Machine learning Algorithm 8/8/21
Machine learning Algorithm 8/8/21
ok Weatest Word land
DAGEN IS NO.
- Easy to implement supervised mading learning
that can be used to solve both chase Contract by
- Easy to implement supervised machine learning algorithm that can be used to solve both classification is regression problem.
Posting it down
Beating it down who have the
- Supervised Machine Immine Alanili
- Supervised Machine Learning Algorithm is one that
depends on -> lakeled in put data
to learn a function that produces
an appropriate autout whom
given new unlabeled data.
given new untabeled douta.
-> classification problem problem
Al have discrete value as its output.
Mot a mitte matical operation 11.
have predictor (a set of predictor) & a label.
Live in the second
the will be the place and the world for

Arion Agra

· Participation

12/8/8 Madine looming Algorithm -> A regression problem no has a real number, Bits outpost. - ! As have one or more independent von: able or a dependant variable: Mod MIN-> - Cosy to implement all poer and whiteher from more had some that outso star bout and in tool ". A dary 1 page out -> An unsupervised Machine learning. CHO water use of input data without any toit so a labels inplice palmont and one becaused are Tries to learn the basic structure of the and hobbe de give us unove insight into the note that gate of bluebuilds of standardo we. onen new unlabeled duta. -> KNN appears the idea of similarity (distance proximity or close ness) to the > K-NN raignoithm stores; althoughte aveilable dorta: 3 classifies a rew obta point based On Einilarify. -> New datas can easily be classified into well suite category

-> K-NN is a non-parametric algorithm if does not
-> K-NN is a non-parametric algorithm, it does not make any assumption on underlying dota.
server with the formand of the analysis of the server of
-> It's called lazy learner algorithm
1 It does not learn from training set,
instead it stores the dataset 3 at a time
of classification, it perform an action on darboset.
-> K-NN algorithm at the training phase.
astores the ditaset
On when it gets new data it classifies that data into
a category that is much similar to new dota-
(5,8) 6 25 (5,8) 6 + (0,8) 6: diagont of specifical
-> Types of distance Metrics.
amintousti Distonce
ap Makhattan Distance
Co Euclidean Distance * Most popular.
Cornell pistance election sit deligion del
are Cheby Over Dirtance !
parts () and down to g
Alle 하는 100 House Control (1980) - 100 House Control (1980) - 100 House Control (1980) - 100 House Control (19

of and market smoth

Detence Method Concepts

0	Mintowski Dictore
	- Intented for real-valued vector space.
_	- Calculate distance in a normed lever space.
	ap spaces where distances can be represented
	as a vector that has a langth is the
	lengths cannot be repative.
	the contraction of the contraction is
3	Conditions Conditions
	nd Non-negativity: d(xy)>=0
	and Identity; d(x,y)=0 if yould if
	$\alpha = -y$
	ad (y,y) = d (y,x)
_	AD Trangle Inequality : d(xy) + d(y,z) >= d(x,z)
	Company has constituted
	(n y: -11 P) / P
	[[]]]]
	We manipulate this famula with different privatures.
	p=1 ~ Milli With different privatures.
	p=1 ~ Mahal Monhattan Distonce.
1	Collicean Distance.

Manhattan Distance	A1156A
- Known As Taxical / City block distance.	17.
- Distance between two points is the own or	file
absolute differences of their Contesion cook	rdinake.
The state of the s	
from (= 12 12 12) /p	
TEI TONG TONG TONG TONG TONG TONG TONG TONG	
$ \frac{\partial^2 \mathbf{z}}{\partial z} = \sum_{i=1}^{n} \mathbf{x}_i - \mathbf{y}_i $,/
i=1	
eg/ red (4, 4) 1 & green (1, (1) 1) 601	Ų.
eg/red(4,4) green (1,1):	
- (1-10) = (110) = (110) b	
-> This distance is preferred in case of	high
dimensionality.	
J	

...

Euclidean Distance - It measure of the true straight line distance between two points of fuclided Space. if red (4,4) 3 green (1,1) 4 1 22 y2 -11 = 6 No:N

Cosine distance	individe wind of
- Calculate similarity before	een two vectors
- Measure by the cosine	of the angle between
ofwo vectors pointing in	herisand direction.
	elymone does not e
- This distance gives us	a new perspective
to a business problem b	find some hidden
information. 2 2000 oit	sandary all by It
· restor	other hardren or at
$\cos \theta = a.b$	
1 all all all all all all all all all al	a company of a first
De will get 2 values (0	4:01/3/18/ 18/07 2:3/3.4
De will get 2 values (0	\$1) 0 = 100% similar
	1 = 0°/0 similar
15472 NJ- 1011 2511/194	
	· wine of
	to elected off papers of
-29 infants 3 131 73 -10 701	
	1100 10 231 57 71 .73
would be to now all no	21.37

	Some distance
,	T lealed (1) allows
1	The Practice of the Control of the C
; I.	Load the data in suizes of make
2.	In the 120 0 Know your Copanie.
_	I acomple in the or
3.1 A	Detace Detace
	· collection:
3.2	Add the distance & the index or the example.
	to an ordered collection.
	a
4	Sort the ordered collection of distances by
	indices from smallest to largest (in ascarding order)
30)	by dictorces. (18 0) wind a top 11000
40	1 = (1° /2 = 1
5.	Pick the first K entries from the sorted
	Collection.
6.	Get the labels of the selected Kentries.
-	
म	. If regression, return the moon of K lebek.
- 0	10
8.	If classification, return the mode of the k lobels.

