21

Bazat e teknologjive - Big data

Fokusim në machine learning

1. Hyrje në Big data
2. Big Data Analytics
3. Hyrje në "Machine Learning"
4. Unsupervised - Supervised Learning
5. K - Means
6. K - Nearest Neighbor (k-NN)
7. Regression - për klasifikime dhe prediction
8. Decision Trees
9. Naive Byes classifier (punon me probabilitet)
10. Neuroni artificial
11. Discrete Perception, Perceptron Learning Rule
12. MLP (Multi Level Perception)
13. Gradient Descent Algorithm

Vlerësimi:
- 100% - Provimi final — teori (përgjithshme), Detyra (5-6)
- 30% - Projekti (opsionale)
- ±10% - Vijueshmëria

DeepLearning - generative AI
Machine learning - duhet vet mi nxerr vetitë
Soft Computing

CamScanner ile tarandı

- Unsupervised: me trajnu makinën pa mësëdhënie (si kallxojna gysh me kon output)
Se dino rezultatin e dukur
Punëm me përdor për Quantizim me reduktu numrin e pixels.

- Supervised: me kallxu edhe rezultatin
  -
  - Prediction

problem: feature extracture

Volume (volumi): TB, PB, ZB, ...
Variety (llojë): Structured (database, RMS)
Semi-Structured (comma seperator) (CSV file)
Unstructured (audio, video)

Velocity (shpejtësia): Data Streams, Real Time

Veracity (cilësia): Quality of Data

Value

Big Data Sources:

1. E-Government
2. Companies Data
3. Social media
4. IoT
5. E-health
6. E-baking
7. E-commerce
8. Video
9. Audio

CamScanner ile tarandı

Machine Learning

Supervised Learning: Classification, Prediction

Gjithmonë dalja duhet të jepet tek makinat

X	y
3	12
5	20
6	24
8	32
9	36

modeli

$\boxed{y = 4x}$ → Estimation

→ Prediction

kur kemi

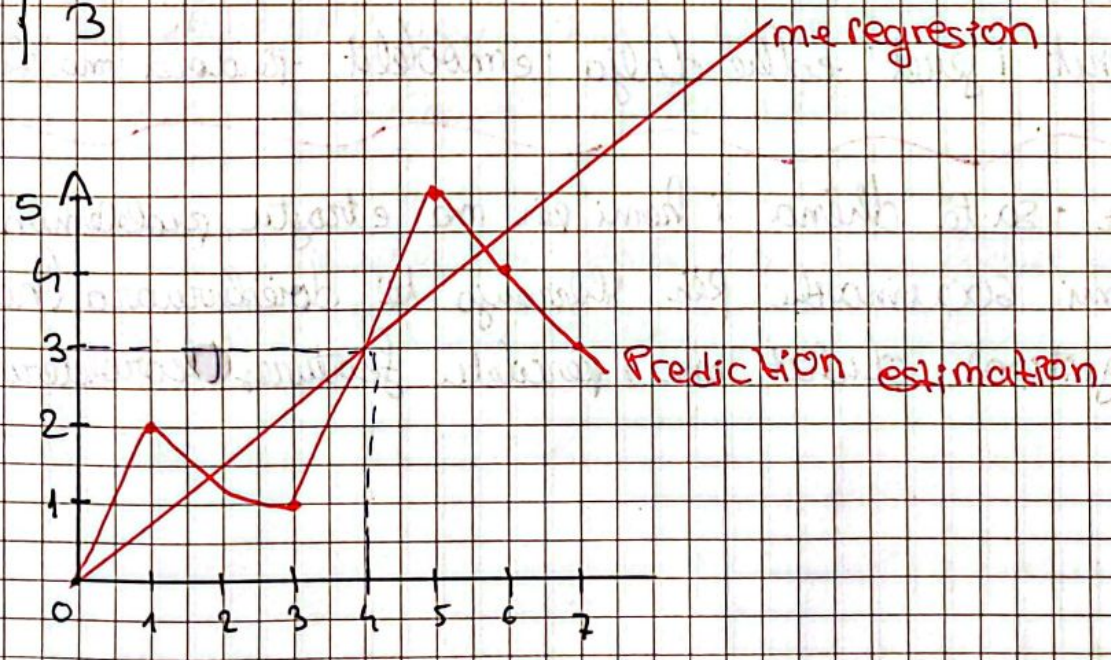
$x = 2.5$

atëher $y = 10$

Shëq estimation

X	y
1	2
3	1
5	5
6	4
7	3

→ Përdoret Machine Learning për me minimizu gabimin



CamScanner ile tarandı

Ligjerata 2 Unsupervised, Supervised Learning

V- Volume

V- Variety (të strukturuara, të pa strukturuara (images videos ox te spitali me word), gjysëm të strukturuara (me excel te spitali))

V- Velocity (Shpejtësia e reagimit kur bojna grumbullim të të dhënave psh real time, offline, data streams (duhet drejt për drejt mi procesu))

V- Veracity (sa janë të kompletume dhe të besueshme)

V- Value (duhet me diẗë sa vlerë ka edhëna dhe sa navjn gajo e dhënë)

• Për me bo analiza të këtyre Big dataave duhet me i përdor teknikat e machine Learning.

Unsupervised: Clustering (Pa mbështetje) - për të grupuar të dhëna, për quantizim.

• Kur dalja nuk dihet e përdorim unsupervised learning

• prej të dhënave në input tenton mi gjet njëjashmërite dhe në bazë të tyre i grumbullon (na duhet me i bërsu tani)

Supervised: Classification, Prediction (Me mbështetje)

algoritmit i jepet edhe dalja e modelit - përdoret ma shumë.

Dataset - sa të dhëna i kemi për me e trajtu problemin psh kemi 600 imazhe për shkronja të dorëshkruara (handwri en recognition) duhet me i përcaktu features (karakteristikat)

CamScanner ile tarandı

640x980

1 [A] $f_{1,1} \dots f_{1,10}$

⋮

600 [A] $f_{600,1} \dots f_{600,10}$

601 [B] $f_{601,1} \dots f_{601,10}$

⋮

1000 [B] $f_{1000,1} \dots f_{1000,10}$

- Pixelat nuk kanë vlerë dhe s'munëm me i jep si input.
- Si features te deep learning, algoritmi i mson vet automatikisht, kurse te machine learning duhet nja coktu vet
- psh, këtu mostravën e pixels
- $f_{1,1}$ - mostra 1 feature e 1
  $f_{600,1}$ - mostra 600 feature e 1
- Imazhet sjepin informat por si input ja jepim features

- Pjesëmarrësi i Supervised learning osht që përpos inputit edhe Daljo i jepet algoritmit.

D (dalja)
A
⋮
A
B
⋮
B

- Kurse te Unsupervised learning nuk e dinë (nuk jepet) dalja, vetëm në bazë të features që jepen si data set (input) bohen grupime
- Nëse punojmë me pixel problemi bëhet shumë dimensional e me features e reduktojmë

~~X 457,8 - mostra e 457~~
~~6 features e 78~~

Shembull. Për diabetin, prej dataset i mxerrim të dhënat (features) që na krijojn punë. Si psh traslëgimia, moshë analizat, psh i kemi 6 features që munën me e përcaktu a e ka diabetin.

- Klasifikimi për 400 pacient. Psh 150 e kanë diabetin dhe me S.L e jepim me 1 kur e ka me 0 kur s'ka pra ja japim daljen. Kurse nëse e bëjna me U.L nuk i japim dalje por në bazë të njajshmërisë mundet me bë grupim.

	D
$f_{1,1} \dots f_{1,6}$	1
⋮	⋮
$f_{150,1} \dots f_{150,6}$	1
$f_{151,1} \dots f_{151,6}$	0
⋮	⋮
$f_{400,1} \dots f_{400,6}$	0

- Duhet me pas kujdes sa mostra sa features
- DMTH osht problem 2 dimensional

D X y ϵ

$\boxed{i\text{-mostra}}$

D_i - Dalja e dëshiruar

y_i - Dalja aktuale e modelit

$X_{i,j}$ — X_i - mostra në hyrje (që random te algoritmi)
— X_j - features në hyrje

ϵ - Error / gabimi

$\epsilon = y_i - D_i$ - kur dalja aktuale osht mbrapsht prej daljes së dëshiruar

(Nëse del 0 d.m.th krejt jen dal mirë 100%.)

$X_{457,8}$ - mostra e 457
- feature e 8

te shembulli paraprak i kemi 400 mostra dhe 6 features prej:
- i prej 1-400
- j prej 1-6

X - bëhet krejt dataseti (input seti) me i ej.

$$X = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

me matricë
400 rresht
6 kolona

Për: $X_{457,8}$, D_{457} (dalja e dëshiruar), Y_{457} (dalja për të gjitha features) dhe $E_{457} = Y_{457} - D_{457}$ (errori)

Train set 70-80% Test set 20-30%

Input seti ndahet në 2 pjesë dhe gjithmonë më shumë për pjesën e trajnimit sepse osht mia mirë

Ndarja bëhet rastësisht gjë me e testu saktësinë e algoritmit psh për 1000 i kemi 800 për train dhe 200 për test.

psh nëse pas testimit 100 jon me sukses kurse 100 ja (për 200 duth) atëherë mënkuptohet se saktësia e algoritmit eshtë 50%.

CamScanner ile tarandı

Confusion Matrix (për supervised)

TP ✓ TN ✗
FP ✗ FN ✓

} Për me e vlerësu sa mirë e kon bë klasifikimin algoritmi
✗ - keq me pas trashë
✓ - tentojmë me e rrit (maksimizu)

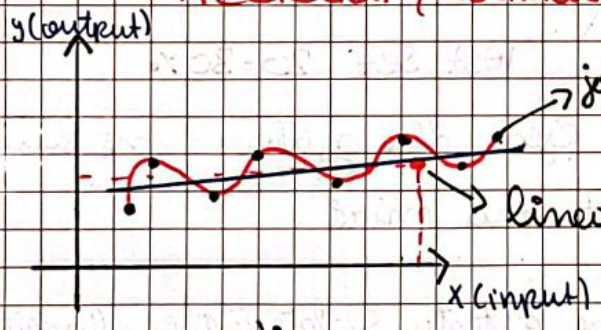
TP - dalja A - algo. e ka klasifiku si A

TN - dalja A - su dal A

FN - dalja so A - su dal A

FP - dalja so A - u dal A

Prediction, Estimation



jolinear - osht ma zor me e procesu (galeimi minimizohet dhe prediction o ma i saktë)

Estimation
$y=4x$

x	y
2	8
4	16
5	20
6	24
8	32

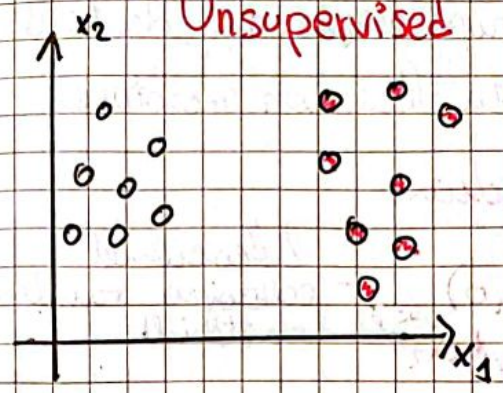
Estimation - me gjetë modelin për me minimizu galeimin

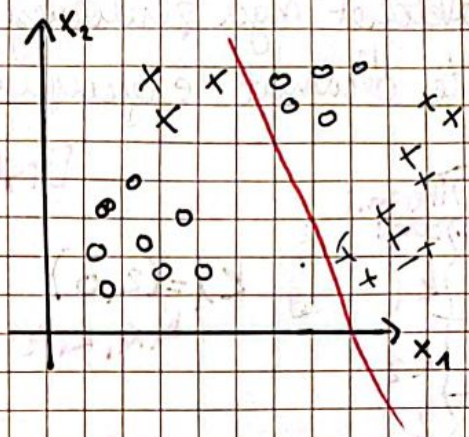
Prediction - për hyrje të cilat nuk i kemi përdorë në estimation, dhe e japim një dalje për to osht prediction
psh për $X=7 \Rightarrow y=28$ po parashikojmë modelin.

