Total No. of Questions: 6

Total No. of Printed Pages:3

Enrollment	No
------------	----



Faculty of Engineering End Sem (Odd) Examination Dec-2022 CB3EL01 Machine Learning

Programme: B.Tech. Branch/Specialisation: CSBS

Duration: 3 Hrs. Maximum Marks: 60

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d.

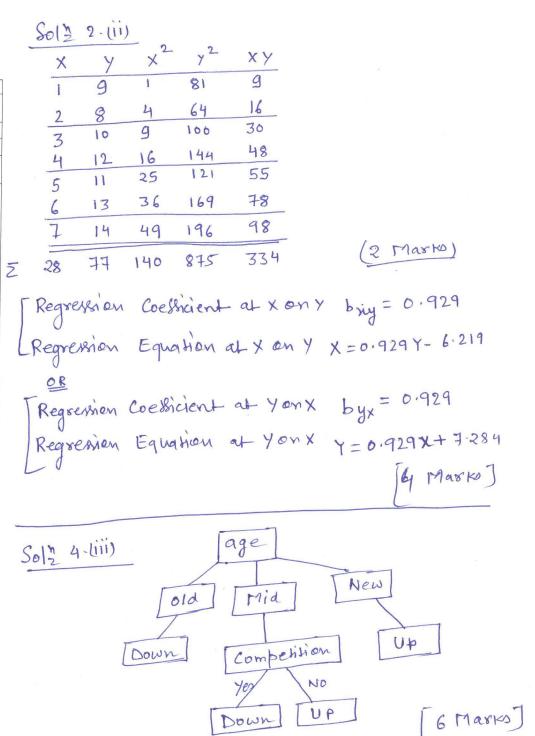
Q.1 (1	MCQs)	should be written in full instead of only a, b, c or d.	
Q.1	i.	ML is a field of AI consisting of learning algorithms that? (a) Improve their performance (b) At executing some task (c) Over time with experience (d) All of these	1
	ii.	In the mathematical equation of Linear Regression Y = β 1 + β 2X + ϵ , (β 1, β 2) respectively refers to	1
		(a) (X-intercept, Slope) (b) (Slope, X-Intercept)	
		(c) (Y-Intercept, Slope) (d) (slope, Y-Intercept)	
	iii.	The sample of data used to provide an unbiased evaluation of a model fit on the training dataset while tuning model hyperparameters-	1
		(a) Complete dataset (b) Test dataset	
		(c) Validation dataset (d) None of these	
	iv.	A is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test	1
		data for which the true values are known.	
		(a) Dataset (b) Confusion matrix (c) Classifier (d) Normal of those	
	* 7	(c) Classifier (d) None of these Which of the following are true about begging?	1
	V.	Which of the following are true about bagging? (a) In bagging, we choose random subsamples of the input points with replacement	1
		(b) If we use decision trees that have one sample point per leaf, bagging never gives lower training error than one ordinary decision tree	
		(c) Both options correct	
		(d) None of these	

P.T.O.

	vi.	Select the correct statements related to Naïve Bayes (NB) classifier.	1					
		(a) NB has low variance and high bias						
		(b) NB has low variance and low bias						
		(c) NB has high variance and low bias						
		(d) NB has high variance and high bias						
	vii.	Select the incorrect statements related to "support vector machine"	1					
		with "linear and nonlinear classification".						
		(a) SVM can be used for linear classification						
		(b) SVM can efficiently perform non-linear classification						
		(c) SVM can not perform non-linear classification						
		(d) SVM uses kernel trick for non-linear classification						
	viii.	How does the state of the process is described in HMM?	1					
		(a) Literal						
		(b) Single random variable						
		(c) Single discrete random variable						
		(d) None of these						
	ix.	Select the correct statements related to kNN classifier.	1					
		(a) kNN has high variance and low bias						
		(b) kNN has high variance and high bias						
		(c) kNN has low variance and low bias						
		(d) kNN has low variance and high bias						
	х.	Cluster is a collection of-						
		(a) Similar types (b) Non-similar types	1					
		(c) Linear types (d) Non-linear types						
Q.2	i.	How does the learning rate affect the performance of a gradient	4					
		descent algorithm?						
	ii.	Calculate the regression coefficient and obtain the lines of regression	6					
		for the following data:						
		X 1 2 3 4 5 6 7						
		Y 9 8 10 12 11 13 14						
OR	iii.	Can we have overfitting and underfitting at the same time? Explain.	6					
		How do you overcome overfitting and underfitting problems?						
		and the first of the control of the						
Q.3	i.	Discuss methods of evaluation for a machine learning algorithm.	4					
(ii.	In k-fold cross-validation, what happens when it is too big or too	6					
	,	small? How should one decide what value of K to assume for the	-					
		dataset?						

