Garl A

	DATE:
	Assignment?
)	Explain différent motrics used for evaluating
	classifier performance.
a Avs	A curacy -
11	This is the most basic measure for classifier
	performance, and simply measures the proportion
	of correct predictions made by the classifier.
9 ii	Precision -
	This meths measures the proposition of the
	positives out of all instances predicted to be
	positive by the classifier.
B V	Positive 39
(iii)	Pl'swre-in-
	This is the narmonic mean of the previous
1000	and recall and provides a way to paratice the
	two metrics.
- 12	
(IV	Confusion matrix-
- 18th	Training the matrix that character the number of
2 3 1 · ·	the positives true negatives false positives and
and the same	false negatives for a dassifier and provides
W	the positives, true negatives false positives and provides false negatives for a dassifier and provides as more detailed view of its performance.
tory 11 m.	the felling many the state of t
	What is linear regression? Explain how it is
	different from classification.
pos	different from dassification. Vincar regression is one of the easiest and
Fr.	most populor machine laming algorithms. It is a
Straffe Was A	statictical method that is used for predictive
	most populor machine learning algorithms. It is a statistical method that is used for predictive analysis, linear regression make predictions for for educational use
	, FOR EDUCATIONAL USE

was a supplied of the state of	A particular to the state of th			
and divine a first was ready passed and apparent	continous/real or numeric variables such as sales,			
	age, product prices			
AND THE RESIDENCE OF TH				
全世界的自然的學習和特別的問題。	Regression	Classification:		
Marie Ma	In Regression, the object	Classification: In classification, the output variable must be discrete value.		
一个上述的表现在是一个人	output variable must	ranable much be discrete		
nestra a strain a serie	be of continous nature	value		
CONTRACTOR OF THE STATE OF THE	or real value			
1	Used for continous data.	Died for discrete		
and the same of th		data		
29	Uneare regression	Logistic regression.		
	mill as to be to a code on	and divisive dustria		
3)	ala socialismo	renative and divisive dustria		
Parameter	Agglome-rative	Divisive.		
late goog	Bottom-up approach	Top-down approach:		
Approach	Each data In its owndows	All data points are in a		
Manager and Property and Proper	and the algorithm	stoal witer, and the algorith		
Description of the Property of the Party of	rounciardy mragerite	recursively splits the dister		
	docat pairs of dusters	into smaller subsets until		
	until a single duster	tach point is in Pts own		
	containing all points	cluster		
Onder the Control of	is obtained.	The state of the s		
Law Jalla	The state of the s	Te lose comblex caroner Alvelo		
CHOOLEXIII	7. more complex.	Te lass complex companishedy		
The state of the s		OR VDUCATIONAL USE		

FOR EDUCATIONAL USE

and the second s	DATE:
Outliers	can handle them better pivisive may made sub-
The state of the s	I han divisive clostering dusters around outliers.
4,	The least of the last of the l
	Explain two feature selection measures in
	building decision tree.
ons i)	Toboxonation agin (79)
	It is a widely used feature such on measure
	in decision tree algorithms. It is based on the
Tar	concept at controll which measured importy
•	randomness of a set of clara. I a is coward
	as the medicin on in entropy actived by spaining
T Howard	the data on a particular feature. The higher he
R	It the more informative the feature is considered
1.51 46.5 L	to how it middle school is a charles
	I'm = Entropy before - Mohlen torey after sput)
	split
	Gioi index -
	It is another commonly used peature selection
0	and the similar to entropy. It is rate
	calculated as the probability of miscas misclassification of a randomly choosen element in the dataset when using the peature for splitting.
Terri.	of a randomly choosen element in the awaset
	when using the pearure for spulling.
3 8 3 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
A PART CAR	Cini=11- 5(p)
	p is probability of each class.
	1. de liberad medured by Badina interducte
59)	Hhat are the different measures for finding intercluster distance? Determine the distance between two clusters of (17, 42, 10) and (2(20, 36, 8) using a single linkage minimum, distance, techniques.
The second secon	austrance y Description (2 (20 36 8) Using at alphale links
The state of the s	distance techniques
A second of the	FOR EDUCATIONAL USE

tros.	Evilidean distance-
1	This is the most commonly used distance metric,
	This is the most commonly usual distance metric, and it et calwates the straight distance between
	two points.
11	man hattan destance -
	Also lenger as the 11 distance or city block
16	listand Translated Inches Solly of the most of
λ	differences between the co-ordinates of two points in a grid-like fashion
1.	points in a grid-like fashion
1 1	The state of the s
(::	Min Lowski distance -
	Minkowski distance in ois a generalised distance
	Minkowski distance in ois a generalised distance metric that includes tididean distance and Manhatton
11	distance as the special cases.
	c1 (17,42,10) (2 (20,36, 18)
,	Distance = \[\[\langle \co -17\rangle + \(\langle \co -17\rangle \cdot\) + \(\langle \co -17\rangle \cdot\)
-1	$= \sqrt{3+(-6)^2+2}$
1, 2, 4	= V.9+4+36
Toy by V	All III E TO THE TOTAL TO THE STATE OF THE S
5	Explain Agglomerative (AGNES) clustering algorithm
	with an example. Comment on Denabgraun cold
	elector tormation process.
ms	It is a hierarchical disting againston
	that oberations overal similar westers based
	or a chosen similarity or dissimalinity memo
11.	or a chosen similarity or dissimalisity metric until a stopping condition is met the algorithm, starts, will all points, in metatosis.
Walter Trans	diannthm start will all points in me classic