- Galeimi o me gjet një pjerrtësi dhe me i minimizu galeimet

Clustering (Unsupervised Learning)

(pa mbështetje)
Unsupervised





- Për me gjet njëjashmëri përdoret distanca

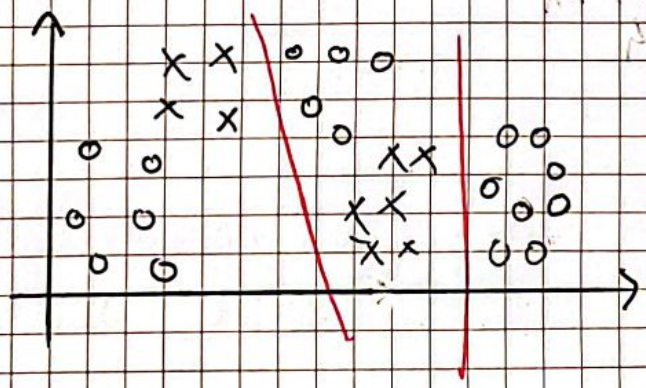
* Sa ma afër jam të njëjashme *

Distanca euklidiane përdoret

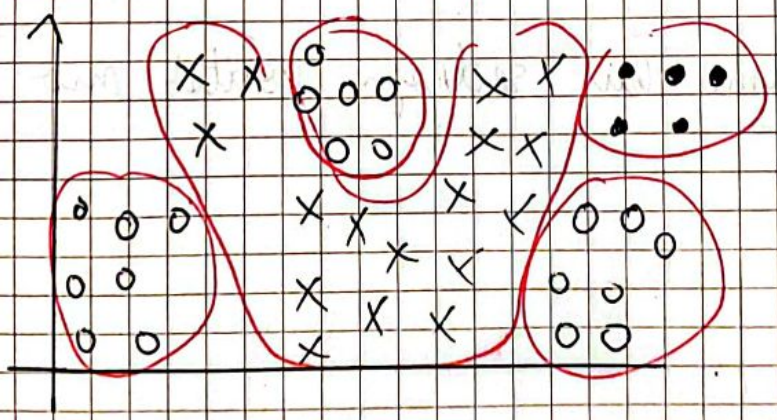
X_{ij} $\quad p=15$
$\quad j=2$ (2 dimensional)

$X_{1,1} \dots X_{1,2}$

Te supervised dalja por dihet, kur bëjmë trajnim testojmë me gjet një drejtëz që i dan klasifikimet.



Linearly separable



Jo lineare

CamScanner ile tarandı

K-Means

K- Perfaktohet nga perdoruesi (në sa grupe dëshiron ni da të dhënat)

Means - të dhënat e grupuara gendra llogaritem mesataren

Distanca Euklidiane

X_1	X_2	Iterimi i parë
1	2	k_1
3	3	k_2
4	3	k_2
-1	3	k_1
-3	2	k_1
-2	-1	k_1

$k_1=(2,0)$ $k_2=(3,0)$ – lokacionet i caktojmë random vërtetësisht

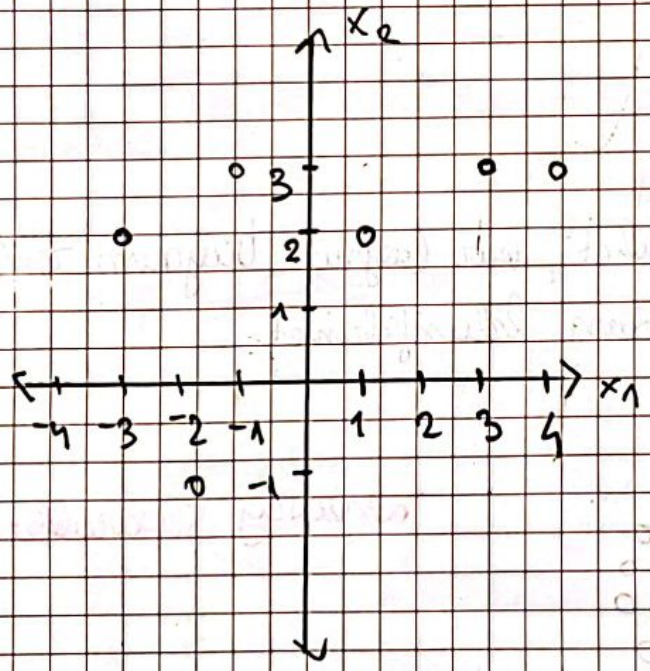
k_1x_1, k_1x_2 k_2x_1, k_2x_2

$$d_{ik_1}=\sqrt{(k_1x_1-x_{i1})^2+(k_1x_2-x_{i2})^2}$$

$$d_{ik_2}=\sqrt{(k_2x_1-x_{i1})^2+(k_2x_2-x_{i2})^2}$$

Për secilin hapje llogaritet distanca

$d_{ik_1} < d_{ik_2}$



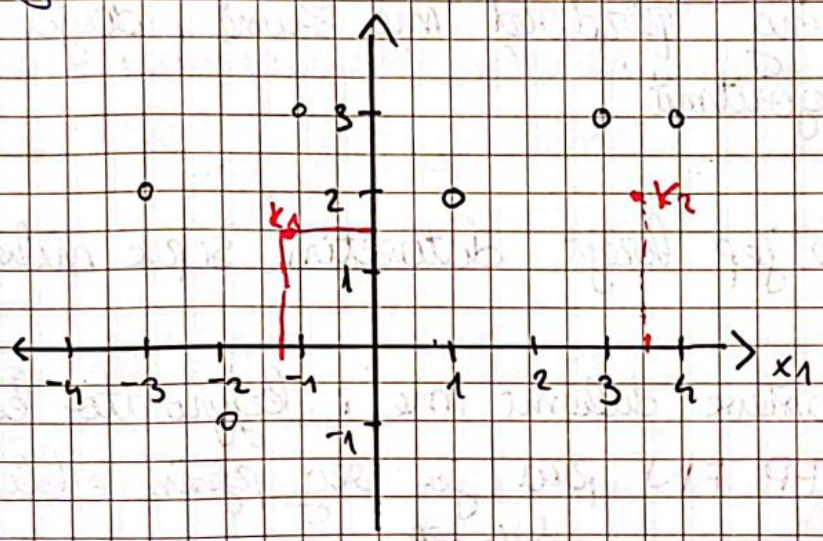
kur krylet algoritmi - kur sukseshëm poritet me

Tani e marrim vlerën e update-tur

$$k_{1x_1} = \frac{1+(-1)+(-3)+(-2)}{4} = -\frac{5}{4}$$

$$k_{2x_2} = \frac{2+3+2+(-1)}{4} = \frac{6}{4}$$

Pasi e gjetëm mesataren i zhvendosim qendrat



Për k_1: $d_{1k_1} = \sqrt{\left(-\frac{5}{4}-1\right)^2 + \left(\frac{6}{4}-2\right)^2}$

$d_{2k_2} = \sqrt{(3,5-1)^2 + (3-2)^2}$

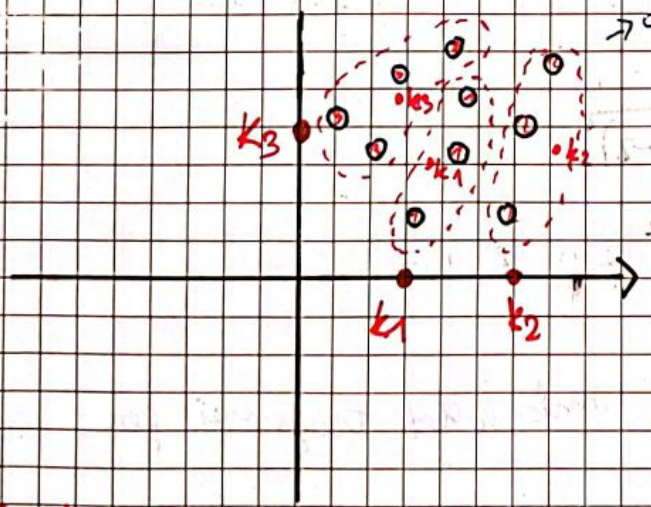
* nëse na lejon maj lyxi gjeli ink bëhet trajnimi por automatikisht grupohet.

Ligjerata 3 K-Means

- Unsupervised e përdorim kur nuk e dimë daljen dhe kur dojmë me grupu. Kur krylet algoritmi daljen me grupime duhem me analizu vet (përdoret për kuantizim të të dhënave)
me i reduktu vlerat

Supervised learning përdoret ma shumë, dalja këtu dihet, i jepet algoritmit

Shpesh direkt me jep krejt datasetin sepse nihet duhet me testu
Në shumicën e rasteve duhemi me i kqyr ato konfuzion matrix TP TN FP FN pra ja veq pjesën e klasifikimit



k-ja i jepet algoritmit
k=3
dmth në 3 grupe ka me munu me i bë me brate të distancës

Lokacionet k_1 k_2 k_3 percaktohen rastësisht

i = 1 ... 10 - mostrat (numullakat)
j = 2 - 2 dimensional

Llogaritet distanca e mostrave prej k_1, k_2, k_3
dhe distanca ma e shkurtë merret si target grup

- pas vërimit të parë gjejnë mesataret dhe i bajnë update lokacionet, mesmes vazhdon njëjtë deri sa nuk ka ndryshime

X_1	X_2	X_3	
2	3	4	k_2
-1	-3	0	k_1
1	5	-1	k_1
2	4	0	k_2
1	-1	-2	k_3

- në provojm lojn veç 1 iterim me lloj

$k=3$

d_{11}, d_{12}, d_{13}

$k_1 \Rightarrow (k_{11}, k_{12}, k_{13}) \Rightarrow (1, 1, 0)$

$k_2 \Rightarrow (k_{21}, k_{22}, k_{23}) \Rightarrow (1, 0, 1)$

$k_3 \Rightarrow (k_{31}, k_{32}, k_{33}) \Rightarrow (0, 1, 1)$

$i = 1..5$

$j = 3$ (3 dimensione)

