Q.B P.7-I

9.1	What is learning? Explain diffe	rent types
	of learning with example.	, to
=> .	body of une had at part lies to	y S
	- Mochine learning is a category of	AI. In
work!	machine learning computers has the	he ability
0 7 5 1	to learn themselves , explicitions	aramming
	is not executived as the market	3
Toe .	Machine learning ofocuses on the	tidy and
, N	development of alanithment of that	rean learn
a las	development of algorithms dethat	dictions on
0000	framoridata and also make pre	The state of
	diddal gradt more los so got	
, ,	Data John Sandaria	4 <i>i</i>
1 6	Computer to B	param
71 -13	Output Computer to Pr	onh
		C.
	- Different types of learning in 1991	(2
a that has	place and primary best veget of	· (=
	Supervised Learning . H 'x' bu	901
. ! .	=> In this type of learning were	ise data
Lable	which tois comprises of input on	doucorresponding
	output.	01
3		n have
24	1 +2 11/0	white come
-	XX Y) (partiment)	New Input
L	TIP-2 01P-2	1
IF.	Learning Learning	Model
4	IP-n op-n	
		Julput Y
1		1

I - r.9 8.2

given an observation 'xhigher new observation 7x' it will try to And out what is T. Theoresponding o'Y's: Enimost soidsom filled For eng., Consider scenario where you have to arbuildigimage recloseifert to differentiate between cats and adogs in has to If you reed with a adapast snot dogs and cats dabelledtininges to the algorithm, - withe machine will bearn to classify between a dog or not from these labelled images. - When we input new dog or not images that it has never seen before, it will use the dog or cot. 2) Unsupervised Learning To unsupervised learning you are only given input 'x' , there is no label to the data and given the dola or different data points, your may want to form clusters or want to find some pattern. interest Learning algorithm Input-n gainross Y - 1 (23 1-01)



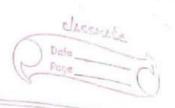
- For e.g., consider that you have dataset that contains information about the purchases you made from the shop. Through clustering, the algorithm can group the same purchasing behaviour among you and other customers which reveals potential customers without predefined labels balaibara silabam at 3) Reinforcement Learning. = In reinforcement learning you have an agent who is acting in an environment and you want to findout what actions the agent must take based on the reward or penalty that the agent gets the In this an agent seeks to learn the optimal actions to take based on evertone outcomes of pact actions. For e.g., AI agent to play game like chess. 4) Semi-supervised learning => It is a combination of supervised and unsupervised learning. In this there is some amount of labeled training data and also you have large amount of unlabeled data and you day to come up with some learning algorithm that convertereven when training idada is not labeled. For eng is long-rage translation modelyalon priningly after the standard



Explain the following terms with the help of appropriate graphics . a) Biagingio att. privataula de mit. .= Bias is defined or the Inability of the model because of that there is some difference or error occuring between the model's predicted value and actual e) Reinforcement trampa. (e -Bias is a systematic export that occur due to wrong accumption in the machine dearning process of book was bon is stand Longo with baptiction and to be one at this of oget relets to lear the Bigs in a social series made soil some follow to the the the cher (i) Semi-superior leaving Lan poisson is Actualian : it is values political bus coguerous balade to bausine aroust of results of the AT . (mus b) s. Variance nou sir dan ajob principale > Variance is the measure of apread in data From its mean position. In ML, variance is the amount by which the performance of a predictive imodel changes when it is trained or different subsets of the training data.



			Data
	Model Data	m	del 1 6
	model	had be the state of	1.
S. Aw	on the under Exting	11 12 12 12	1
1s bo			9
110	Vie Cristian Control	THE POST OF SHADE	· of ob
· - lab	رد داره مدومهدي مة	ranklab antiti	1 -abalt -
	.lahaa	00100001 90	idopra
labori	(other pointing data)	(new	dota)
040	5 4. 1.7 ton 2006	additionly act	4 (0
	10	- Aguana	11900
c	d Over Fitting + + Arabal	· Nadalaha	0 01: 17
⇒.	- A statistical mode	to removald to	be overfitted,
	When we train it		C/
	When model gets		/ - / / - /
	dota, it starts		
	and inaccurate da		
	Then the model		
	data correctly, b	ecause of too	mony details
	and noise.	LLE	Land and
	To a nutshell,	Overfitting - 1	tigh voriance
	and low bias.		
		1	^ "
	X	X/0 0	XX/ODO
	XX	X/. 0 0	X X 0 0 0
	V 000 V	00	× 0000
	$^{\wedge}$ \times $^{\circ}$ \times $^{\wedge}$	XOXXO	000
	X X X X X X X	X O O	X C COO
	$\times \times \times \rightarrow$	× × × × ×	(X)
	,	,	,
	(Under- fixthing)	Appropriate -	Over -
		Cil (too)	Y1111