OR	iii.	•	engineering is importechniques used for fe		_	6
Q.4	i. ii.	Discuss the	e bayes classifier with major drawbacks d how it can be correc	of K-nearest ne	eighbour learning	4 6
OR	iii.	_	sion tree for the g		•	6
			ake PROFIT as the ta			
		AGE	COMPETITION	TYPE	PROFIT	
		Old	Yes	S/W	Down	
		Old	No	S/W	Down	
		Old	No	H/W	Down	
		Mid	Yes	S/W	Down	
		Mid	Yes	H/W	Down	
		Mid	No	H/W	Up	
		Mid	No	S/W	Up	
		New	Yes	S/W	Up	
		New	No	H/W	Up	
		New	No	S/W	Up	
Q.5	i.	Discuss asso	ciation rule mining in	detail		4
Q.J	ii.		em can be solved wit		model? What are	6
	11.	the limitation		ii iiiddeii iviaikov	model: What are	U
OR	iii.		applications of seque	ence classification	9	6
OK	111.	what are the	applications of seque	nee classification	•	U
Q.6		Write short n	notes on any two:			
	i.	Ward's algor	rithm			5
	ii.	K-nearest ne	ighbour clustering alg	gorithm		5
	iii.	Expectation-	maximization algorith	nm		5
			****	:		

		one decide	what value of K to assume for	or the dataset?		T			
OR	iii	Why Feature Engineering is important in model building and lists out some of the							
	_	techniques i	techniques used for Feature Engineering?						
Q.4	i.		ve Bayes Classifier with an			4			
	ii.	Discuss the	major drawbacks of K-neare	est Neighbour lea	rning Algorithm and how it	3+3			
0.0		can be corre	ected?						
OR	111	Design de	cision tree for the given	dataset showing	ng each step properly:	6			
		(Take PROFIT as the target feature(attribute))							
		AGE	COMPETITION	TYPE	PROFIT				
		Old	Yes	S/W	Down				
		Old	No	S/W	Down				
		Old	No	H/W	Down				
		Mid	Yes	S/W	Down				
		Mid	Yes	H/W	Down				
		Mid	No	H/W	Up				
		Mid	No	S/W	Up				
		New	Yes	S/W	Up				
		New	No	H/W	Up				
		New	No	S/W	Up				
					-				
Q.5	i.	Discuss asso	ciation rule mining in detail.			4			
	ii.	Which probl	em can be solved with hidde	n Markov model	? What are the limitations	3+3			
OR		of HMM?	II. i						
OK	111	What are the applications of sequence classification? Explain conditional random fields.							
Q.6			notes on any two:						
		Ward's Algo				5			
			eighbour Algorithm			5			
	iii	Expectation-	Maximization Algorithm			5			



Scheme of Marking

Total No. of Que	estions: 6	Total No. of Printed Pages:2			
		Enrollment No			
ر د در میشانده آندایشورد .	Fa	aculty of Engineering			
601-C4	End Sem (Odd) Examination Dec-2022				
UNIVERSITY Knowledge is Power	Machi	ne Learning-CB3EL01(T)			
	Programme: B.Tech.	Branch/Specialisation:CSBS			
Duration: 3 Hrs.		Maximum Marks: 60			

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d.

Q.1	i)	ML is a field of AI consisting of learning algorithms that?	1
		(a) Improve their performance (b) At executing some task	36
		(c) Over time with experience (d) All of the above	4.3
	ii)	In the mathematical Equation of Linear Regression $Y = \beta 1 + \beta 2X + \epsilon$, $(\beta 1, \beta 2)$ refers to	11
		(a) (X-intercept, Slope) (b) (Slope, X-Intercept)	
		(c) (Y-Intercept, Slope) (d) (slope, Y-Intercept)	
	iii	The sample of data used to provide an unbiased evaluation of a model fit on the training dataset while tuning model hyperparameters	1
		(a) Complete Dataset (b) Test Dataset (d) None of these	
	iv	A is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known. (a) Dataset	1
		(c) Classifier (d) None of these	
	v)		1
		(a) In bagging, we choose random subsamples of the input points with replacement	
		(b) If we use decision trees that have one sample point per leaf, bagging never gives	
		lower training error than one ordinary decision tree	
		(c) Both options correct	