Përcaktimi i lokacioneve fillestare o pak problem për algoritmin, shkaku i performancës

distanca euklidean

$$d_{11} = \sqrt{(k_{11}-X_{11})^2 + (k_{12}-X_{12})^2 + (k_{13}-X_{13})^2}$$

$$d_{12} = \sqrt{(k_{21}-X_{11})^2 + (k_{22}-X_{12})^2 + (k_{23}-X_{13})^2}$$

$$d_{13} = \sqrt{(k_{31}-X_{11})^2 + (k_{32}-X_{12})^2 + (k_{33}-X_{13})^2}$$

d_{11} - distanca mes mostrës së parë dhe lokacionit të parë k_1 (dimensioni)

X_{12} - mostra 1 feature 2

d_{12} - distanca mes mostrës 1 dhe lokacionit 2

i gjejmë 3 distancat për mostrën e parë dhe e marim minimalen (target grupin) Nëse $d_{12} < d_{11} < d_{13}$

mostra 1 i takon k_2

$$d_{21} = \sqrt{(k_{11}-X_{21})^2 + (k_{12}-X_{22})^2 + (k_{13}-X_{23})^2}$$

$$d_{22} = \sqrt{(k_{21}-X_{21})^2 + (k_{22}-X_{22})^2 + (k_{23}-X_{23})^2}$$

$$d_{23} = \sqrt{(k_{31}-X_{21})^2 + (k_{32}-X_{22})^2 + (k_{33}-X_{23})^2}$$

$d_{21} < d_{22} < d_{23}$

me 1.1

pra i takon k_1

$$d_{51} = \sqrt{(k_{11}-X_{51})^2 + (k_{12}-X_{52})^2 + (k_{13}-X_{53})^2}$$

$$d_{52} = \sqrt{(k_{21}-X_{51})^2 + (k_{22}-X_{52})^2 + (k_{23}-X_{53})^2}$$

$$d_{53} = \sqrt{(k_{31}-X_{51})^2 + (k_{32}-X_{52})^2 + (k_{33}-X_{53})^2}$$

* nëse dalin 2 distanca të njëjta e marrim cilën dojmë

$d_{53} < d_{52} < d_{51}$

CamScanner ile tarandı

Nasi që krylet iterimi i parë i ndrojna lokacionet i bojna update.

• Për k_1 e kqyrim kush e ka, e ka mostra 2 dhe 3

$$\hat{k_{11}} = \frac{x_{21}+x_{31}}{2}, \quad \hat{k_{12}} = \frac{x_{22}+x_{32}}{2}, \quad \hat{k_{13}} = \frac{x_{23}+x_{33}}{2}$$

$$k_1 = (k_{11}, k_{12}, k_{13}) \Rightarrow k_1 = (\hat{k_{11}}, \hat{k_{12}}, \hat{k_{13}})$$

• Për k_2 e kqyrim kush e ka, k_2 e ka mostra 1 dhe 4

$$\hat{k_{21}} = \frac{x_{11}+x_{41}}{2} \quad \hat{k_{22}} = \frac{x_{12}+x_{42}}{2}, \quad \hat{k_{23}} = \frac{x_{13}+x_{43}}{2}$$

$$k_2 = (k_{21}, k_{22}, k_{23}) \Rightarrow k_2 = (\hat{k_{21}}, \hat{k_{22}}, \hat{k_{23}})$$

• k_3 e ka veq x_5

$$\hat{k_{31}} = x_{51} \quad \hat{k_{32}} = x_{52} \quad \hat{k_{33}} = x_{53}$$

duhet shumë iterime me bë deri të përfundon algoritmi dhe distancat tani llogaritem me vlerat e grupeve të updatuara

Nëse ojen një mostër (lyrje) e re që nuk osht ndër në trajnim të algoritmit direkt shpërndahet psh $x_{101}, x_{102}, x_{103}$ dhe llogaritem distancat dhe gjendet minimalja dmth pa e trajnu direkt grupohet pa i update lokacionet (mundem edhe me i update por s'ka nevoj so jay mirë)

- Nëpërim: i jepmostrat lokacionet, ma i llogaritem, i iterim dhe i bojna tani update

lokacionet e updatuara	x_1	x_2	x_3	lokacionet fillestare
k_3	2	3	4	k_2
k_2	-1	-3	0	k_1
k_1	1	5	-1	k_1
k_1	2	4	0	k_2
k_1	1	-1	-2	k_3

$k_1=(k_{11},k_{12},k_{13})=(1,1,0)$

$k_2=(k_{21},k_{22},k_{23})=(1,0,1)$

$k_3=(k_{31},k_{32},k_{33})=(0,1,1)$

$d_{11}=\sqrt{(1-2)^2+(1-3)^2+(0-4)^2}=\sqrt{1+4+16}=\sqrt{21}$

$d_{12}=\sqrt{(1-2)^2+(0-3)^2+(1-4)^2}=\sqrt{1+9+9}=\sqrt{19}$

$d_{13}=\sqrt{(0-2)^2+(1-3)^2+(1-4)^2}=\sqrt{4+4+9}=\sqrt{17}$ → k_3 mostra 1:

$d_{21}=\sqrt{(1+1)^2+(1+3)^2+(0-0)^2}=\sqrt{20}$

$d_{22}=\sqrt{(1+1)^2+(0+3)^2+(1-0)^2}=\sqrt{14}$ → k_2

$d_{23}=\sqrt{(0+1)^2+(1+3)^2+(1-0)^2}=\sqrt{18}$

$d_{31}=\sqrt{17}$, $d_{32}=\sqrt{29}$, $d_{33}=\sqrt{21}$ → k_1

$d_{41}=\sqrt{10}$, $d_{42}=\sqrt{18}$, $d_{43}=\sqrt{13}$ → k_1

$d_{51}=\sqrt{8}$, $d_{52}=\sqrt{10}$, $d_{53}=\sqrt{14}$ → k_1

Tash i update-ojmë lokacionet

$k_{11}=\frac{1+2+1}{3}=\frac{4}{3}$ $k_{12}=\frac{5+4-1}{3}=\frac{8}{3}$ $k_{13}=\frac{-1+0-2}{3}=\frac{-3}{3}=-1$

$k_1=(k_{11},k_{12},k_{13})=(\frac{4}{3},\frac{8}{3},-1)$

$k_2=(-1,-3,0)$ $k_3=(2,3,4)$ (sepse s'kam veq ka një)

CamScanner ile tarandı

Perkaktimi i vlerave fillestare ka ndikim në performancën e algoritmit sepse psh me e gjit k_3 larg ——— edhe k_3 mund shumë iterimeve akoma nuk grupohet kurrgjë me to

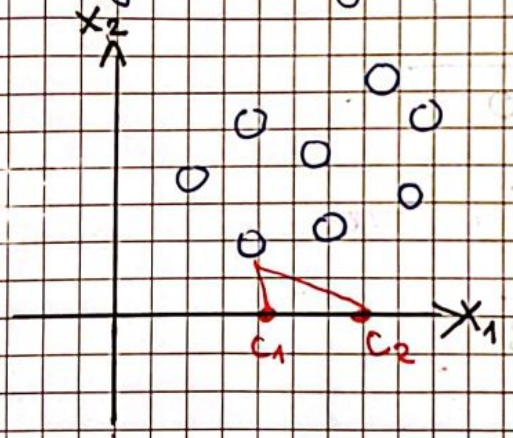
monet me lyp me vrrtatu

Kohonen Self Organizing Map (SOM)

- Edhe kjo osht te unsupervised learning, njejtësim si te k-means edhe kjo përdor distancat për me gjet njejtësi dhe me grupu.
- Pra edhe SOM grupon në bazë të njejtësmerisë
- bazohet në neuronet artificiale

• C - numri i grupeve

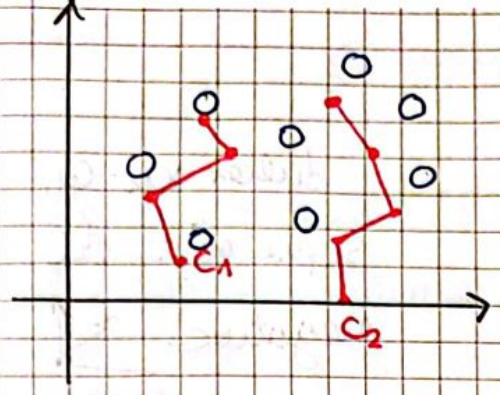
i caktojmë 2 grupe random si lokacione fillestar



• Kur fillon algoritmi kto 2 grupe bojn gara mes vete, ai gë fiton gara e bon update lokacionin dhe shkon për 1 hap ma larë

• por hapi duhet të jetë < 1 sepse nëse osht = 1 atëher bjen të hyjo (rrumullaku)

η - eta e përcakton sa e bon hapin grupi gë fiton garën



- Për të parën hyrje fiton C_1 dhe e lëen një hap drejt hyrjes pra. mendohet si të fiton updatohet
- pastaj për tjetrën hyrje krahasohet me C_1 të updatuar
- dhe kështu me radhë dhe pas një numri iterimesh stabilizohet

- Këtu rëndohet një funksion i cili, nëse jon 10 grupe dhe 1 e fiton garën i mundësonë edhe grupeve të ngjashme (afër) me to me lëë një hap por ma të vogël

X_1	X_2
1	2
-1	0
3	-2
-2	-1
4	5

3 grupe C_1 C_2 C_3

$C_1 = (C_{11}, C_{12})$
$C_2 = (C_{21}, C_{22})$
$C_3 = (C_{31}, C_{32})$

për secilën hyrje llogaritet distanca, cija ma e vogla fiton

$$d_{11} = \sqrt{(C_{11}-X_{11})^2 + (C_{12}-X_{12})^2}$$

$$d_{12} = \sqrt{(C_{21}-X_{11})^2 + (C_{22}-X_{12})^2}$$

$$d_{13} = \sqrt{(C_{31}-X_{11})^2 + (C_{32}-X_{12})^2}$$

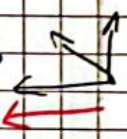
$d_{12} < d_{13} < d_{11}$ pra fiton C_2

CamScanner ile tarandı

C_2 mundet me update

$$C_2 \Rightarrow C_{21} = C_{21} + \eta(x_{11} - C_{21})$$
$$C_{22} = C_{22} + \eta(x_{12} - C_{22})$$

duhet $x_{11} - C_{21}$. Sepse nëse del negative:

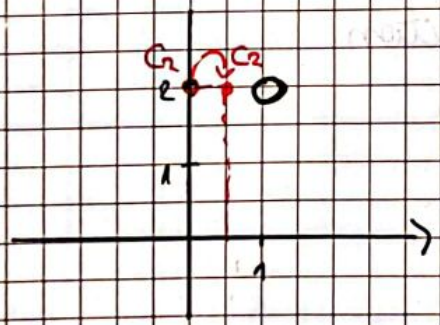


pozitive

$\eta = 0{,}3$ ose $0{,}2$

$$C_{21} = 0 + 0.3(1-0) = 0 + 0.3 = 0.3$$
$$C_{22} = 2 + 0.3 \cdot 0 = 2 \quad \Rightarrow \boxed{C_2 = (0.3,\ 2)}$$