Application of KNN	
- X	Lava to sorthand
	y
- Climate forecasting - Estamating soil water	The warm of the same
- Estamating voil water - Stock nouted forceasting	trade of the Asymptotic State of the State o
- Stock nourted forecasting	parameter.
	CONTROL TO THE PROPERTY OF THE PARTY OF THE
Wiero el oby	d software street so
Bank	20/11/16 2/01/61 C-
Briency Banknin Let &	
xemale	
rates.	Understanding 3 Managing Financial
KNN in	Ret
	1614 . To - work 1
Money lourdering to Finance	Trading Futures
	1 Tel Tennello don
analy ses	1.
analy ses	- Loan management
analyses Book	- Loan management
Bonk custoner	- Loan management
Bonk custoner profiling	- Loan management
Bonk custoner	- Loan management
Bonk custoner profiling	Loan management
Bonk customer profiling	Loan management de la
Bonk custoner profiling	Loan management de la
Bonk customer profiling	Loan management 1-11-11-11-11-11-11-11-11-11-11-11-11-1
Bonk customer profiling	Loan management 1-11-11-11-11-11-11-11-11-11-11-11-11-1
Bonk customer profiling	Loan management 1-11-11-11-11-11-11-11-11-11-11-11-11-1

Mill to mitoring Alventages of KNN -> Rowst to noisy training data -> Effective if training obta is large. -> No training phase training for the -> Learns complex mode is easily JARG Bunfrisp Heles E' probrote 1951 1.01 top ri WW Diadrantages of KNN - Get slower if the number of independent variables incre-ses. - Need to dotermine the value of prometer K. Low Computational Distances between efficiency data objects becomes less dustinct. High Dimensional Data sporsity Data False longer amount of Intition data 3 storage

Choosing the right value of K

- (harge the value of K & run KNN algorithm
 several times to achieve: 3-
 - K that reduces the number of errors
 - Maintaining the algorithm's ability to
 accorately make predictions when it's given
 data it hosn't seen before.
- As we increase the value of K to 1

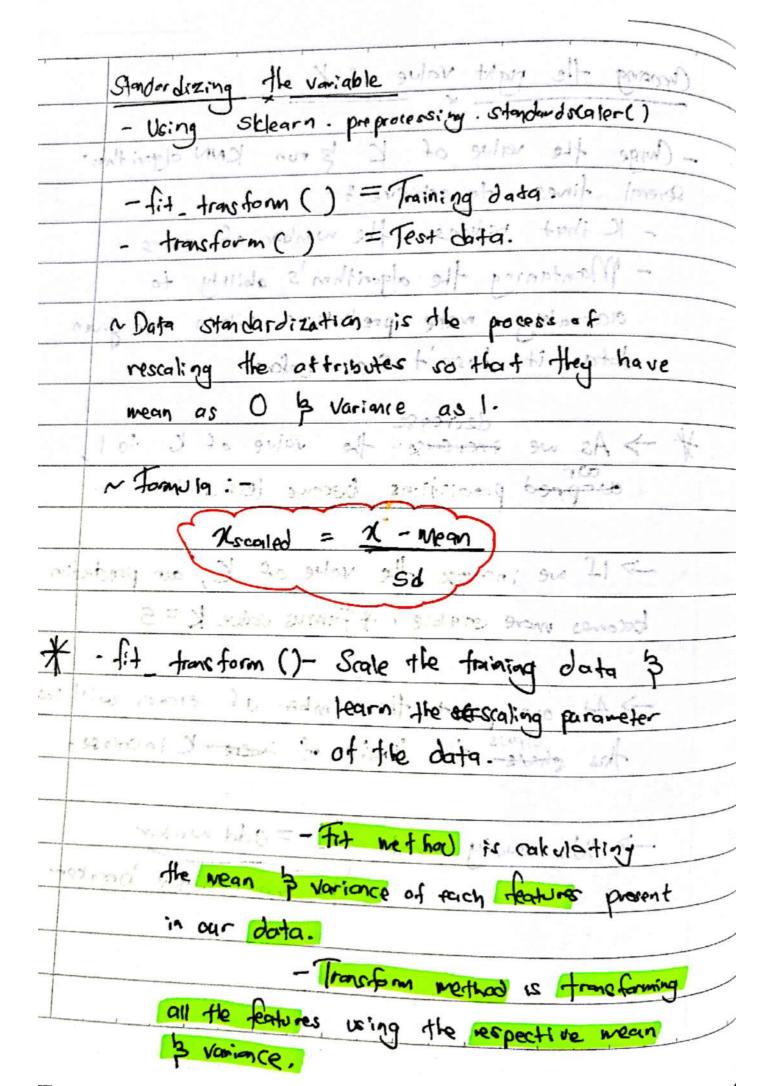
 our pred prodictions become less stable.

6 stoo mirst st die - () and and the

- -> If we increase the value of K, our prediction becomes more stable . # Famous value, K = 5
- -> At one point the number of errors will increase,
 this chose the limit of the K increase.
- -> We usually make K = odd number

 To moke a fire breaker

wantigard 2 to 1 ya magazin



transform () - We can use the same mean 3 Variance as it is coloulated from the training data.

Why we doesn't use sit in our test data?

No It will compute a new mean pranance

that is new scale for each feature.

~ Will teams our model learn about our test