Decision Tree Algorithm. When data is like this you can easily draw a decision boundary as a straight line and you can use linear regression to do the pristing prediction. 2000 de algo elgo privade de sacagnition! To idantify handwritteh systimport matploblit. 12/1/20t as jeen min hen data is complex. sind me clecision boundary in le vits not possible to cloran as ingle dine to create une boundavies. so linear régression digits. dasa [100] is not possible here. desits taget [0:5] digits mages (0) of pet-mathow (diggits images [5]) Model creation. What is a Decision Pree ? How closs it works Decision tree is a type of supervised learning algorithm (having a pore-defined farget rariable) that is mostly used in classification problems. Li It works for both categorical and continuous Lis In toris technique, we split the population or sample into two or more homogeneous sets (or sub- pro populations) besed on most significant splitter / Lifterediator in improvement

o Categorical Variable Decision Tree: Decision Tree which has categorical target variable then it called as categorical variable decision tree. Example: - In above sentence scenario of student problems where the troop not " vie VES or NO: 102 play croicket or O Continuous Variable Decision Tree: Decision Tree has continuous target variable then it is called as Montinuous Variable Decision. O Root Node: It supresents entire population or sample and this further pets divided of Splitting. It is a process of divided into two or more sub-nodes. ODecision Node: Where a sub-node splits into further sub-modes, then it is called decision D'Leaf/Terminal Nock: Nodes de not split is called Leaf or Terminal nocle. O Penning: when we remove sub-nodes of a decision node, this process is called pouning. You can say opposite process of splitting. O Branch / Sub-Tree - A sub section of entire tree is called branch or sub-tree.

O Parent and Child Nocle: A node, which is divided into sub-nodes is called parent mode of sub-modes where as sub-modes are the child of parent node crowell sonbeen LX At strat, all the training examples on

Classification by Decision Tree Induction O Decision tree Took of the free structure in the structu Internal mode denotes a test on one attribute.

5) Branch represents an outcome of the Decision tree generation consists of two phases.

O Tree construction.

At start, all the training examples are in the root of the root of · Partition exomples succursively based on selected attoibutes. · Identity and remove branches that weetlect moise or outlivers

① Use of decision tree classifying an unknown sample.

Test the attribute values of the sample against the decision tree.

Algorithm for Docision Topo Tombustion Algorithm for Decision Tree Incluction

(a) Basic algorithm (a greedy algorithm) receisive divide - and - conquer monner L> At start, all the toaining examples are La Attailures are categorical (if continuous-valued, they are discretized in advance)

A. M. W.	nles	are partitioned	recursivel	y base 0.
10 Examples are partitioned recursively based on				
selected attaibutes.				
12 Test attachings and Only On the basis of a				
17 1est of the sole solected on the busis of				
in test attributes are selected on the basis of a heuristic or statistical measure (e.g., information				
conditions for de la gain)				
Conditions for stopping partitioning				
17 All samples for a giving mode belong to the same class.				
The street of the				
Destributes for further class. partitioning-majority voting is employed for dessifying the leaf.				
partitioning attributes for further				
does of Jessayority voting is employed &				
classifying the leaf = show ments				
5 Thaxa				
are no samples only				
Example I/ problem I mus siz some slamest Attaibutes Gender Car				
temale and side owns - 10 selections				
Attoibules				
Gender	Car	Cheap	- Indiana	classes
	2 Contraction	Lost Cost Co)	Income	Tromsportation
Male	0	Charles Error	or rever	smoole
Malo	257.55.0H	Cheap show	000	BUS
Female	1		-75.0)	Bus
Female	0	Cheap	redium	Town
Male		cheap	Medium	1945
Male	0	Stendard	Medium	Bus 3
Female		Stendard	reclium	Train
Female	MESI	Expensive		Cars
Male	wE2 mx	Expensive	Medium	
teraule	ri mi72	Expensive	High	C
Decision Tree for Exemple + sland provide				
Meas wes of Time smilly				
W. O.	TE	arel		
the wintravel cost (Km) reported to every is aldot A				
- O Note about a cost				
Expensive S Cheap colo large				
(Care) Stembard Chardonia				
3/5/10/10/10/10/10/10/10/10/10/10/10/10/10/				
		Fouin	Male	Temale
ig 3 (Bus) = robort (Car Ournesship)				
	75	G -		
		as = I - week & like	Fication barn	Train

Rules Crenerated barrothrong son 2019 mones selected attaibutes. If bavel cost/km is expensive then Mensistic ses = caspone l'anordens (cas) information If Travel cost / Km is standard then III same class. There are no remaining attributes on see It Franch Cost/Kont is cheap and gender in male then, mode = bas al ant Britig 2200 If Travel cost 1/km is choop and genderie female and she owns no car then, mode = bus It travel cost/Km is chaps and gender is Classification wing Decision Tree Travellest Income (\$)/km Level Standard High powstage Person Gender Corr name Comership Mode Alexas Male 7 Leain of Malerville 011 Cheap of Medium Buddy Buy M Female 1 Cheap 1X High Cherry A table is pure or homogenous it contains only a single class. If a clasa table contains several class, then we can say that the table is impure or heterogenouselen Entropy July - Ri log Pol Glysamin Chini Index = Ind Sp Classification Esson = I - more 7 p.7

Probability of classes Prob(Bus) = $\frac{4}{10} = 0.4$ Prob(Car) = $\frac{3}{3} = 0.3$ risot

Prob(Train) = $\frac{3}{10} = 0.3$ risot Entropy = = - 12 log 2 p = 10 d o d = stolo - d Entropy = -0.4 log2(0.4) -0.3 log 2(0.3)-0.3 log2(0) (200 12 man 9:52 8 8 -10:55211+0.5211=1.537) Crini Inden Crini Inclex = 1 - 2 pg labor - 1 Ango) Chini Index = 1 - (0.42 + 0.32 + 0.32) = 0.660 Classification Error Classification Essas = I - mox 263 Classification Error Index = I - Max 50.4,0.3,0.3 In notebook primale = 0.60 all ent import pandas as poli a server for tillong of Lover sklearen import tree x00 @ 200691 from skileam, preprocessing import Labeltneader Label Encoding Label Encoding sefers to converting the labels into a numeric form so as to convert them the machine-readable form. Machine learning algorithms com then decide in a better way how those labels must be operated. It is an important pre-processing step for the structured alabate

Load the data. 1922010 go etilidodos? data = pol. read - csv ('salaries csv') dor inputs = clasa. otrop [['salary-more then-100 Kill, axis = (columns!) -) dong target = data[salary_more_then_100K] Data preprocessing = Rdoepus ele_data = LabelEncoder () company l= le_claba. fit_tooms form (injoute of job-l= le data fit toonsferon (inputs [1/0]) degree_l=le-data. fit-toansform (imputs ['degree') isbutz [combown] = combown] Create model poredict. _ relocation model= tree. Decision Tree Classifier() rabor Train model - fit (inputs, target) model. score (inputs, danget) model. spredict ([[2,2,0]]) moitoitissol) Marsificotion Foror Index = I - Max 50.4, 03,03 Assignment. In this file asing following columns build amodel to predict if person would survive or not. home sklegan, prepracessing import Labertneder Label Emeding tabel Emcoding sofers to converting the labels into a mumeric from so as to convert them the machine-readarble from. Machine learning