• Nëse η osht ma e vogël duhet ma shumë iterime por osht ma e sakt, ge osht ma e madhe ma pak iterime por nuk shkon shumë në gendër



hyrja 1,2

Ligjerata 4

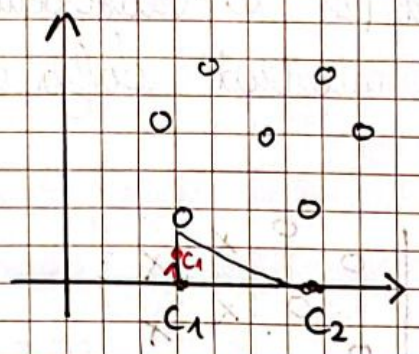
Kohonen SOM

1° K mu percaktu (lokacionet rastësisht)

2° Me kalkulu distancën për secilën mostër

3° lokacionet update-olen në bazë të mesatarave

k-means



x_1	x_2	x_3
2	-3	1
1	4	2
5	0	-1

dataset

d_{11} d_{12} d_{13} d_{21} d_{22} d_{23}

$$C_1 = (C_{11}, C_{12}, C_{13}) = (0, 1, 2)$$
$$C_2 = (C_{21}, C_{22}, C_{23}) = (-1, 0, 1)$$
$$C_3 = (C_{31}, C_{32}, C_{33}) = (2, 1, -3)$$

lokacionet përcaktohen rastësisht

Kur vjen mostra e parë kalkulohen d_{11} d_{12} d_{13}

d_{11} - distanca prej mostrës 1 te C_1

$$d_{11} = \sqrt{(C_{11} - x_{11})^2 + (C_{12} - x_{12})^2 + (C_{13} - x_{13})^2}$$

$$d_{12} = \sqrt{(C_{21} - x_{11})^2 + (C_{22} - x_{12})^2 + (C_{23} - x_{13})^2}$$

$$d_{13} = \sqrt{(C_{31} - x_{11})^2 + (C_{32} - x_{12})^2 + (C_{33} - x_{13})^2}$$

$d_{12} < d_{13} < d_{11}$

d.m.th C_2 ka fitu menilin e update-ës

η duhet mes 0 - 1

$$C_{21} = C_{21} + \eta(x_{11} - C_{21})$$
$$C_{22} = C_{22} + \eta(x_{12} - C_{22})$$
$$C_{23} = C_{23} + \eta(x_{13} - C_{23})$$

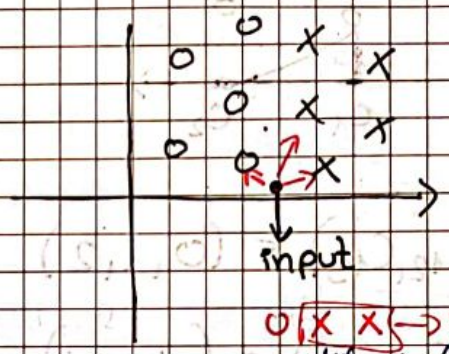
CamScanner ile tarandı

Fakt për secilën hyrje shkon njëjtë
(mund me ora në provim)

Supervised learning
K-NN (Nearest Neighbors)

- bazohet te distanca, përdoret për të klasifikuar ~~shërbt~~ të dhënat. Kur i kemi mostrat dalja dhe klasat dihen

X_1	X_2	X_3
3	2	1
-1	0	-2
1	3	5
4	2	1

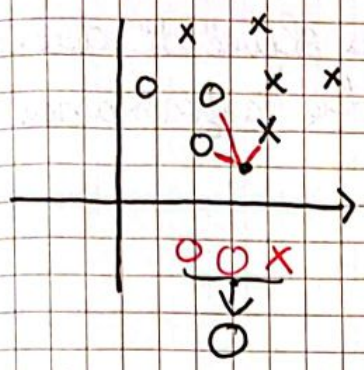


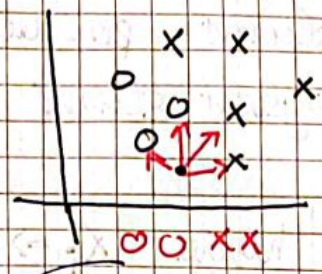
O X X → 2
atëherë klasifikohet si X

$k = 3, 5, 6, \ldots$

Duhet me e përcaktu k (k-sa klasa përdoren për me i klasifiku të dhënat) numri të k-ja e përcakton performancën e algoritmit

- kur vjen një hyrje kalkulohet distanca të të gjitha mostrat, pra kur e kemi $k=3$ i llogaritim 3 distancat më të shkurta për inputin
- Nëse $k=5$ i kalkulojm 5 distanca prandaj k ka ndikim në performancë





atëherë i kërkojmë distancat, nëse edhe distancat janë të njëjta atëherë e marrim vlerësimin njëjtë

• te K-NN nuk përdorim iterimet, nuk ndrojmë vlerat por vetëm bëhet klasifikimi.

X_1	X_2	X_3	D
3	2	1	A
-1	0	-2	B
1	3	5	A
4	2	1	B

Input: $X_{5j} = (2, 1, 0)$

• $k=2$

• $k=3$

d_{15} - prej mostrës 1 deri te kërkojë 5

$$d_{15} = \sqrt{(x_{11}-x_{51})^2 + (x_{12}-x_{52})^2 + (x_{13}-x_{53})^2}$$

$$d_{25} = \sqrt{(x_{21}-x_{51})^2 + (x_{22}-x_{52})^2 + (x_{23}-x_{53})^2}$$

$$d_{35} = \sqrt{(x_{31}-x_{51})^2 + (x_{32}-x_{52})^2 + (x_{33}-x_{53})^2}$$

$$d_{45} = \sqrt{(x_{41}-x_{51})^2 + (x_{42}-x_{52})^2 + (x_{43}-x_{53})^2}$$

$$d_{25} < d_{45} < d_{35} < d_{15}$$

$k=2$ merret klasa B

$k=3$ edhe ktu klasifikohet si B | $\boxed{BBA}$

ose nëse kemi $d_{25} < d_{35} < d_{45} < d_{15}$

për $k=2$ klasifikohet si B sepse d_{25} ka min e vogla

$k=3$ ngap si B

KNN përdoret shumë por vetëm duke ka shumë të dhëna duhet mu llogarit shumë distanca dhe performanca lejen tani

X_1	X_2	X_3	D
0	1	-1	A
2	3	-2	C
4	1	5	B
3	2	1	C
4	5	-1	A
-1	-2	0	B

hyrja $X_i \Rightarrow (-1, 2, 0)$

$k = 3, 4, 5, 2$

$d_{1i} = \sqrt{(0+1)^2 + (1-2)^2 + (-1-0)^2} = \sqrt{3}$

$d_{2i} = \sqrt{(2+1)^2 + (3-2)^2 + (-2-0)^2} = \sqrt{14}$

$d_{3i} = \sqrt{(4+1)^2 + (1-2)^2 + (5-0)^2} = \sqrt{51}$

$d_{4i} = \sqrt{(3+1)^2 + (2-2)^2 + (1-0)^2} = \sqrt{17}$

$d_{5i} = \sqrt{(4+1)^2 + (5-2)^2 + (-1-0)^2} = \sqrt{35}$

$d_{6i} = \sqrt{(-1+1)^2 + (-2-2)^2 + (0-0)^2} = \sqrt{16}$

$d_{1i} < d_{2i} < d_{6i} < d_{4i} < d_{5i} < d_{3i}$

$k=3$ A C B - klasifikohet në A

$k=4$ A C B C - klasifikohet në C

$k=5$ A C B C A - i kalkulojm totalet

$A = \sqrt{3} + \sqrt{35}$

$C = \sqrt{14} + \sqrt{17}$

$C < A$ atëherë klasifikohet në C

$k=2$ A C - klasifikohet në A

CamScanner ile tarandı

Pytje teori: supervised VS unsupervised
Si bëhet testimi i algo
të spjegohet algoritmi

Byes Theorem → supervised learning
(naive Byes Classifier)

$P(A) \Rightarrow 0, \dots 1$ (probabiliteti për një ngjarje A

$P(B) \Rightarrow 0, \dots 1$ $1 \to 100\%$ $0 \to 0\%$ Sea shancë me ndodh një ngjarje

psh për coin osht 0,5

Me probabilitet i klasifikojm ngjarjet me kushte

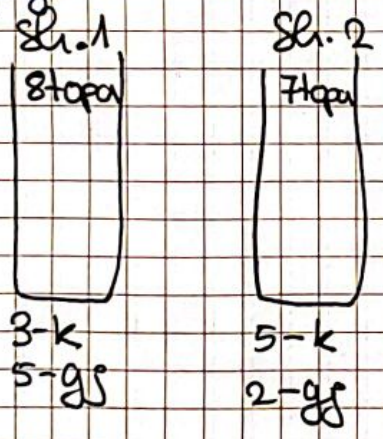
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Condition Probability (Probabiliteti me kushte)

ngjarja gë vërtet nga ngjarja B

e dim gë kandodh

Nëse ka ndodh B sa janë gjasat me ndodh A-ja



•psh te Sh.1 sa janë gjasat me nxorr top të kaltër

$P(A) \Rightarrow$ Sh.1

$P(B) \Rightarrow$ Sh.2

$P(C) \Rightarrow$ sa janë gjasat me zgjedh top të kaltër

$$P(C) = P(A) \cdot P(C|A) + P(B) \cdot P(C|B)$$

$$= \underbrace{\frac{1}{2}}_{0.5} \cdot \frac{3}{8} + \frac{1}{2} \cdot \frac{5}{7} = \frac{3}{16} + \frac{5}{14} = \boxed{0{,}54}$$

CamScanner ile tarandı

Formulë që i lidh probabilitetet

$$\boxed{P(A|B) = \frac{(A \cap B)}{P(B)}} \Rightarrow A \cap B = P(A|B) \cdot P(B)$$

$$\boxed{P(B|A) = \frac{(A \cap B)}{P(A)}} \Rightarrow A \cap B = P(B|A) \cdot P(A)$$

$$\boxed{P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}}$$

$$\boxed{P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}}$$

Kjo përdoret te supervised learning nëse psh e klasifikojmë A ? kalkulojmë gjasat pastaj e çë B gjajë që ka gjasat ma shumë klasifikohet

Ligjërata 5

Bayes Theorem - Naive Bayes Classifier

Probabiliteti me kusht

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

→ kur e dimë që ka ndodhë B-ja si është probabiliteti me ndodhë njëjtë A

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

Shembull

0.8% kancer (e populacionit janë me kancer)

99.2% jo-kancer (i rejstrojnë me 100)

e kemi bë një test

0.8% ⟶ 98% +

99.2% ⟶ 97% −

$P(\text{kancer}|+) = 0{,}008$ $P(\neg\text{kancer}) = 0{,}992$

$P(+|\text{kancer}) = 0{,}98$ $P(+|\neg\text{kancer}) = 0{,}03$

$P(-|\text{kancer}) = 0{,}02$ $P(-|\neg\text{kancer}) = 0{,}97$

→

CamScanner ile tarandı

Një pacient e ka bë një test i ka dal + tash duhet me kalkulu a ka kancer a jo

$$P(\text{kancer}|+) = \frac{P(+|\text{kancer}) \cdot P(\text{kancer})}{P(+)} = \frac{0{,}98 \cdot 0{,}008}{0{,}0376} = \boxed{0{,}208}$$

$$P(\neg\text{kancer}|+) = \frac{P(+|\neg\text{kancer}) \cdot P(\neg\text{kancer})}{P(+)} = \frac{0{,}03 \cdot 0{,}992}{0{,}376} = \boxed{0{,}791}$$

Pra pacienti s'ka kancer (sepse te 99,2 3% jon + kurse te ~~98~~ 0,8 jon 98% kjo ma shum popullsi)

$$P(+) = P(\text{kancer})P(+|\text{kancer}) + P(\neg\text{kancer}) \cdot P(+|\neg\text{kancer})$$
$$= 0{,}008 \cdot 0{,}98 + 0{,}992 \cdot 0{,}03 = \boxed{0{,}0376}$$

kur punon me probabilitetet nëse sot del > 1 dije gë ka bë gabim.

