

Logistisk regresjon

Først litt linær regresjon

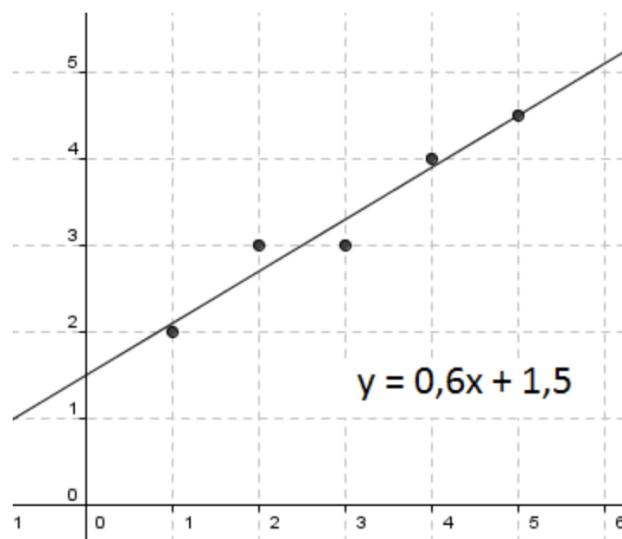
“Hva blir prisen på boligen”
gitt...

Konkret tall

Dersom man har har følgende observasjoner

x - verdi	1	2	3	4	5
y - verdi	2	3	3	4	4,5

Plotter man disse i et koordinatsystem får man:



Dataprogrammet gir oss funksjonsuttrykket til den linjen som passer best med måledataene. I dette tilfellet er det $y=0,6x+1,5$

Linær regresjon formel

$$y = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b$$

=

$$y = w^T x + b$$

$$y = \sigma(\mathbf{w}^T \mathbf{x} + b)$$

↓

hvor $\sigma(z) = \frac{1}{1 + e^{-z}}$

Logistisk regresjon

Ikke lenger et konkret tall, men «hva er sannsynligheten for at (objektet havner i klasse a eller b)»

Hvordan gå fra konkret tall til sannsynlighet??

Sigmoidfunksjonen tvinger regresjonsresultatet til å bli en sannsynlighet, altså tall mellom 0 og 1

Hva betyr alt i formelen?

- y : er outputen fra klasifikasjonsmodellen. Kan være alt mellom 0 (ikke medlem) og 1 (medlem).
- σ : er altså en funksjon som får resultatet vårt til å ligge mellom 0 og 1.
- w : er vekten til en feature
- T : svarer til antall vekter
- x : er en feature
- b : er skjæringspunktet dvs verdien av y hvis alle features x er 0

$$\frac{1}{1 + e^{-(w \cdot x + b)}}$$

$$y = \sigma(\mathbf{w}^T \mathbf{x} + b)$$

↓

hvor $\sigma(z) = \frac{1}{1 + e^{-z}}$

Sentimentanalyse...

Pixels

Nei.

— Publisert 05.08.15, kl 18.00. / Oppdatert 06.08.15, kl 13.19. —



Adam Sandler og Kevin James er tilbake igjen, og *Pixels* er den dårligste filmen jeg har sett på kino noen gang.

Komedien er ikke morsom. Figurene er like spennende som dem du finner i en middels god oppvaskmiddelreklame, og Sandler slentrer så giddesløst rundt foran kamera at det virker som han demonstrativt surmuler over størrelsen på lønnssjekken.

It's **bad**. There are virtually no surprises, and the writing is **poor**.
So why was it so **enjoyable**? For one thing, the cast was **great**.
Another **nice** touch is the music.
I was overcome with the urge to get off the couch and start dancing.
It sucked **me** in, and it'll do the same to **you**.