d) Under Fitting => A statistical model or a machine learning algorithm is said to have under fitting when it connot apture the underlying trend of - Under Fitting destroys the accuracy of data our machine learning model. It's occurence simply means that our model or the algorithm does not fit the dola well enough. - In a nutshell, undersitting - High blace bather and down variance for los som as this most stop boom and learning from the

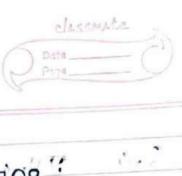


Q.3 Discuss how to evoluate a ML model for over fitting and under fitting i explain using diagram What measures needs to be taken in mue of overfitting and underfitting. => Under Fitting A stastical model or ML algorithm is said to have under fitting when it cannot capture the underlying toend of the data. Under Fitting dectroys the accuracy of our ML model. - In a nutshell, underfitting - High bias and low variance. · Overfitting =) A statistical model is said to be overfitted, When we train it with a lot of data. - When a model gets strained with so much of data, it starts learning from the noise and inaccurate data entries in our dataset. - Then model does not rategorize the data correctly, because of too many details and - In a nutshell, Overfitting - High variance inanda low biasis in inter plant - To evaluate whether a model is overfitting or underfitting , we typically use learning CHANES. (3) A learning curve plats. the model's performonre on the training and validation

datagets as a function of the number of

classmate Pose

training epoches the size of training set Under hitting batting a sol of bins si labora la Epoches A down on this Earlypath the labora the mostoppings 16 ant seimpator han soch isham anti-Measures to address over Fitting ! monitor) Rejularization => Apply regularization techniques such as 12 and 12 regularization to penalize large coefficients in the model. or underlitting our typically use leproin 2) Cross-validation Divse techniques like k-fold cross-volidation to ensure that the model generalizes of well different subsets of the data.



· Measures to address underfitting! - month signife o soil hosses tosible) Increase Model Complexity! I Use more complex model that can bapture the underlying patterns in the data. the dougast and pathod boad ante buil 2) Reduce regularization stone atob = The regularization is too strong, it can contains the model too much intending to under fitting no consider reducing the regularisation white strengthers sidt sciminion et bosse adjusting the model parameters. The model in simple linear appropriate is appresented of d + XD = 'Y



Thustoate the process of learning with the 0.4 gradient descent for a simple linear regression using bell shaped error curve. In simple linear regression, the goal is to find the best-fitting line through the data points, which iminimizes the error between the predicted values and actual to dies the model the mich sailors to Gradient descent is an optimization dechnique used to minimize this error by iteratively adjusting the model parameters. The model in simple linear regression is represented as -Y' = ax + b y'= predicted value, a = slope of the line, b = interrept X = input feature The error function typically used is the Mean Squared Error (MSE). $MSE = \frac{1}{n} \left[\frac{2}{3} \left(\frac{3}{3} - \frac{1}{3} \right)^{2} \right]$ Where, n = no. of data points

91 = predicted value

actual value



Gradient descent aims to find the values of a and b that minimizes the MSE.

The update rules for the parameters using gradient descent are:

 $\Rightarrow a = a - \alpha \cdot \frac{0 \text{ MSE}}{2a}$

 $b = b - \alpha \cdot \delta MSE$

where, d is learning rate

Bell Shaped Errox (urre Tragine plotting the error (MSE) as a function of the model parameters a andb

The error surface might look like a

bell-shaped curve.

The minimum point on this curve represents
the optimal values of a and b, where

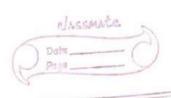
- Gradient descent moves the parameters iteratively downhill an this surface until

the minimum point is reached.