CamScanner ile tarandı

Naive Bayes Classifier

$$S = \arg\max \left(P(s_j) \cdot \prod_{i=1}^{N} P(v_i | s_j) \right)$$

target klas

$P(s_j)$ → gjasat mu klasifiku si një klasë

M - duhet me kalkulu për secilën feature në cilën klasë

V_1	V_2	V_3	V_4	S
Po	Jo	Jo	Po	A
Jo	Po	Jo	Jo	B
Po	Po	Jo	Po	B
Po	Jo	Jo	Jo	B
Jo	Jo	Jo	Jo	A
Po	Po	Po	Po	A
Po	Jo	Po	Po	A
Jo	Po	Po	Po	A

→ supozojm klasat

mostra 1 4 features secila kan 2 fleje rëregjëgjjeve Po ose Jo

Nëse gjen një lyrje (Jo, Jo, Jo, Po.)

Mi gjet gjasat që me taku A dhe me taku B

klasifikim si A $\frac{5}{8}$ herë s'A ka

$$P(S=A) \cdot P(V_1, V_2, V_3, V_4 | S=A) =$$

sa o prob. që $V_4 = Po$ nëse klasa osht A

$$= \frac{5}{8} \cdot P(V_1|A) \cdot P(V_2|A) \cdot P(V_3|A) \cdot P(V_4|A)$$

$$= \frac{5}{8} \cdot \frac{2}{5} \cdot \frac{3}{5} \cdot \frac{2}{5} \cdot \frac{4}{5} = \frac{240}{5000} = \boxed{0{,}048}$$

$$P(S=B) \cdot P(V_1, V_2, V_3, V_4 | S=B) = \frac{3}{8} \cdot P(V_1|B) \cdot P(V_2|B) \cdot P(V_3|B) \cdot P(V_4|B)$$

$$= \frac{3}{8} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot 1 \cdot 0 = \boxed{0}$$

So mirë me dal 0 e mbrojmë një Jo te rreshti 3 (në provim e gjen mirë ax nëse del 0 mbi mbrojmë)

→

CamScanner ile tarandı

V_1	V_2	V_3	V_4	S
Po	Jo	Jo	Po	A
Jo	Po	Jo	Jo	B
Po	Po	Jo	Po	B
Po	Jo	Po	Jo	B
Jo	Jo	Jo	Jo	A
Po	Po	Po	Po	A
Po	Jo	Po	Po	A
Jo	Po	Po	Po	A

tosh:

$$P(S=B)\cdot P(V_1|B)\cdot P(V_2|B)\cdot P(V_3|B)\cdot P(V_4|B)=$$

$$=\frac{3}{8}\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot 1\cdot\frac{1}{3}=\boxed{0{,}013}$$

Për A probabiliteti $\boxed{0{,}048}$

Për B probabiliteti 0,013

d.m.th hyrja (Jo, Jo, Jo, Po) => A

klasifikohet si A

(Po, Po, Po, Jo) => B

$$P(S=A)\cdot P(V_1|A)\cdot P(V_2|A)\cdot P(V_3|A)\cdot P(V_4|A)=$$

$$=\frac{5}{8}\cdot\frac{3}{5}\cdot\frac{2}{3}\cdot\frac{3}{5}\cdot\frac{1}{5}=\boxed{0{,}012}$$

$$P(S=B)\cdot P(V_1|B)\cdot P(V_2|B)\cdot P(V_3|B)\cdot P(V_4|B)=$$

$$=\frac{3}{8}\cdot\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{2}{3}=\boxed{0{,}037}$$ pra klasifikohet si B

Për të dhëna numerike

$$P(v)=\frac{1}{\sqrt{2\pi\cdot\sigma^2}}\cdot e^{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2}$$

$\mu \rightarrow$ Mesatarja

$\sigma \rightarrow$ Devijimi standard

$$\sigma=\sqrt{\frac{\sum(v_i-\mu)^2}{(N-1)}}$$

Kur kemi të dhëna vlera numerike maspari nevojitet me e gjet mesatarën

$\sigma \rightarrow$ na jep një matje, sa i kemi të dhënat larg prej mesatares

CamScanner ile tarandı

Psh për 2 student

$\frac{50+50}{2} = 50$
$\frac{100+0}{2} = 50$ } dmth mesatarja s'po tregon shumë info

kurse psh $\frac{100+0}{2} = 50$ Krap mesatarja njëjtë

$\frac{20+80}{2} = 50$ por devijimi standard σ osht më i madh te 100+0

• Varianca është katrori i devijimit standard

Devijimi standard: $\sigma = \sqrt{\frac{\Sigma (v_i - \mu)^2}{(N-1)}}$

Variance: $\sigma^2 = \frac{\Sigma (v_i - M)^2}{(N-1)}$

* Nëse e kena për krejt dataset atëher e marrim vetëm /N jo N-1

V_1	V_2	V_3	V_4	V_5	S
Po	Jo	Jo	Po	37	A
Jo	Po	Jo	Jo	42	B
Po	Po	Jo	Po	45	B
Po	Jo	Po	Jo	48	B
Jo	Jo	Jo	Jo	32	A
Po	Po	Po	Po	35	A
Po	Jo	Po	Po	40	A
Jo	Po	Po	Po	41	A

$\boxed{A}$ → N=5 se i kena 5 A-ja

• $\mu = \frac{37+32+35+40+41}{5} = \boxed{37}$

• $\sigma^2 = \frac{(37-37)^2 + (32-37)^2 + (35-37)^2 + (40-37)^2 + (41-37)^2}{N-1}$

$\sigma^2 = \frac{54}{4} = \boxed{13,5}$

• $\sigma = \sqrt{\sigma^2} = \sqrt{13,5} = \boxed{3,67}$

$P(V_5 = 46 \mid S = A) = \frac{1}{\sqrt{2\pi \cdot 13,5}} \cdot e^{-\frac{1}{2}\left(\frac{46-37}{3,67}\right)^2} = \boxed{0,005}$

Për (Jo, Jo, Jo, Po, 46)

CamScanner ile tarandı

$\boxed{S=B}$

$\downarrow$

$N=3$

- $\mu = \frac{42+45+48}{3} = \frac{135}{3} = \boxed{45}$

- $\sigma^2 = \frac{(42-45)^2+(45-45)^2+(48-45)^2}{2} = \frac{(-3)^2+0^2+3^2}{2} =$

$= \frac{9+9}{2} = \frac{18}{2} = \boxed{9}$

- $\sigma = \sqrt{\sigma^2} = \sqrt{9} = \boxed{3}$

$$P(V_5=46 \mid S=B) = \frac{1}{\sqrt{2\pi \cdot 9}} \cdot e^{-\frac{1}{2}\left(\frac{46-45}{3}\right)^2} = \boxed{0{,}125}$$

Domth $(Jo, Jo, Jo, Po, 46) \Rightarrow B$ (kjo është edhe më logjike sepse te B kërcet vlerat janë më afër 46)

- $P(S=A) \cdot P(V_1|A) \cdot P(V_2|A) \cdot P(V_3|A) \cdot P(V_4|A) \cdot P(V_5|A) =$

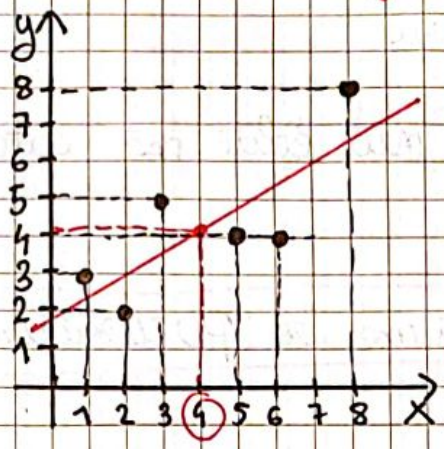
$= \frac{5}{8} \cdot \frac{3}{5} \cdot \frac{2}{3} \cdot \frac{3}{5} \cdot \frac{1}{5} \cdot 0{,}005 = 0{,}012 \cdot 0{,}005 = \boxed{0{,}00006}$

$P(S=B) \cdot P(V_1|B) \cdot P(V_2|B) \cdot P(V_3|B) \cdot P(V_4|B) \cdot P(V_5|B) =$

$= \frac{3}{8} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot 0{,}125 = 0{,}037 \cdot 0{,}125 = \boxed{0{,}0046}$

CamScanner ile tarandı

Leksionata 6 Regression



→ në këtë form ka gabime por minimizohen

KNN dhe Bayes përdoren vetëm për klasifikime, ojen një hyrje dhe e mson ku me klasifiku (te Bayes me probabilitet

Kurse me Regression mundemi edhe me parashiku edhe me klasifiku.

x	y
1	3
2	2
3	5
5	4
6	4
8	8

x - hyrja y - dalja e dëshirvar

- E kemi përmend Estimation ku krijojna një model për të cilin për të gjitha vlerat e x-it e fitojmë vlerën e y-it, kurse për vlerat e x që nuk i kemi në dataset bëjmë Prediction. psh

x	y
2	4
5	10
6	12

⇒ $\boxed{y=2x}$ Estimation

⇓

- Por këtë vrast nuk mundemi me Estimation sepse nuk mundemi me gjet një gë për çdo x i get vlerat e y, andaj ktu e përdorim Regressionin ku mundohemi me e gjet një drejtëz për me minimizu katrorin e gabimeve

$$y = w_0 + w_1 \cdot x$$ → zgjidhja gjenerale

w_0 - Constant (sa duhet të zhvendoset drejtëza në formë vertikale)

w_1 - Slope (rrjedhësia e drejtëzës optimale me të cilën minimizohet gabimi te të gjitha mostrat ekzistuese)

CamScanner ile tarandı

psh te 1 ekemi 3, kurse me drejtëz del 2,5 pra o gabim e na gata tentojm me e minimizu

