt svar her

Angi formelen for Precision, og regn ut Precision for prediksjonene med hensyn til "Green" kategori. Du kan skrive utregningen i matematikkverktøyet eller som enkel tekst.

Skriv ditt svar her

Angi formelen for Recall, og regn ut Recall for prediksjonene med hensyn til "Green" kategori.

Skriv ditt svar her

Maks poeng: 6

4 P R 3

Vi bruker de samme dataene som i P R 1:

a) Angi et nytt sett med prediksjoner som gjør at både Precision og Recall blir på 0,5, med hensyn til "Green".

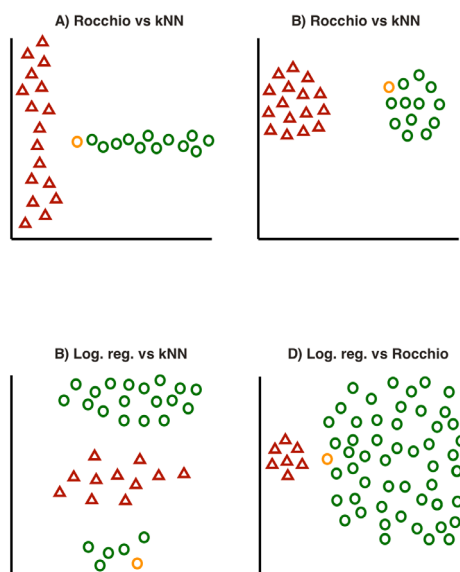
b) Angi for hver prediksjon du har oppgitt, om den er en True Positive (TP), True Negative (TN), False Positive (FP), eller False Negative (FN).

Du kan kopiere tabellen under inn i besvarelsen, og fylle ut prediksjoner "Blue" eller "Green" som oppfyller ovenfor nevnte krav. samt om denne er TP, FP, TN, eller FN.

Sample (gold label)	Prediction	TP / FP / TN / FN
Sample-01 (Blue)		
Sample-02 (Blue)		
Sample-03 (Green)		
Sample-02 (Blue)		
Sample-03 (Green)		
Sample-02 (Blue)		
Sample-02 (Blue)		
Sample-02 (Blue)		
Sample-02 (Blue)		
Sample-02 (Blue)		

Skriv ditt svar her

5 Klassifikatorer



Dette er en refleksjonsoppgave med 4 deloppgaver (A, B, C, D).

Under ser du fire diagrammer som viser treningseksempler for et binært klassifikasjonsproblem i to dimensjoner. Treningseksempler for de to klassene vises som *trekanter* (i rødt) og *sirkler* (i grønt). I tillegg vises et test-eksempel (i gult) som i henhold til gullstandarden tilhører sirkel-klassen. For hvert diagram oppgis også to ulike algoritmer for klassifikasjon. For hvert diagram skal du argumentere for hvorvidt én algoritme eventuelt vil være mer egnet enn den andre for å klassifisere test-eksemplet.

6 Logistisk regresjon

1. Hvordan kan man anvende logistisk regresjon for klassifiseringsproblemer med flere enn 2 kategorier? Svar med 3–5 setninger, og angi formelen som brukes for logistisk regresjon med $K > 2$ klasser. Forklar formelen.
2. Hva er rollen til tapsfunksjonen (*loss function*) i trening av maskinlæringsmodeller som logistisk regresjon? Forklar i 2–3 setninger.

Skriv ditt svar her

Format ▾

B *I* U x_2 x^2 I_x          Σ



Words: 0

Maks poeng: 8

7 Hidden Markov Models

Vi ønsker å estimere en liten HMM for å predikere dagens temperatur (med 2 mulige tilstander: *varm* eller *kjølig*) basert på visuelle observasjoner tatt fra vinduet (med 3 mulige verdier: *sol*, *regn* eller *overskyet*).

Vi har derfor samlet værdata fra 3 (uavhengige) uker:

Uke 1:

Dag	Tilstand	Observasjon
1	Varm	Sol
2	Varm	Overskyet
3	Varm	Sol
4	Kjølig	Regn
5	Kjølig	Overskyet
6	Kjølig	Overskyet
7	Kjølig	Overskyet

Uke 2:

Dag	Tilstand	Observasjon
1	Kjølig	Overskyet
2	Kjølig	Regn
3	Kjølig	Sol
4	Kjølig	Overskyet
5	Varm	Sol
6	Varm	Overskyet
7	Kjølig	Overskyet

Uke 3:

Dag	Tilstand	Observasjon
1	Kjølig	Regn
2	Varm	Sol
3	Kjølig	Overskyet
4	Varm	Sol
5	Varm	Sol
6	Kjølig	Regn
7	Kjølig	Overskyet

1. Estimer transisjonmodellen for værtilstanden (med to ulike verdier: varm eller kjølig) fra en dag til den neste, basert på de 3 sekvensene over. Vis transisjonmodellen du har estimert i en tabell.
2. Estimer emisjonsmodellen mellom tilstandene og observasjonene basert på de 3 sekvensene. Vis også emisjonsmodellen din i en tabell.
3. Oppdater nå emisjonsmodellen for å inkludere Laplace smoothing med $\alpha = 1$, og vis den i en ny tabell.

Skriv ditt svar her

Format

B
I
U
 x_2
 x^2
 I_x







 Ω


 Σ



8 Dependenssyntaks 1

Hva kjennetegner såkalt dependensgrammatikk? Forklar kort hvordan dependensgrammatikk skiller seg fra frasestrukturgrammatikk.