(MSE)

in this point

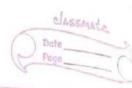
Parameters (a,b)



8.5 Explain the steps of developing ML opplications 2) Collection of Dota 211- ct mile > You could collect the samples from website and extracting data. . antoloroni From Rsc Feod or an API . or Fromordevices to mollect wind speed measurement Publicly available data + xan 2) Preparation of the input data Donce you have the input data, you need to check whether it's in a useable format in our dot over exercise and deline - Some algorithms can taccept target il variables and Features as string; some need than Hos beimintégéres ton son voy it Some algorithms accepts Features in special format. niopo testo bas prito sono 3) Analyse the input data time! - > > Looking at the dota you have passed in a dextoreditornoto check collections and preparation of input idata isteps are properly working and you don't have a bunch of empty values. - Plotting data in 1,2,0x3 dimensions can also help. 4) The impostance of this step is that it

makes you understand that you don't have

any garbage ralue coming in.



5) Trains the algorithm =) Good clean data From the first two steps is given to the algorithm. The algorithm extracts the information or knowledge. . . stab postars has This knowledge is mostly stored in formal that is readily usable by machine for next 2 stepsotob aldolino ploildia -6) Test thei algorithm => Tri this step the information learned in the previous istep is tused. Is to - When you are checking an algorithm, you ati satted who boild at the tipe the tipe Works an properly for not some has - IF you are not satisfied with result, you can again go back to step 4, change some thing and test again . I man 3) Use it who is it so propose (8) =) To this step real programis developed to do some task and once agan it is checked if all the previous steps worked 10 1951 you expected. ... how pointed Platting date in 1.0.00 3 dimensions can thook not took hosters how not so

primon sulpa sen

9.6	Exploin issues in ML
=>-	it has now took by tragent took set 12
	Issues in ML:
	highes objection on mile
	D. Poor quality of data -
	=) Data plays a eignificant role in ML process
	one of the significant pasues that ML
	professionals face is the absence of good
40,	quality dota. all to and al sint to
.10	Unclean and noisy dota can make the
	whole mineral and and expansion
1	whole process extremely exhausting.
	We don't want our algorithm to make
-45	inaccurate or faulty predictions.
	Hence the quality of data is essential to
	embance the output man time the
	and att revitab observations box
	2) Under Fitting of training data light
	=> This process occurs when data is unable
	to establish an accurate relationship
	between input and output variables.
	- It signifies the data is too simple to
	establish a precise relationship.
	3) Over Fitting of training data
	-> Overfitting refers to machine learning model
	trained with a massive amount of data
	that negatively affects its performance.
	- This means algorithm is trained with
	noisy and biased data, which will affect
	110



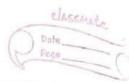
4) Lade of training data => The most important task you need to do in the ML process is to train data 1 to achieve an accurate output. - Less amount of training data early produce inoccurate or too blased predictions IM foot mornings and for and. 5) Slow simplementation => This is one of the common issues fared by machine learning professionals. - The machine learning models are highly efficient in providing accurate results, but it takes tremendous amount of - Hence the quality of data is mit or til of - Further, it requires constant monitoring and maintainance to deliver the best output iskab printed to pritted about 6 2) Cyc. Lidding of the days (2) equitively attended its perferoneme



9.7 Write short note on applications of ML. Applications of MI: H is tubors some at sol transport oute 1) nImages Recognition en lander => Image recognition is one tof the most common application of machine learning. It is used to identify objects, persons, places it digital timages retc. 2012 The populari ruse case of image recognition and face detection is Automatic Friend -tagging suggestions priving -7192. (A) etolilago entistra teom ant 70 and (= + 2) Traffic prediction 192 2: 1M 30 DITTE interwant ito visit new place, we take help of Google Maps , which shows us the correct paths with the shortest and it predicts the itraffic conditions + It, predicts they traffic conditions such as whether traffic is cleared inslow-moving or heavily congested with the help of two ways: i) Real time location of vehicle from google 30 compape and sensors point flight Average time has taken on part days in to oriation the war estimen witingoos - Speech reception is a process of 3) Product ire commondations 3 ML' is widely used by various re-commerce

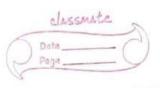
and entertainment companies such as Amazon,

Netflix. retention product recommendations.



