21921	DWDM
	Jutorial-5
	U18co021 - Suhil Bonde
91	Converting decevior tree to rules is preferred because:
Sales Section Control of the Control	Converting to rules allows distinguishing among different contexts in which a decesion node is used
	It removes the distinction between attribute tests occurring at the rea root & leaves of the tree
	It improves readability
	and the second of the second o
	The second of th



92	
a	To consider count of data tude, it
	needs to be order included in
.)	To consider count of data tuple, it seeds to be trobe included in the attribute selection measure
.3.	(such as information gain)
.5.	2000 20 10 20 10 10 10 10 10 10 10 10 10 10 10 10 10
6.	Solary Rayer: A = {26-30 kg, B= {31-35 kg
	C= {36-40k3, D= {41-45k3
	E = {46-50 k3, F= £66-7043
8.3	5.0 1 1 1 1 53.0 4 2
	{A,B, 03 (Sulary) {C, F3
0	
	{5}
	jowier Senior
	(Department)
€5	ecretary, system? { Sales, marketing?
	owor Sewor
	7
C	X = { Systems, 26 30, 46-50k}
	P(X journ) = 0.182
	P(x jouior) = 0.182 P(x serior) = 0
	C162 D D D D
	NB predicts janor

<i>Q</i> 3	
Tople	Class P. TP FP TTN EN TPR FPR
1	P 0.95 1 0 5 4 0.2 0
2	N 0.85 1 1 4 4 0.2 0.2
3	P 0.78 2 1 4 3 0.40-2
4	P 0.76 3 1 4 2 0.60-2
.5	N- 0.60 3 2 3 2 0.60.4
6	P 0.55 4 2 3 1 0.8 0.1
7	N 0.53 4 3 2 1 0.8 0.6
8	N 0.52 h h 1 1 08 0.8
9	N 0.51 4 5 0 1 0.8 1
10	P 040 5 5 8 0 1 0
	1 2 833
	Leines Senier
	D-8 (Nepatinent) 8-0
5.	
7.	6:6 m. of 6 & Super of the 63.
	0·h
	0.2
	. 531.0 : (
	0 0.2 0.4 0.6 0.8
	ROC Carre
	13 Madrille imir
-	
1	

04	Methode has class into laces.
	Methods for class imbalance:
	On the Part of the
	tersampling: resample sometom tuples from
	Oversampling: Resample soundon tuples from numority class
	4
ii	Mudersampling: Delete random taples from
	Teamine of al
	majorety dass.
	7 10 11 (1)
icl	I hreshold Moving: Shifting the threshold
	to seduce the cartle
	Take neerting essen
-	Threshold Moving: Shifting the threshold to reduce the costly false negative errors.
	Cat Will the Al
_	Touthing . Indesing the Cost of
	mes-classifying the
	Cost-Modifying Incresing the cost of minority class
سرا	Eusenble mothods : Using boseting and
	The state of the s
	Ensemble methods: Using booting and random foreste.
	1
	5 VM & MLP would perform better on
	embalanced datasts than tree based
	5VM & MLP would perform better on imbalanced datasts than tree laser classifiers. SVM is not sensitive to imbalance as support vectors only look for boundries between clusters.
	into la con as convert yestore only lank
	moveme de suppero vectore and mor
	for voundries between cluster.
	Oversampling perfors better for numeric datasets like a credit and fraud transaction. Cost Modifying would help too.
	like a createt and hourd traverte
	CA 11 111
	case roomsying walla help too.