	(d) Non	e of the	above							
vi				nents rel	ated to 1	Naïve Ba	yes (NE	3) Classi	fier	1
1	/									
	(a) NB	has low	varian	ce and h	igh bia	S				
vii	Select tl	ne incorr	ect stat	ements i	elated to	"Suppo	rt vecto	r machin	e" with "Linear and	1
**					Clarea .	o orpr				
	Nomine	di Ciuso.	nounc.	1						
	(a) SVN	A can be	used fo	or linear	classific	ation				
							sificatio	n		
:								9		1
VII	How do	oes the s	tate or	tne proce	ess is de	scribed i	n miviivi			1
	\ I !	1							*	
-					iable					
						1	1.79			-
ix	Select th	he correc	et state	nents rel	ated to	kNN clas	ssifier.			1
	/									
	(c) kNI	N has lov	v varia	nce and	low bias					
	(d) kN	N has lo	w varia	nce and	high bia	S				
x)	Cluster	is a colle	ection o	of:						1
	/									
	(a) Sim	ilar Tvr	es	(b) Non	-similar	Types				
	(c) Liii	cai rype		(4) 11011		JPG				
	Han d	aga tha	Lagrani	na moto	offoot	the part	Cormono	of a	Gradient Descent	4
1.			iearm	ng rate	arrect	the peri	ormance	e 01 a	Gradient Descent	1
ii.	Calcula	te the r	egressi	on coef	ficient	and obta	in the	lines of	regression for the	6
	followi	ng data:								
	X	1	2	3	4	5	6	7		
	Y	9	8	10	12	11	13	14		
iii	Can we	have o	verfitti	ng and i	ınderfitt	ing at th	e same	time? E	xplain. How do you	3+
111	overcor	ne overf	itting a	nd under	fitting r	roblems	9		1	
1 1	OVCICUI	HC UVCH	itting a	iid diidei	iittiiig į	rootems	•			-
										-
i.			s of ev	aluation	for a ma	chine lea	arning a	lgorithm		4
	vii vii ix x) i.	vi Select the select t	vi Select the correct (a) NB has low (b) NB has low (c) NB has high (d) NB has high (d) NB has high vi Select the incorr Nonlinear classi (a) SVM can be (b) SVM can ref (c) SVM can no (d) SVM uses ke vi How does the s a) Literal b) Single randor c) Single discrete d) None of the no ix Select the correct (a) kNN has high (c) kNN has low (d) kNN has low (d) kNN has low (d) kNN has low (e) Linear Type (i) How does the algorithm? ii. Calculate the note of the	(a) NB has low varian (b) NB has low variane (c) NB has high variane (d) NB has high variane (d) NB has high variane vii Select the incorrect state Nonlinear classification (a) SVM can be used for (b) SVM can efficiently (c) SVM can not perfor (d) SVM uses kernel tr vii How does the state of (a) Literal b) Single random variate (Single discrete rand (d) None of the mention ix Select the correct states (a) kNN has high variate (c) kNN has low variate (d) kNN has low variate (d) kNN has low variate (c) Linear Types (c) Linear Types i. How does the learnity algorithm? ii. Calculate the regressification of the correct states (a) Similar Types (c) Linear Types (c) Linear Types	vi Select the correct statements rel (a) NB has low variance and ho (b) NB has low variance and ho (c) NB has high variance and ho (d) NB has high variance and ho vi Select the incorrect statements in Nonlinear classification" (a) SVM can be used for linear (b) SVM can efficiently perform (c) SVM can not perform non (d) SVM uses kernel trick for not how does the state of the process. vi How does the state of the process. a) Literal b) Single random variable co Single discrete random variable co Single discrete random variable co Single discrete random variance and (b) kNN has high variance and (c) kNN has high variance and (d) kNN has low	vi Select the correct statements related to be a NB has low variance and low bias (c) NB has high variance and low bias (d) NB has high variance and high bias (d) NB has high variance and high bias vi Select the incorrect statements related to Nonlinear classification" (a) SVM can be used for linear classific (b) SVM can efficiently perform non-linear (d) SVM uses kernel trick for non-linear (d) SVM uses kernel trick for non-linear b) Single random variable c) Single discrete random variable d) None of the mentioned ix Select the correct statements related to b) kNN has high variance and low b) (b) kNN has high variance and low b) (c) kNN has low variance and high bia) (c) kNN has low variance and high bia) Cluster is a collection of: (a) Similar Types (b) Non-similar (c) Linear Types (d) Non-linear T i. How does the learning rate affect algorithm? ii. Calculate the regression coefficient affollowing data: X 1 2 3 4 4 7 9 8 10 12	(a) NB has low variance and high bias (b) NB has low variance and low bias (c) NB has high variance and low bias (d) NB has high variance and high bias vi Select the incorrect statements related to "Suppor Nonlinear classification" (a) SVM can be