DATE: its own doster and then successively merges dosters until a single duster containing all points is frimed. eg (A, B, C, D, E, F). i) Initialization: AU points porming it= own cluster in start 19), 18), (C), (D), (E), (F). ii) Compute similarity I dissimalinity
(al whate similarity / dissimilarity between dusters

based on a choosen metric. iii) Merge dustersMerge two most similar dusters into a single duster. Lets say in) > (B) have the highest smilarity, so they merge into a new duster in) Updale the similarity-recalledte the similarity between newly tramed dusters (AB), (c), (D), (=), and (F) 1) Ropeot step 2-4 till only one choiter is left in the end vi) Dendogram -During dostering a tree like storcture is presented. It shows the biesarchied storcture of the dosters and the order on which they

	DATE:
, 3 ,-	were merged
·	
4 vi	Debster bornation-
	The final clusters are formed by witting the
	dendogram at a certain height or distance,
	which corresponds to desired number of
	dusters.
e Ten	The second of th
m T	A database has five transactions. Let min
* 1	I sub = 50% and min confidence = 100 1. That
	all prequent algorithms using Aprion algorithm.
V	and I have the self the self of the self o
	Them-bought 71 RMONKE, 73
	71 RMONKE, TY
	72 \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \
	13 \$ (A, K, E)
F	74. SCOPIK, IEY.
113 113	75 5 M, U, C, K, EY.
	support = 50 % = 50 X 5 = 3.
, , , , , , , , , , , , , , , , , , ,	1- itemset -
1	Item frequency.
1	A
	-1. 2.1.
	D 4
	9
	5
Wat was	3
- · . · . ·	2
di- di	FOR EDUCATIONAL USE
VATO	

		DETEN	
	1- 1+em set = & E, K, 1	1,0,13	
* 1.0	2 - item set -		
		A company of the same of the s	
	9 tem	frequency.	AST AST
	EM	2	Description of Publish
		2	
	EY.	2	
0	K M	3	
		3	
	MD		The state of the s
	M	2	
	4.7	2	
15/11	2 - itenset = { (6, x)	EX SE 04 SK MG SK C	2 68 42
	2 1101130 - 1017		ana kanangan kanan ka kanan banga
	3-item set-		
	i tem	regional	
	E,k,O	Frequency	
The state of the s	E,kM	2 1	
	E, K, Y	2 1	
11121	E, O, M		
	E, 0, Y	7	
	K,M,O		
	KM, Y	2	
	x0,7	2	
T I	3-itemset = 7 8 E,	x,043	
	pian . Vi		Mallan.
	4 4 1		The state of the s
			THE REPORT OF THE PROPERTY OF
	FOR	EDUCATIONALUSE	Contract of the State of the St

And the second of the second o			And the second s
8)	A dataset has	5 transactions. Let	min cup
	and min contid	ence = 500/ 80 1. Find	all
	prequent "temsete	ence = 50° / 80 'l. find veing fp growth	
	es & Switchers of	7	
	410	them bought.	
	7100	am, on keyy	The second secon
	7200	& D.D. IV. K. F Y?	Legit C
	7300	\$ MA KEY.	- 10-
	7400	SMUELKY 3	(CE)
	7500	\$ MA KEY. \$ MUELKY? \$ L, O, O, K, RE?	
		\/ ·:	And and
	Items	coont.	The second second
	Α (KA .	
	C	2	
	William De Party	(s 1 - 1 - 1 - 1 - 2 - 2 - 2 - 2 - 2 - 2 -	MAN STATE
	E	4	Traff.
	ĭ	, a for the start	
	11K 10 11 Test	5 111	The state of the s
	M	3 0 4 7	
	~	2 (1)	
	D	4 (8 9	
	J	1 18 1 1 1 1 1 1 1	
	7	3 4 1 0 1 1 4	
	1- item set = FE	K M, 0, 73	
and the second s	710	list of items.	
	7100	SE, x, nn, 0, y3	
Walter Commencer	7200	\$ E, X, 0, NY	
- V-	7300	EF,K,MJ.	
1 12	7400	5K W , 4 3	
	7500	IKE, K, OY	
All and the second seco		FOR EDUCATIONAL USE	an eagle, eather the annual and a several teath of the several and a several and a several and a several and a

MILO MILO VIIO

612 C NULL
612 C NULL
011 C C 111
011 C C 111

E:40 ,0m:1 rn:20 20:2 0:10 707:1

Making FP tree -

* * *

DATE:

	Itemset	Conditional	Conditional	Frequent	
		Pattern	FP-tree	Pattern	
		*		-2	
	ξ Y 3.	8 K, E, M, 0113	{k!33	{k, 7:34.	
		くた、E、O.Y. るお、Min.	2.9		
		& K, M'g.	* * * * * * * * * * * * * * * * * * * *		
		ř	9	C1 0122	
	401	3 K, E : 23	\$ K!3 E:33	2K,0:39	
		5 k, E:24		? E, 0.39	
		1- 1		5	
Å	a m 3	SK, E:23	5 £ : 3 3	9 k, M: 33	
		\$ K:13			
		the section is a	64.4.7	\$\$,€:49	
	र्व ६ ५	3 K : 49	\$K!49	2 F, C . 41	
		<u> </u>	3 1 21	# 03.00 T	
	a x 4		74	F (6))	
			1.0	T. C.	
		1 Arch 1 V			
		7.		To the	
			77 **-	A A A A A A A A A A A A A A A A A A A	
		1-10-1-1			
		<u> </u>	~ (N)	140 mm	
		1 101	7	Armana .	
		7-1-3/2	<u> </u>	T	
V. III				4 0 100g	
DAME.			era anidon	16.	
		2001	1 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		

FOR EDUCATIONAL USE