- Nëse vjen një hyrje për 4 gë nuk është në dataset e bojna Paraslikim

Prandaj me Regression mundemi me bo paraslikim edhe klasifikim

O LEAST SQUARES

$$W_1 = \frac{N \cdot \Sigma x \cdot y - \Sigma x \cdot \Sigma y}{N \cdot \Sigma x^2 - (\Sigma x)^2}$$

$$W_0 = \frac{\Sigma y - W_1 \cdot \Sigma x}{N}$$

→ ky ekuacion përdoret për me fitu pjerrësinë e drejtëzës astu gë me u zvogëlu kutrori i gabi

- pra qëllimi është me minimizu gabimin

$\Sigma x \cdot y = 3+4+15+20+24+64 = \boxed{130}$

$\Sigma x = 1+2+3+5+6+8 = \boxed{25}$

$\Sigma y = 3+2+5+4+4+8 = \boxed{26}$

$\Sigma x^2 = 1+4+9+25+36+64 = \boxed{139}$

$(\Sigma x)^2 = 25^2 = \boxed{625}$

$$W_1 = \frac{6 \cdot 130 - 25 \cdot 26}{6 \cdot 139 - 625} = \frac{780 - 650}{834 - 625} = \frac{130}{209} = \boxed{0,62}$$

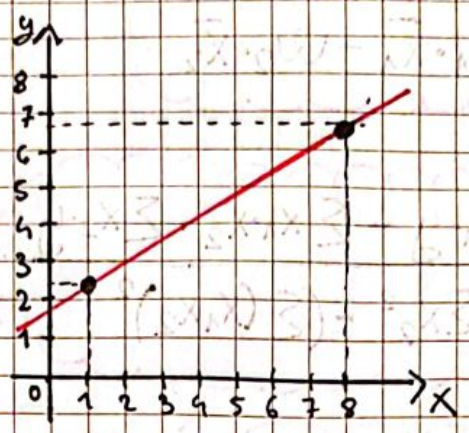
$$W_0 = \frac{26 - 0,62 \cdot 25}{6} = \frac{26 - 15,5}{6} = \frac{10,5}{6} = \boxed{1,75}$$

CamScanner ile tarandı

Pjerrësia poritive → drejtëza shkon tu u rritë
Pjerrësia negative → drejtëza shkon tu u zvogëlu

$$y = w_0 + w_1 \cdot x$$

$$\boxed{y = 1{,}75 + 0{,}62 \cdot x}$$



kur i bashkojmë pikat e fitojmë një drejtëz që është më e rëndësi për krejt pikat

$\boxed{x = 1}$

$y = 1{,}75 + 0{,}62 \cdot 1 = \boxed{2{,}37}$

$\boxed{x = 8}$

$y = 1{,}75 + 0{,}62 \cdot 8 = \boxed{6{,}71}$

Nëse bëjm psh një hyrje 20 e gjesim në formulë dhe e parashikojna daljen për to

Multiple Regression

- problemi këtu ndryshon sepse mundemi me pas shumë features

x_1	x_2	y
2	3	3
1	4	1
5	2	2
4	1	4
1	5	6
3	4	3

- pra këtu me 2 parametra munohemi me bo parashikim
  psh x_1 - leitet, x_2 - investimet, y - shitjet
- Nëqë problemi është 2 dimensional duhet me gjet sipërfaqe që i afrohet krejt mostrave.
- Gjithmonë duke tentu me minimizu katrorin e gabimeve

$$\boxed{y = w_0 + w_1 x_1 + w_2 x_2}$$

CamScanner ile tarandı

$$y = W_0 + W_1 \cdot X_1 + W_2 \cdot X_2$$

$$W_0 = \bar{y} - W_1 \cdot \bar{X}_1 - W_2 \cdot \bar{X}_2$$

$$W_1 = \frac{\sum x^2 \cdot \sum x_1 \cdot y - \sum x_1 \cdot x_2 \cdot \sum x_2 \cdot y}{\sum x_1^2 \cdot \sum x_2^2 - (\sum x_1 x_2)^2}$$

$$W_2 = \frac{\sum x_1^2 \cdot \sum x_2 \cdot y - \sum x_1 \cdot x_2 \cdot \sum x_1 \cdot y}{\sum x_1^2 \cdot \sum x_2^2 - (\sum x_1 \cdot x_2)^2}$$

Nga tabela e jepur:

$\sum x_2^2 = 9+16+4+1+25+16 = \boxed{71}$

$\sum x_1 \cdot y = 6+1+10+16+6+9 = \boxed{48}$

$\sum x_1 \cdot x_2 = 6+4+10+4+5+12 = \boxed{41}$

$\sum x_2 \cdot y = 9+4+4+4+30+12 = \boxed{63}$

mesatarët

$\bar{y} = \frac{3+1+2+4+6+3}{6} = \frac{19}{6} = \boxed{3,16}$

$\bar{X}_1 = \frac{2+1+5+4+1+3}{6} = \frac{16}{6} = \boxed{2,66}$

$\bar{X}_2 = \frac{3+4+2+1+5+4}{6} = \frac{19}{6} = \boxed{3,16}$

$(\sum x_1 \cdot x_2)^2 = (41)^2 = \boxed{1681}$

$\sum x_1^2 = 4+1+25+16+1+9 = \boxed{56}$

$$W_1 = \frac{71 \cdot 48 - 41 \cdot 63}{56 \cdot 71 - 1681} = \frac{3408 - 2583}{3976 - 1681} = \frac{825}{2295} = \boxed{0,35}$$

$$W_2 = \frac{56 \cdot 63 - 41 \cdot 48}{56 \cdot 71 - 1681} = \frac{3528 - 1968}{3976 - 1681} = \frac{1560}{2295} = \boxed{0,67}$$

$$W_0 = 3,16 - (0,35 \cdot 2,66) - (0,68 \cdot 3,16) = \boxed{0,11}$$

$$y = 0,11 + 0,35 \cdot X_1 + 0,67 \cdot X_2$$ → ky ekuacion përdoret për paradikim optimal

CamScanner ile tarandı

$x_1 = 7 \qquad x_2 = 10$

$$y = 0{,}11 + 0{,}35 \cdot x_1 + 0{,}67 \cdot x_2 = 0{,}11 + 0{,}35 \cdot 7 + 0{,}67 \cdot 10 =$$
$$= 0{,}11 + 2{,}45 + 6{,}7 = \boxed{9{,}26}$$

Regresioni mundet me u përdor edhe për Klasifikim

x_1	x_2	y	y_1	y_2
2	3	B	0	1
1	4	A	1	0
5	2	B	0	1
4	1	B	0	1
1	5	A	1	0
3	4	B	0	1

- Duhet me i kodu daljet, për klasën A marrim me përdor y_1 për B y_2
- E caktojmë një kufi psh 0,5
- Nëse rezultati i regresionit del mbi 0,5 atherë merret si 1.
- Qay dalje sai ki gay e ekuacione kemi me pas. Në këtë rast 2 ekuacione:

$$y_1 = W_{0_{y_1}} + W_{1_{y_1}} \cdot x_1 + W_{2_{y_1}} \cdot x_2 \qquad \Sigma x_1 y_1 = 0+1+0+0+1 = \boxed{2}$$

$$y_2 = W_{0_{y_2}} + W_{1_{y_2}} \cdot x_1 + W_{2_{y_2}} \cdot x_2 \qquad \Sigma x_2 y_1 = 4+5 = \boxed{9}$$

$$W_1 = \frac{71 \cdot 2 - 41 \cdot 9}{56 \cdot 71 - 1681} = \frac{142 - 369}{3976 - 1681} = \frac{-227}{2295} = \boxed{-0{,}09}$$

$$W_2 = \frac{56 \cdot 9 - 41 \cdot 2}{56 \cdot 71 - 1681} = \frac{422}{2295} = \boxed{0{,}18}$$

$$W_{0_{y_1}} = 0{,}33 \overset{-(-0{,}09)}{+} 0{,}09 \cdot 2{,}66 - 0{,}18 \cdot 3{,}16 = 0{,}33 + 0{,}24 - 0{,}57 = \boxed{0{,}0006}$$

$\bar{y}_1 = \frac{2}{6}$

$$y_1 = 0{,}0006 - 0{,}09 \cdot x_1 + 0{,}18 \cdot x_2$$

CamScanner ile tarandı

$\sum x_1 \cdot y_1 = \boxed{14}$ $\sum x_2 \cdot y_2 = \boxed{10}$

$$W_{1y_2} = \frac{71 \cdot 14 - 41 \cdot 10}{2295} = \frac{994 - 410}{2295} = \frac{584}{2295} = \boxed{0{,}25}$$

$$W_{2y_2} = \frac{56 \cdot 10 - 41 \cdot 14}{2295} = \frac{560 - 574}{2295} = \boxed{-0{,}006}$$

$\bar{y}_2 = \frac{4}{6}$

$$W_0 = 0{,}66 - 0{,}25 \cdot 2{,}66 + 0{,}006 \cdot 3{,}16 = 0{,}66 - 0{,}66 + 0{,}01 = \boxed{0{,}01}$$

$$y_2 = 0{,}01 + 0{,}25 \cdot x_1 - 0{,}006 \cdot x_2$$

* rë krejt këto raste superojm se janë linear separable mes vete

Provojmë për hyrjet $x_1 = 1$ $x_2 = 4$

$y_1 = 0{,}006 - 0{,}09 \cdot 1 + 0{,}18 \cdot 4 = \boxed{0{,}71} > 0{,}5 \Rightarrow 1$ A

$y_2 = 0{,}01 + 0{,}25 \cdot 1 - 0{,}006 \cdot 4 = \boxed{0{,}23} < 0{,}5 \Rightarrow 0$ B

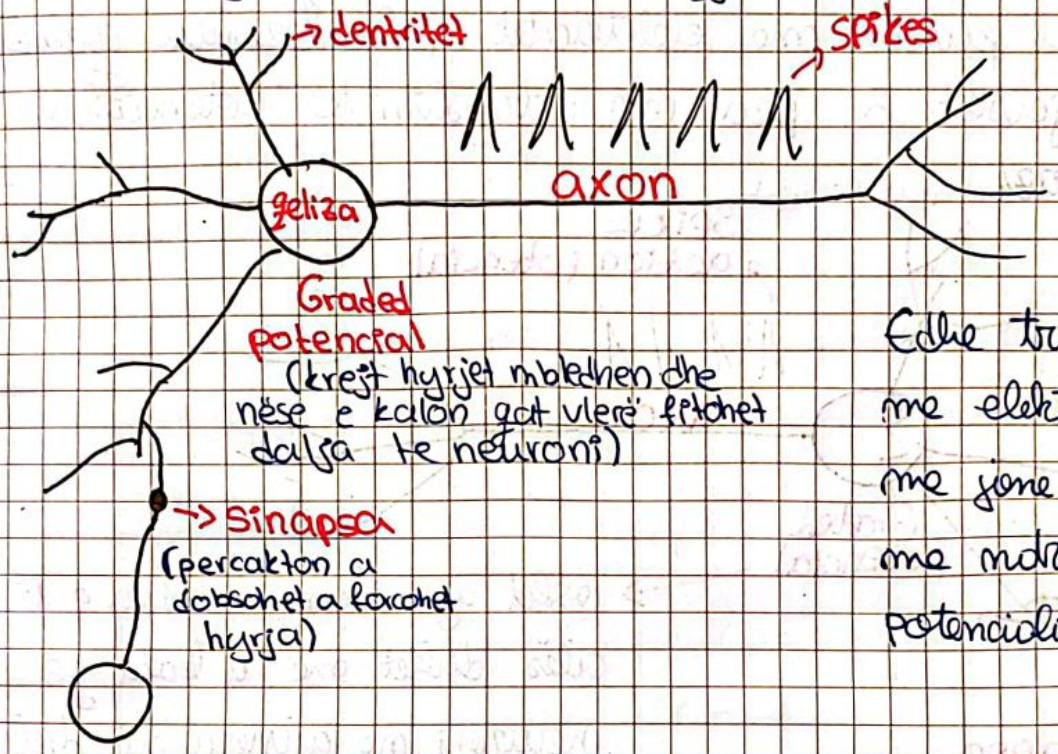
- prej rezultatit e fitojmë përputhjen me tabelë

