

Refer to table given below. Here, male is denoted with numeric value 0 and female with 1. Find in which class of people, Angelina will lie whose  $k$  factor is 3 and age is 5.

6

Name	Age	Gender	Class of Sports
Ajay	32	0	Football
Mark	40	0	Neither
Sara	16	1	Cricket
Zaira	34	1	Cricket
Sachin	55	0	Neither
Rahul	40	0	Cricket
Pooja	20	1	Neither
Smith	15	0	Cricket
Laxmi	55	1	Football
Michael	15	0	Football

Suppose you tracked the weather conditions for 14 days and based on the weather conditions, you decided whether to play golf or not play golf. Draw the frequency table and conditional probability table and on the basis of that decide whether to play golf or not.

6

Outlook	Temperature	Humidity	Windy	Play?
sunny	hot	high	FALSE	no
sunny	hot	high	TRUE	no
overcast	hot	high	FALSE	yes
rainy	mild	high	FALSE	yes
rainy	cool	normal	FALSE	yes
rainy	cool	normal	TRUE	no
overcast	cool	normal	TRUE	yes
sunny	mild	high	FALSE	no
sunny	cool	normal	FALSE	yes
rainy	mild	normal	FALSE	yes
sunny	mild	normal	TRUE	yes
overcast	mild	high	TRUE	yes
overcast	hot	normal	FALSE	yes
rainy	mild	high	TRUE	no

Given an  $n$ -character word, we want to predict which character would be the  $n+1$ th character in the sequence. For example, our input is "predictio" (which is a 9 character word) and we have to predict what would be the 10th character. Which neural network architecture would be suitable to complete this task and How?

5

8. Briefly discuss the idea of a support vector machine. What will happen upon choosing a very small value of  $C$  ( $\sim 0$ ) in the SVM?

5



# PUNJAB ENGINEERING COLLEGE

## Mid-Term Examination

Programme: B.Tech

Course Name: Machine Learning for Data Science

Maximum Marks: 20

Note: All questions are compulsory.

Year/Semester: 21221

Course Code: DSC 303

Time allowed: 1 Hour

Date: 4<sup>th</sup> October, 2021

Q. No.		Marks												
1. (a)	<p>The values of independent variable <math>X</math> and dependent value <math>Y</math> are given below:</p> <table><tr><td><math>X</math></td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td></tr><tr><td><math>Y</math></td><td>2</td><td>3</td><td>5</td><td>4</td><td>6</td></tr></table> <p>a) Find the least square regression line <math>Y=aX+b</math> b) Estimate the value of <math>Y</math> when <math>X</math> is 10</p>	$X$	0	1	2	3	4	$Y$	2	3	5	4	6	3
$X$	0	1	2	3	4									
$Y$	2	3	5	4	6									
(b)	<p>Explain the sum of squared error. Why SSE is not considered an optimal model evaluation metric?</p>	2												
2.	<p>Justify (with Yes or No) the following statements along with a valid reason:</p> <p>a) In least-square linear regression problem, adding regularization can decrease the error of the solution on the training data. b) As the number of training examples goes to infinity, your model trained on that data will have high variance and high bias.</p>	4												
3.	<p>Take a small size dataset example and explain <u>dummy variable</u> and one-hot encoding method for handling categorical data for model learning.</p>	6												
4. (a)	<p>What is overfitting and why is it a problem in machine learning? How can we avoid overfitting?</p>	2.5												
(b)	<p>Consider that you are building a spam filter to distinguish between genuine e-mails and unwanted spam e-mails. Assuming spam to be the positive class, which among the following would be more important to optimise and why?</p> <p>a) Precision b) Recall</p>	2.5												



**PUNJAB ENGINEERING COLLEGE**  
**End-Term Examination**

Programme: B.Tech  
Course Name: Machine Learning for Data Science  
Maximum Marks: 50

Year: 4<sup>th</sup>  
Course Code: DSC 303  
Time allowed: 02 Hours  
Date: 8<sup>th</sup> Dec, 2021

Note: All questions are compulsory.