Skriv ditt svar her

Format

B

I

U

x_2

x^2

I_x

























Words: 0

Maks poeng: 5

9 Dependenssyntaks 2a

I dependensgrammatikken er det regler for om en dependensgraf må være syklisk eller må være asyklisk. Det samme gjelder for egenskapen transitiv. Også når det gjelder om en graf må være projektiv eller ikke projektiv, har vi av hensyn til parsingen satt en regel.

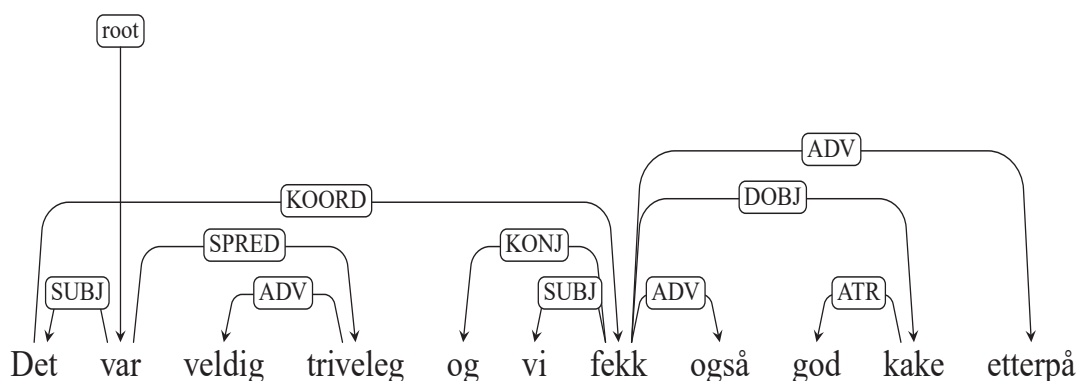
Vedlagte dependensgraf inneholder en feil som gjør at den bryter med reglene vi har for dependensgrafer. Hvilken egenskap er det?

Velg ett alternativ:

- ☐ Grafen er syklisk
- ☐ Grafen er asyklisk
- ☐ Grafen er projektiv
- ☐ Grafen er ikke projektiv
- ☐ Grafen er transitiv
- ☐ Grafen er antitransitiv



Maks poeng: 3



10 Dependenssyntaks 2b

I dependensgrammatikken er det regler for om en dependensgraf må være syklisk eller må være asyklisk. Det samme gjelder for egenskapen transitiv. Også når det gjelder om en graf må være projektiv eller ikke projektiv, har vi av hensyn til parsingen satt en regel.

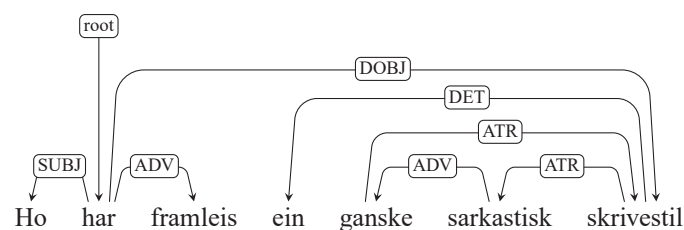
Vedlagte dependensgraf inneholder en feil som gjør at den bryter med reglene vi har for dependensgrafer. Hvilken egenskap er det?

Velg ett alternativ:

- ☒ Grafen er syklisk
- ☐ Grafen er asyklisk
- ☐ Grafen er projektiv
- ☐ Grafen er ikke projektiv
- ☐ Grafen er transitiv
- ☐ Grafen er antitransitiv



Maks poeng: 3



11 Dependensparsing

Ta utgangspunkt i vedlagte graf og angi en fullstendig transisjonssekvens som vil resultere i denne grafen ved bruk av "arc standard"-algoritmen for dependensparsing.

Denne har tre mulige transisjoner: SHIFT , LEFT-ARC , or RIGHT-ARC

Du skal for hver transisjon angi:

- innholdet i stack
- innholdet i inputbuffer
- valgt transisjon
- eventuell relasjon som legges til grafen.

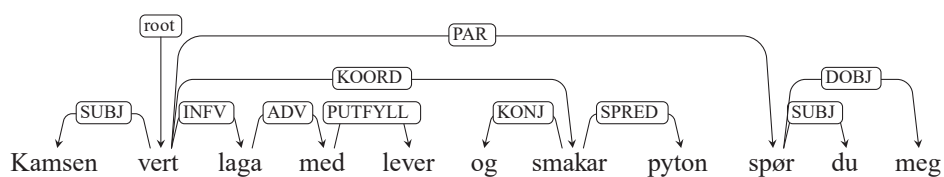
Merk at du skal starte og slutte med *kun* rot-etiketten på stacken.

Når vi starter har inputbufferen denne setningen:

Kamsen vert laga med lever og smakar pyton spør du meg

Skriv ditt svar her

Format
|
B
I
U
x₂
x²
I_x
|
📄
📋
|
↶
↷
↺
|
≡
≡
|
Ω
📊
|
✎



- Skriv ditt svar her**

13 Dialogsystemer

1	hei !
2	hei ! er du Anna ?
3	ja jeg er Anna
4	er du sulten Anna ?
5	ja !

er du sulten ?

Hva vil vår IR-basert chatbot svare på dette spørsmålet? Vis beregningene dine.