Te unsupervised learning e patëm : K-means dhe SOM

Te supervised learning $\underbrace{\text{K-NN, Bayes}}_{\text{klasifikim}}$, $\underbrace{\text{Regression}}_{\text{klasifikim + paraslykim}}$

Neuroni artificial

- fitohet prej Neuroneve biologjike



Edhe truni komunikon me elektricitet por vetëm me jone pozitive, pra me ndryshimetë potencialit (voltazha)

Λ - action potencial : nëse ka frekuencë të lartë gjenerohet më shumë Λ // Λ ma shpesh kurse nëse o më e vogël gjenerohen më rallë

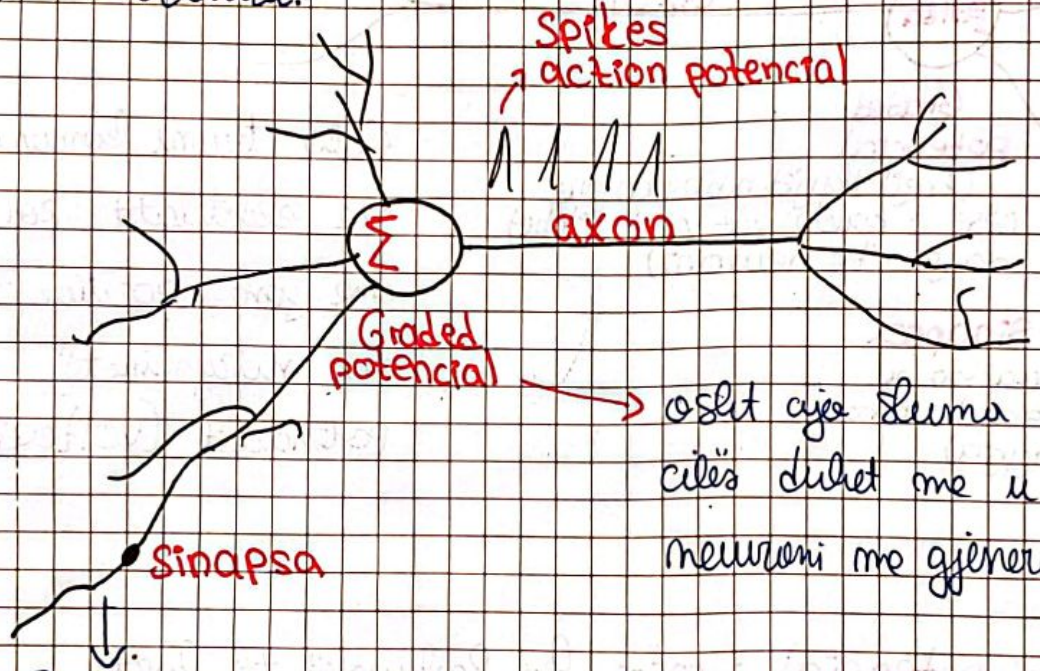
- fidhjet sinaptike forcohen duke mësu

Ligjerata 7 — Neuroni artificial

• **Neuroni artificial** përdoret edhe për paraqitjen edhe për klasifikim

Dentritet mbledhin të gjitha sinjalet prej neuroneve tjera

Edhe truni punon me elektricitet, pshs brenda gelizës i kemi 2 jashtë 3 pra me ndryshim të potencialeve fitohet voltazh.



→ osht aja shuma vlera e të cilës duhet me u kalu që neuroni me gjeneru një dalje

• Sinjal prej një neuroni tjetër, ojen tensioni (voltazhi), lidhja sinaptike e percakton sa voltazh duhet me kalu te neuroni, pra ose e lejon me kalu te neuroni nëse o mjaftueshëm ose nuk e lejon.

• në fillim lidhjet sinaptike janë të dobta kurse me kohë duke mësuar

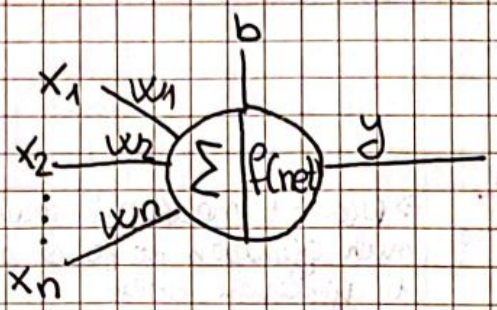
• **Spikes** për një kohë të shkurtë vrritet zvogëlohet voltazhi dhe gjenerohen pulse

» Nëse lidhja sinaptike është e dobët sinjali dobësohet dhe nuk del anej, nëse është e fortë del te neuroni tjetër

CamScanner ile tarandı

Neuroni artificial

Dendritet modelohen me anë të hyrjeve.
Hyrjet mbledhen dhe kur e kalojnë një vlerë kufitare bias b



- f(net) - është funksioni aktivizues që gjeneron daljen

- Ktu në dalje nuk gjenerohen spikes por 0 ose 1
0 nëse nuk e kalon kufirin, dhe 1 nëse e kalon kufirin
- Kemi: dalje të vazhdueshme, dalje jo lineare
- Nëse daljet nuk i fitojnë gjysh dojna, atëherë mundemi me i ndru te hyrjet me lisje të ndryshme i update-ojm peshat

$$y = f(\underbrace{x_1 \cdot w_1 + x_2 \cdot w_2 + \ldots + x_n \cdot w_n + b}_{net})$$

bias

funksioni aktivizues

• Ekzistojnë disa të njohura të funksioneve aktivizuese

• Step activation function

Θ - teta

$$f(net) = \begin{cases} 1 & net \geq \Theta \\ 0 & net < \Theta \end{cases}$$

→ Pra te dalja e neuronit marrim me fitu ose 0 ose 1 pra me dit a gjeneron dalje ose jo

• Sign activation function

$$f(net) = \begin{cases} 1 & net \geq \Theta \\ -1 & net < \Theta \end{cases}$$

→ Pra -1 neuroni marrim me thonë nuk gjeneron dalje 0 ne gatesi kurse 1 gjeneron dalje

Këto dy funksione Step dhe Sign gjenerojnë 2 dalje dalje dhe zakonisht përdoren për klasifikim

• Pure linear function

→ me fitu dalje në mënyrë të vazhdueshme

$$f(net) = \begin{cases} net & -\infty < net < \infty \end{cases}$$

• Sigmoid

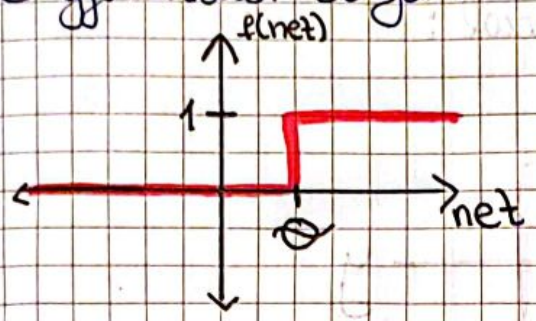
$$f(net) = \begin{cases} \dfrac{1}{1 + e^{-net}} \end{cases}$$

• Hyperbolic tangent

$$f(net) = \begin{cases} \dfrac{1 - e^{-2 \cdot net}}{1 + e^{2 \cdot net}} \end{cases}$$

CamScanner ile tarandı

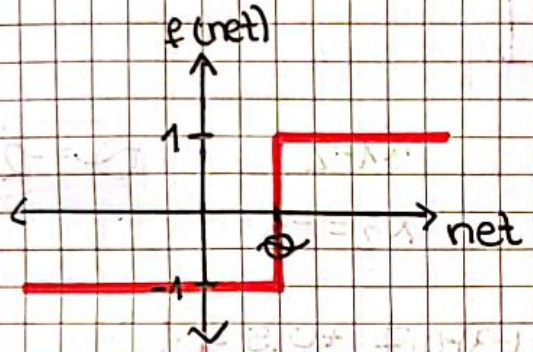
Net - mbledhja e sinjaleve dhe të shumëzuara me peshat përkatëse i shtohet biasi dhe kështu futet në funksionin aktivizues dhe gjenerohet dalja.



Step

$< \theta \Rightarrow 0$

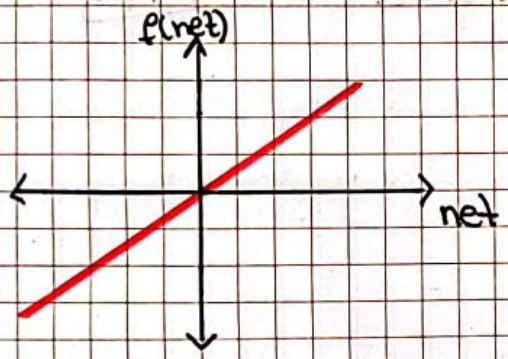
$> \theta \Rightarrow 1$



Sign

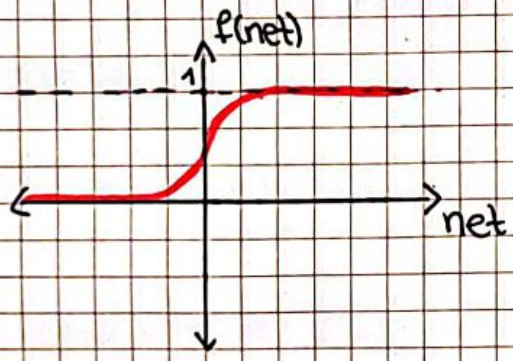
$< \theta \Rightarrow -1$

$> \theta \Rightarrow +1$

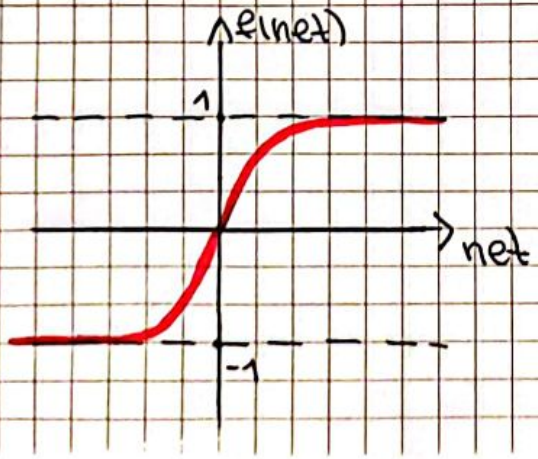


Pure linear (e vazhdueshme)