Q. No.		Marks																																				
1.	<p>The sales of a company (in million dollars) for each year are shown in the table below:</p> <table border="1"> <thead> <tr> <th>x (year)</th><th>2005</th><th>2006</th><th>2007</th><th>2008</th><th>2009</th></tr> </thead> <tbody> <tr> <td>y (sales)</td><td>12</td><td>19</td><td>29</td><td>37</td><td>45</td></tr> </tbody> </table> <p>(i) Find the least square regression line <math>y = ax + b</math>.            (ii) Use the least squares regression line as a model to estimate the sales of the company in 2012.</p>	x (year)	2005	2006	2007	2008	2009	y (sales)	12	19	29	37	45	4																								
x (year)	2005	2006	2007	2008	2009																																	
y (sales)	12	19	29	37	45																																	
2.	<p>The Back-Propagation (BP) algorithm is often used for training feed-forward neural networks. Why do we need to calculate the gradient in the BP algorithm?</p> <p>When a BP algorithm is used, the error function must be differentiable. Why?</p>	4 2																																				
3.	<p>Why should we use CNN? Explain. Further, let's take a 6 X 6 grayscale image.</p> <table border="1"> <tbody> <tr><td>3</td><td>0</td><td>1</td><td>2</td><td>7</td><td>4</td></tr> <tr><td>1</td><td>5</td><td>8</td><td>9</td><td>3</td><td>1</td></tr> <tr><td>2</td><td>7</td><td>2</td><td>5</td><td>1</td><td>3</td></tr> <tr><td>0</td><td>1</td><td>3</td><td>1</td><td>7</td><td>8</td></tr> <tr><td>4</td><td>2</td><td>1</td><td>6</td><td>2</td><td>8</td></tr> <tr><td>2</td><td>4</td><td>5</td><td>2</td><td>3</td><td>9</td></tr> </tbody> </table> <p>where, the Filter is 3 X 3, i.e.,</p> <p>1 0 -1            1 0 -1            1 0 -1</p> <p>Now, Convolve the entire image and draw the output image.</p>	3	0	1	2	7	4	1	5	8	9	3	1	2	7	2	5	1	3	0	1	3	1	7	8	4	2	1	6	2	8	2	4	5	2	3	9	6
3	0	1	2	7	4																																	
1	5	8	9	3	1																																	
2	7	2	5	1	3																																	
0	1	3	1	7	8																																	
4	2	1	6	2	8																																	
2	4	5	2	3	9																																	
4.	<p>What improvements have been done to RNN? Also, explain the architecture of LSTM?</p> <p>b) Answer the following precisely with respect to KNN algorithm:</p> <p>(i) How to handle categorical variables in KNN?            (ii) How to find best K value?            (iii) Why KNN is non-parametric?</p>	6 6																																				





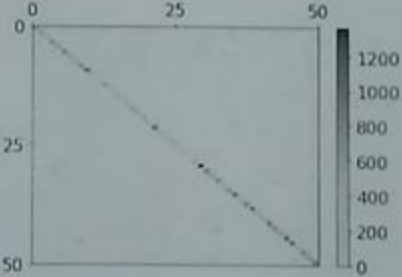
**PUNJAB ENGINEERING COLLEGE**  
**(Deemed to be University)**  
**Mid-Term Examination**  
**(Session 2019-20, Semester 19202)**

Programme: B.Tech (CSE)  
 Course Name: Machine Learning  
 Maximum Marks: 30

Year/Semester: 2<sup>nd</sup>/4<sup>th</sup>  
 Course Code: CSN-211  
 Time allowed: 1.5 Hours

**Notes:**

- All questions are compulsory.
- The candidates, before starting to write the solutions, should please check the question paper for any discrepancy, and also ensure that they have been delivered the question paper of right course code.

Q. No.		Marks
1. a)	Assume you are preparing a data frame for a supervised learning task and you notice that the target label classes are highly imbalanced and multiple feature columns contain missing values. The proportion of missing values across the entire data frame is less than 5%. Which is best option from following to minimize bias due to missing values? Give comment for each option. (i) Replace each missing value by the mean or median across non-missing values in same row. (ii) Delete observations that contain missing values because these represent less than 5% of the data. (iii) Replace each missing value by the mean or median across non-missing values in the same column. (iv) For each feature, approximate the missing values using supervised learning based on other features.	04
b)	What is a hyperparameter and why is it needed in training and learning process? What is the risk with tuning hyperparameters using a test dataset?	04
2. a)	Assume you are given data $\{(x^1, y^1), \dots, (x^n, y^n)\}$ and you are planning to train an SVM. You define a kernel $k$ and obtain a kernel matrix $K$ presented in figure below, where $K_{ij} = k(x^i, x^j)$ . 	06
	(i) What is the issue here? (ii) How can we address this issue? (iii) Why is it important to scale the inputs when using SVMs?	
b)	A data scientist runs a principal component analysis on given data and observes that the percentage of variance explained by the first 3 components is 80%. How this percentage of variance explained is computed? Outline a broader process for same.	04
3.	Given regression (linear or logistic) and naïve bayes classifier, differentiate the purpose of using these classifiers for a given dataset. Which of these two is considered to be high bias/low variance classifier? Justify with appropriate reason.	06
4.	With respect to different feature subset selection strategies, answer following precisely: (i) State the feature subset selection problem mathematically. (ii) How do the objective function in <i>filters</i> and <i>wrappers</i> differ? (iii) Analyze the primary motivation behind Bidirectional search.	06