to the inverce about had alient - Whenever we search some product on. amazon, then we started getting on advertisement for the same product while internet surfing on the same browsex and this is because of ML. . As similar, when we use Metflix, we find some recommendations for entertainent series, movies, etc. and this is also done with the helpsoform Lowgog out Lasi a itamotus si noit 4) Self-driving corsidered application of ML is self-driving cars? on - Min plays ia significant role in self-driving tope delle de service de la contrata del contrata de la contrata de la contrata del contrata de la contrata del contrata de la contrata de la contrata del contrata de la contrata del contrata - Teslo, the most popular can manufacturing company is working on self-driving cors. - It is using unsupervised learning method to in train the cor models to detect peoples and a objects while driving 5) Speech Recognition > While using Google, we get an option of search by voice 1 1t comes under speach recognition , and it's popular application of ML. - speech recognition is a process of converting voice instruction into the text and it is also known as "speech to text Google assistant, siri and Alexa are using

speech recognition technology.



0.8	C	0 15 15 5 1
9.0	Compare decision to	ee classification and
	with logistic regress	ion classification.
5)	i had awadan a of	phairing 14
	Decision Hree	Logistic Regression
*	ingniclassification of	notification
	dertreat	0 01 19/12/2015
	1) A decision tree is a	1) Logistic regression is
	type of supervised	also type of supervised
	learning algorithm that	learning algorithm used
- C-	is commonly used in	for classification tasks
	ML to model and predict	where the goal is
	outcomes based on	to predict probability that
	input data.) = -	an instance belongs to a
	input data.) = -	given class or not.
•	2) Highly interpretable,	2) Coefficients can be
77	os the tree structure	interpreted to understand
	clearly shows the decisions	the influence of each
yo ho	mode at each node.	Feature on the outcome.
		3) Less proone to overfitting
		compared to decision
(3.)	inis at a Combine in	
	4) IX is probugt to	4) It is majorly offect
1 (**	inoise 1	ablighter and
	abon and tils	حاددادا في ا
. 2.	4) It is majorly offerled	4) It is robust to
Ue .	by noise	· noise.
,	5) Can be trained on	s) Requires a lorge enough
	small training set.	training datasel.
netto.	عدد المستعدد منوا	
	250000 110 2200	po bohustolicib



9.9. Explain the properties of GINITIALEX. The Gini index is a measure used in decision trees to evaluate quality of a aplit. It grantifies the degree of impurity or disorder in a dataset. The ginis used in classification and Regiession Tree (CART) The dotacet Trantains example from classes p gini index p gini (T) is défined -i loos - the avadle whom boo labore at IM outcomes based on to predict probability that, port enoughini (T) about - E (p) the tugoi In the above equation it represents tools somether trelative : Frequency of class join ? After splitting To in two subsets To and - some Tall withingstres : No and No region index estitione cof-seplit rodala isistiliano et anos (e gini split (T) = $\frac{N_1}{N}$ gini (T₂) - The attribute with smallest girisple (T) is selected to split the node. The Ginitindex ranges from 0 to 0.5. denons : molicates perfect pourity is meaning all . + 0.5 represents maximum impulaty, indicating that the binstances are uniformly distal buted across all classes.



Q.10 Discuss in brief pruning is Decision tree 0 Pruning in decision tree is a technique used to reduce the size of a decision tree by removing sections of tree that provide little to no additional predictive power. The main good of pruning is to prevent overfitting and improve the model's ability to generalize to unseen data. - Decision tree can grow very longe and complex, capturing noise and outliers in the training data, which leads to overfitting. - Pruning helps this mitigate by simplifying the tree. - By removing unnecessary branches, pruning reduces the complexity of the model, leading to better performance on new, unseen data There are two main types of pring in decision trees:) Pre-pruning This involves halting the growth of the tree, before it becomes too complex. - It sets a limit on tree depth, minimum samples per leaf, or minimum information gain required for a split. 2) Post-pruning

=> This involves first allowing the tree to grow fully and then pruning back unnecessary branches.