$net = f(net)$



Sigmoid (edhe kjo osht e vazhdueshme por nuk është linear)

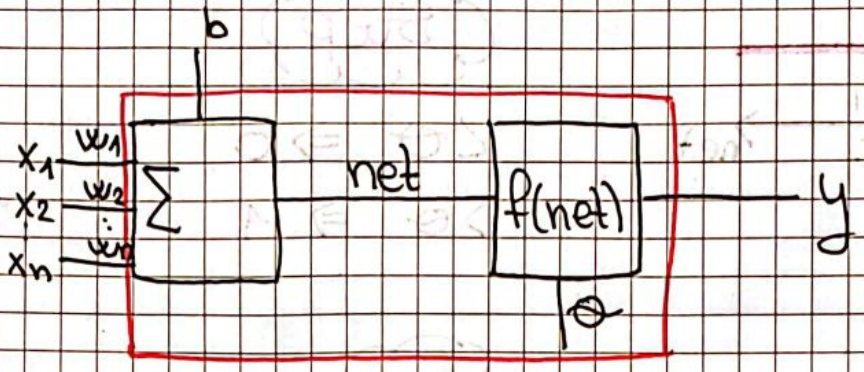


Hyperbolic tangent

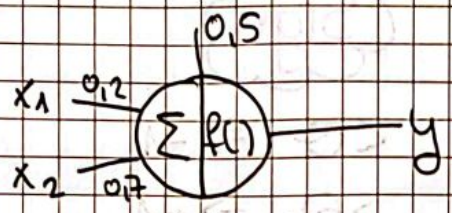
CamScanner ile tarandı

$$y = f\left(\sum_{i=1}^{n} x_i w_i + b\right)$$

Pamje më e detajuar e neuronit artificial:



Shembull.



$x_1 = 2$ $\quad$ $\boxed{\theta = -2}$

$x_2 = -3$

$net = 2 \cdot 0{,}2 + (-3) \cdot 0{,}7 + 0{,}5 = \boxed{-1{,}2}$

Step: $f(net) = \boxed{1}$

Sign: $f(net) = \boxed{1}$

Purelin: $f(net) = \boxed{-1{,}2}$

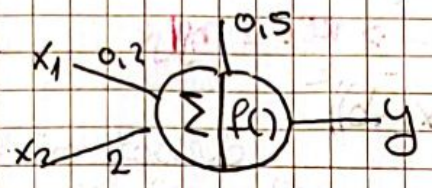
sigmoid: $f(net) = \dfrac{1}{1+e^{1.2}} = \boxed{0{,}23}$

hyperbolic tangent: $f(net) = \dfrac{1-e^{2/4}}{1+e^{-2/4}} =$

CamScanner ile tarandı

Psh për hyrje 2 me -3 dalja e dëshiruar është 0 por po na del 1.

Atëherë mundemi me i ndru peshat ose edhe biasin por hyrjet nuk bën me i ndru asnjëri



$x_1 = 2 \quad x_2 = -3$

Duhet peshën w_2 nga 0,7 e bëjmë 2

$net = 2 \cdot 0{,}2 + (-3) \cdot 2 + 0{,}5 = -5{,}1 < -2$ tani dalja na del $\boxed{0}$

Për këtë ekzistojnë ligje për me i update peshat

Perceptron: dhe Delta learning

- Në vrastin kur pesha w_2 është 0,7 gabimi është -1
- Në vrastin kur pesha w_2 eshtë 2 gabimi është 0

gabimi — dalja e dëshiruar — dalja e fituar

$$e(t) = (d(t) - y(t)) \rightarrow$$ nëse gabimi del i ndryshëm prej 0 atëherë duhet me i update peshat

$$W_i(t+1) = W_i + \eta \cdot e_i(t) \cdot x_i(t)$$

$$b(t+1) = b + \eta \cdot e(t)$$

- $t+1$ → duhet e updatuar ose e re
- η - eta është learning rate apo shkalla e mësimit. duhet të kemi kujdes gjat përcaktimit të saj sepse η e madhe neuroni mson shpejtë η e vogël neuroni mson kadal

me zëvendsim të $e(t) = d(t) - y(t)$ kemi:

Update me anë të Perceptron:

$$W_i(t+1) = W_i + \eta(d-y) \cdot x_i \qquad b = b + \eta(d-y)$$

për peshë — për bias

- Nëse e trajnojm neuronin mas një kohë tërësisht bët me i pështat me daljet e dëshiruara.

CamScanner ile tarandı

Adaptive linear - Delta Learning Rule

LMS - Least Means Square Rule

$$W_i(t+1) = W_i - \frac{1}{2}\eta \cdot (\underbrace{\nabla E(e^2)}_{*})$$

gradienti (derivati parcial) e fitojmë një sipërfaqe të gabimit

* $\nabla E(e^2) = (d-y)^2 = (d-(w_1x_1+w_2x_2+b))^2$ → Për 2 hyrje

derivati me bazë të w_1

$$\frac{\partial E(e^2)}{\partial w_1} = 2(d-(w_1x_1+w_2x_2+b))\cdot(-x_1) =$$

$$= -2\cdot\underbrace{(d-y)}_{e(t)}\cdot x_1 = \boxed{-2\cdot e(t)\cdot x_1}$$ → njejt del edhe për w_2 por me x

$$W_i(t+1) = W_i - \frac{1}{2}\eta(-2\cdot e(t)\cdot x_i) = \boxed{W_i + \eta\cdot e(t)\cdot x_i}$$

Për linearë shumëm me derivu, sepse derivati paraqet shkallën e ndryshueshmërisë e për linear na del infinit

- te ADA → Adaptive Linear fitohen gabime të vazhdueshme

$$e(t) = d(t) - y(t)$$

- Perceptroni Zakonisht përdoret për step ose signum ku gabimi mund mi pas vetëm p.e. 3 vlera $\{-1, 0, 1\}$ kurse te funksionet e vazhdueshme dalin gabime të vazhdueshme

Për gabim = -1, dalja u prit 0 por ka dal 1
Për gabim = 0, dalja u dal aktiv sig u prit dmth ska gabim
Për gabim = 1, dalja u prit 1 por ka dal 0.

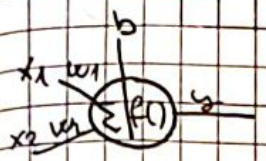
CamScanner ile tarandı

Shembulli.

	x_1	x_2	D
hyrja 1:	1	0	1
hyrja 2:	0	1	0

$\boxed{w_1 = 1}$ $\boxed{w_2 = 2}$ $\boxed{\eta = 0{,}5}$ $\boxed{b = 0}$

peshat fillestare perdoruesi vendosë kurse biasi mundet me fillu me 0



Struktura e neuronit me 2 hyrje f() e perdorim Step sepse daljet janë 1 ose 0

Fillojmë me iterime, fillimisht i fitojmë daljet, i krahasojmë me daljet e dëshiruara dhe nëse gabimi s'o 0 duhet mi update peshat

hyrja 1:

$$f(net) = \begin{cases} 1 & net > 0 \\ 0 & net \leq 0 \end{cases}$$

$net = x_1 \cdot w_1 + x_2 \cdot w_2 + b =$
$= 1 \cdot 1 + 0 \cdot 0 + 0 = \boxed{1}$

$f(net) = \boxed{1}$

$e(t) = 1 - 1 = \boxed{0}$ – pra për hyrjen e parë gabimi është 0 prandaj s'ka nevojë me i update peshat.

hyrja 2: $net = x_1 \cdot w_1 + x_2 \cdot w_2 + b = 0 \cdot 1 + 2 \cdot 1 + 0 = \boxed{2}$

$f(net) = \boxed{1}$

$e(t) = 0 - 1 = \boxed{-1}$ → gabimi negativ duhet mi update-u

*• e(t) = dalja e dëshiruar – dalja aktuale nga neuroni rëndësi

Updates: $w_1 = 1 + 0{,}5 \cdot (0-1) \cdot 0 = \boxed{1}$

$w_2 = 2 + 0{,}5 (0-1) \cdot 1 = \boxed{1{,}5}$

$b = 0 + 0{,}5 (0-1) = \boxed{-0{,}5}$

Ndërsa w_1 s'ka nevojë me update, sepse edhe s'kishte te hyrja e dytë $x_1 = 0$ pra s'ka ndonjë ndikim

• Algoritmi nuk përfundon deri sa fitohen krejt daljet si daljet e dëshiruara.

Iterimi i dytë:

hyrja 1:

$net = 1 \cdot 1 + 0 \cdot 1{,}5 - 0{,}5 = \boxed{0{,}5}$

$f(net) = 1$ $\boxed{e(t) = 0}$

hyrja 2:

$net = 0 \cdot 1 + 1 \cdot 1{,}5 - 0{,}5 = \boxed{1}$

$f(net) = \boxed{1}$

$\boxed{e(t) = -1}$ → prap duhet me update-u peshat

$\boxed{W_1 = 1}$ → s'ndryshojm kësj sepse $x_1 = 0$

$W_2 = 1{,}5 + 0{,}5\,(0-1) \cdot 1 = \boxed{1}$

$b = -0{,}5 + 0{,}5\,(-1) = \boxed{-1}$

hyrja 1:

$net = 1 \cdot 1 + 0 \cdot 1 - 1 = \boxed{0}$

$f(net) = \boxed{0}$

$e(t) = 1 - 0 = \boxed{1}$ → duhet me update

$W_1 = 1 + 0{,}5 \cdot 1 \cdot 1 = \boxed{1{,}5}$

$W_2 = \boxed{1}$ → nuk e ndryshojmë sepse $x_2 = 0$

$b = -1 + 0{,}5 \cdot 1 = \boxed{-0{,}5}$

hyrja 2:

$net = 0 \cdot 1{,}5 + 1 \cdot 1 - 0{,}5 = \boxed{0{,}5}$

$\boxed{f(net) = 1}$

$e(t) = 0 - 1 = \boxed{-1}$

update: $W_2 = 1 + 0{,}5\,(-1) \cdot 1 = \boxed{0{,}5}$ $b = -0{,}5 + 0{,}5\,(-1) = \boxed{-1}$

$W_1 = \boxed{1{,}5}$

CamScanner ile tarandı

hyrja 1:

$net = 1 \cdot 1{,}5 + 0 - 1 = \boxed{0{,}5}$

$\boxed{f(net) = 1}$

dalsi mire

$e(t) = 0$

hyrja 2:

$net = 0 \cdot 1{,}5 + 0{,}5 \cdot 1 - 1 = \boxed{-0{,}5}$

$\boxed{f(net) = 0}$

dalje mire

$e(t) = 0$

CamScanner ile tarandı