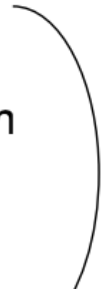
Eksempel

Sannsynligheten for positiv
eller negativ anmeldelse

Feature	Ord i dokument	Antall	Vekt
X1	Positive ord	3	2,5
X2	Negative ord	2	-5.0
X3	1 if «no» ∈ doc 0 otherwise	1	-1.2
X4	Pronomen	3	0.5
X5	1 if «!» ∈ doc 0 otherwise	0	2.0
X6	Log(antall ord)	Ln(66)=4.19	0.7

Hvordan finne vektene w og skjæringspunktet b ?

Gradient descent

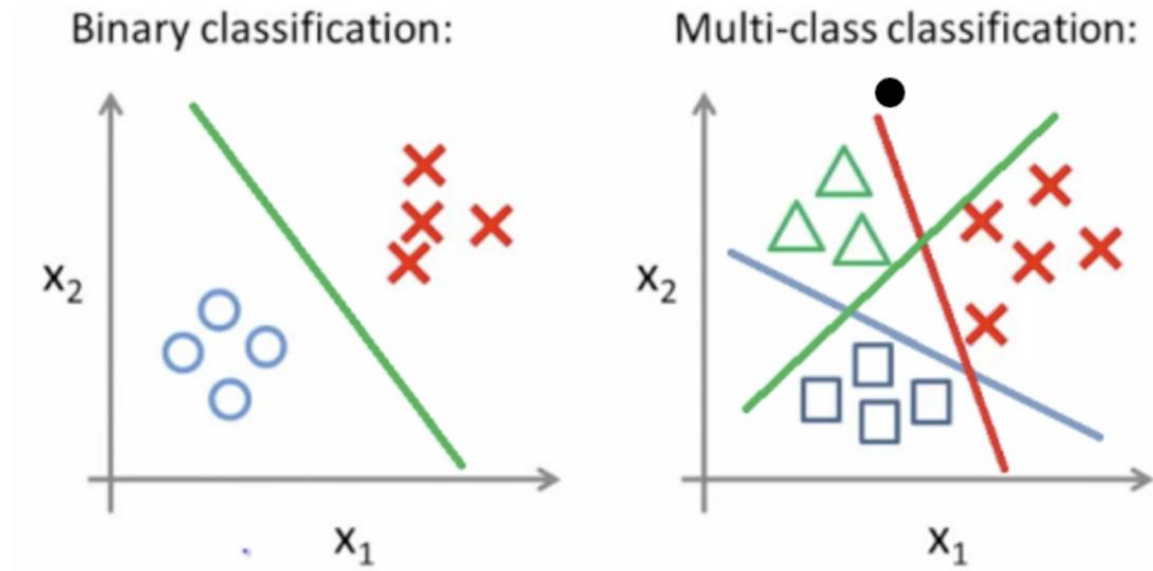
- ▶ **Gradient descent** er en populær optimeringsalgoritme
 - ▶ Gradient descent fungerer slik:
 1. Vi starter med tilfeldige verdier for θ
 2. Vi tar en treningseksempel $(x^{(i)}, y^{(i)})$
 3. Vi beregner hvordan vi best kan endre θ slik at vi reduserer tapet på eksempelet
 4. Vi endrer deretter θ verdiene med en liten inkrement i denne retningen
 5. Vi gjentar for alle treningseksemplene
- 

Logistisk regresjon: multiklasser

- Men hva om vi har tre klasser: positiv, negativ og nøytral? → SOFTMAX
- Erstatte sigmoidfunksjonen!

$$P(y = i | \mathbf{x}) = \text{softmax}(\mathbf{w}_i^T \mathbf{x} + b_i)$$

↙ hvor $\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$



Hvilken funksjon kan vi bruke i multinomial (=“multi-klasser”) logistisk regresjon?

- a) Softmin
- b) Argmax
- c) Softmax
- d) Argmin

Fordeler med logistisk regresjon

1. Kan trenes på små datasett (nyttig spesielt i norsk språkteknologi)
2. God generaliseringsevne (lite overtrening)
3. Rask & skalerbar
4. Forklarbar