The pring techniques are :-2) Cost - Complexity Pruning This method gets assigns a poice to each subtreet primarily based on its accuracy and complexity , then celects the subtree with the ablowest of eeing to mare land seceived S6 in the mi 2) Reduced Error Bruning => Removes the branches that do not eigni-Discontly affect the overall accuracy: · Advantages:- (1) moxa 1) Simplifies the model, improving its generalization to new data. 2) Roults in Smaller Imore interpretable tree. 1817 2,400 178 35 3) Reduces the size of the model, leading to Factor predictions Disadvantages :- 0128 1) IF prining is too agressive, it may remove important branches, leading to a model that is too simple to capture the 2) Not suitable for predictions of continuous aftribute. 3) Computationally expensive to train



D II	10												30
0.11	140	blem	ba	sed	0	- Si	mil	e l	line	No. IT	0	_	1
=	Following table shows the midterm and final exam grades obtained for student in							200					
	Follo	wing	tabl	e she	SWC	th	P -	mid	1+00	m.	11 1		/
	exe	am g	rades	oob	tair	red !	5	or s	stuic	lont	409	17	اهر
100	da-	tobase	· cou	v S P	Úse	7/4	he	me	tho	dia	5 1	<u>a</u>	(
24.10	240	4107	USING	MED	260	51011	-)	YEA	10-	11		
	80	al ex	am –		rad	٠,	R.L.	9.5	thd	ent	The s	_	\
	rec	ceived	8	6 ir	1	00	mic	1 4	erm	0	· W	20	/
	1		- 2	pp:nv	G.	e Ores	2	Ás.	14.5	, 9	G		/
Jasin	Mid	termi m(x)	72 5	0 81	74	94	86	وي	83	86	33	Do	1
	exo	m(x)	10:	oko o	. H		₉ -3-3	n d	百千	ינפו	7	30	8
	Fice	al	84 5	3 77	78	90	75	49	79	77	52	71.	1
	exa	m(x)	A section					500	o to	651		1	20
=)				191							2		
lasa	: امی	- enim	cact out	, 1950	C 64	1.44		-14		2.	1		
				stab	W.	3/1	rie	-	S 1-4	15			
-14-	ho on i	×	· 3 mm	XX				53	1 100	3	(2)		
		72	84	604	8	218	34		3 Y 2				
enibo	91.	.50		265	50	25	00	100	Ser D	. (e.		
		81	77	623	7	.65 6	5.1						
		74	78	577	12	54	36						
		94	90	846	56	88	36		5.4-1-	10		4	
		86	75	645		73							
Poor!	+:	5,9	49	289		34	18		3 7	. /			
- bi	enth	83	79	6 55	7	683	89	3	1 31%				
Section.	200	. 86	77	662	2	7			1260	Co			
		33	52	171	6	. lo	89	را ا	ob v	iJ.			
deliter .	3,	881	67.4	651		. 7:			1-10	(
		81	30	72		65			1.70				
-	10-101	887	878	.672	05.	691	113	500	v+50	Ç	2		



The equation for the regression line is-The thirth y al Fabax +ob - non and an animal tagest alto the an astate ad this are is in a Einghizxy - EX. EY n 5x2 - (5x)2 = 12 (67205) - 887 × 878 12(69113) - (887)2 1: 90 Hamoro . 65 mg . 62 3/13 rax 7/1092 34000 240092 532 box (2/2 a 2x) batrograp P strongs nottal (2878 × 28.010.65 × 887) batcomate 12/02 wollsy 1. 16 am = 25/12 wolls " bedroams VUZ 509 01 : 0 65 X + 25:12 The final grade of a student who received 86 in the mid term exam, 41 = 86 × 0.65 + 25.12 1. y' - 81.02



9.12	Proble	em based	on dec	ision do	cee digarith.				
	Suppose we want ID3 to decide whether the								
	car will be stolen or not. The darget classification is " nor - is stolen?" which non by the or No. 12 - 12 - 12 - 12 - 12 - 12 - 12 - 12								
	Car No.	h	@ Types	Oxigin	Stolen?				
	1	Red	Sports.	Domestic					
	3	Red	sports Sports	Domestic	No				
	4	Red Yellows	Sports	Domestic					
	5	Yellow	Sports	Imposted					
1	(88)	2 yellow?	535UV 1	Imported	No				
	7	Yellow	50121	Imported					
	8	Yellow	SUV						
	10	Red CIRad +	Sports :	Imported Imported					
			<u> </u>						
19 m Do	odus toobide of to shore load all								
		desir evan	time all		\$				
	21.20 + 59.9×93 = 14								
			20·12	· 'V ·	7				