used for linear classification (b) SVM can efficiently perform non-linear classification (d) SVM uses kernel trick for non-linear classification vii How does the state of the process is described in a) Literal b) Single random variable c) Single discrete random variable d) None of the mentioned ix Select the correct statements related to kNN class (b) kNN has high variance and low bias (b) kNN has high variance and low bias (c) kNN has low variance and high bias (c) kNN has low variance and high bias (d) kNN has low variance and high bias (c) Linear Types (d) Non-similar Types (c) Linear Types (d) Non-linear Types i. How does the learning rate affect the perfalgorithm? ii. Calculate the regression coefficient and obtation following data: X 1 2 3 4 5 5 7 9 8 10 12 11 11 11 11 11 11	vi Select the correct statements related to Naïve Bayes (NE (a) NB has low variance and high bias (b) NB has low variance and low bias (c) NB has high variance and low bias (d) NB has high variance and high bias vi Select the incorrect statements related to "Support vector Nonlinear classification" (a) SVM can be used for linear classification (b) SVM can efficiently perform non-linear classification (c) SVM can not perform non-linear classification (d) SVM uses kernel trick for non-linear classification vi How does the state of the process is described in HMM a) Literal b) Single random variable c) Single discrete random variable d) None of the mentioned ix Select the correct statements related to kNN classifier. (a) kNN has high variance and low bias (b) kNN has high variance and high bias (c) kNN has low variance and high bias (d) kNN has low variance and high bias x) Cluster is a collection of: (a) Similar Types (b) Non-similar Types (c) Linear Types (d) Non-linear Types i. How does the learning rate affect the performance algorithm? ii. Calculate the regression coefficient and obtain the following data: X 1 2 3 4 5 6 7 9 8 10 12 11 13	vi Select the correct statements related to Naïve Bayes (NB) Classi (a) NB has low variance and low bias (b) NB has low variance and low bias (c) NB has high variance and low bias (d) NB has high variance and high bias vi Select the incorrect statements related to "Support vector machin Nonlinear classification" (a) SVM can be used for linear classification (b) SVM can efficiently perform non-linear classification (c) SVM can not perform non-linear classification (d) SVM uses kernel trick for non-linear classification vi How does the state of the process is described in HMM? a) Literal b) Single random variable c) Single discrete random variable d) None of the mentioned ix Select the correct statements related to kNN classifier. (a) kNN has high variance and low bias (b) kNN has high variance and high bias (c) kNN has low variance and high bias (d) kNN has low variance and high bias x) Cluster is a collection of: (a) Similar Types (b) Non-similar Types (c) Linear Types (d) Non-linear Types i. How does the learning rate affect the performance of a algorithm? ii. Calculate the regression coefficient and obtain the lines of following data: X 1 2 3 4 5 6 7 7 7 9 8 10 12 11 13 14 14 14 14 14 14	vi Select the correct statements related to Naïve Bayes (NB) Classifier (a) NB has low variance and high bias (b) NB has high variance and low bias (c) NB has high variance and low bias (d) NB has high variance and high bias vi Select the incorrect statements related to "Support vector machine" with "Linear and Nonlinear classification" (a) SVM can be used for linear classification (b) SVM can efficiently perform non-linear classification (c) SVM can not perform non-linear classification (d) SVM uses kernel trick for non-linear classification vi How does the state of the process is described in HMM? a) Literal b) Single random variable (c) Single discrete random variable (d) None of the mentioned ix Select the correct statements related to kNN classifier. (a) kNN has high variance and low bias (b) kNN has high variance and high bias (c) kNN has low variance and high bias (c) kNN has low variance and high bias (c) kNN has low variance and high bias (d) kNN has low variance and high bias (e) Linear Types (f) Linear Types (g) Linear Types (h) Non-similar Types (c) Linear Types (d) Non-linear Types i. How does the learning rate affect the performance of a Gradient Descent algorithm? ii. Calculate the regression coefficient and obtain the lines of regression for the following data: X 1 2 3 4 5 6 7