



**NAMIBIA UNIVERSITY  
OF SCIENCE AND TECHNOLOGY**

**FACULTY OF COMPUTING AND INFORMATICS**

**DEPARTMENT OF INFORMATICS**

<b>QUALIFICATION: Bachelor of Informatics Honours (with specialisations in Web Informatics and Business Informatics)</b>	
<b>QUALIFICATION CODE: 08BIFH/08BIHB</b>	<b>COURSE LEVEL: NQF LEVEL 8</b>
<b>COURSE: Data Science and Analytics</b>	<b>COURSE CODE: DSA821S</b>
<b>DATE: NOVEMBER 2022</b>	<b>SESSION: 1</b>
<b>DURATION: 2 Hours</b>	<b>MARKS: 60</b>

<b>FIRST OPPORTUNITY EXAMINATION QUESTION PAPER</b>	
<b>EXAMINER(S):</b>	<b>DR LAMECK MBANGULA AMUGONGO</b>
<b>MODERATOR (S):</b>	<b>MS EMILIA SHIKEENGA</b>

**THIS EXAMINATION PAPER CONSISTS OF 5 PAGES  
(INCLUDING THIS FRONT PAGE)**

**Instructions for the students**

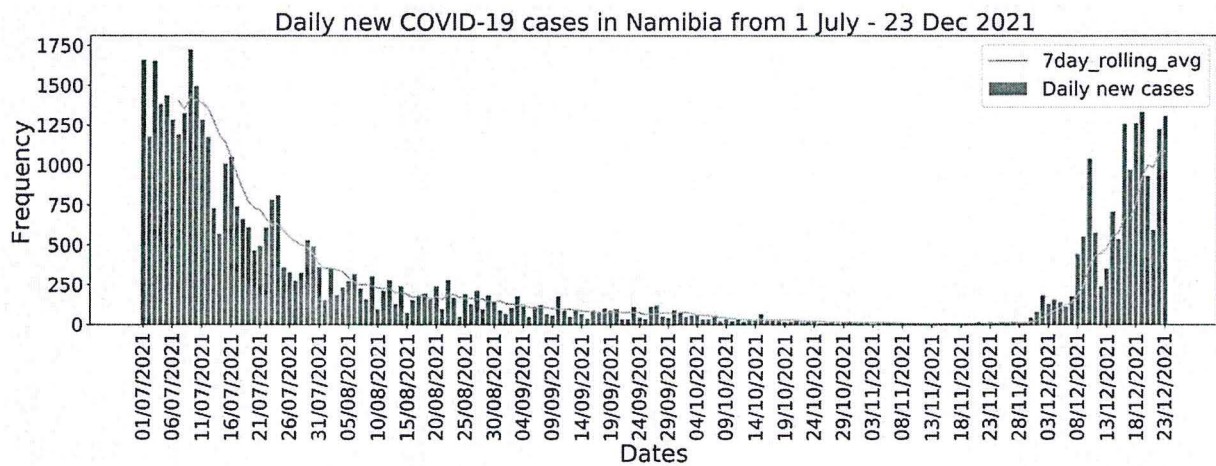
- 1. Answer ALL the questions.**
- 2. Write clearly and neatly.**
- 3. Number the answers clearly.**



**Question 1: Short questions**

**[18]**

- a) **True or False:** When performing unsupervised learning we know the number of clusters beforehand. **[1]**
- b) **True or False:** Big data is initially characterized by 3 Vs: Volume, Veracity and Variety? **[1]**
- c) **True or False:** Data Science is the same as Data analytics. **[1]**
- d) **True or False:** Business can utilise insights from data to maintain competitive advantage. **[1]**
- e) **True or False:** The below Figure is an example of Inferential statistics? **[1]**



- f) **True or False:** If two variables X and Y are correlated, then we must be able to specify the cause i.e, X is the cause or Y is the cause. **[1]**
- g) **True or False:** In a classification problem statement after training followed by testing, we get an accuracy of 99.7%, we can necessarily conclude that it is a good model. **[1]**
- h) Which of the following are correct about Activation Functions in neural network? **[2]**
  - a. Derivative of a sigmoid activation function  $g(z)$  is  $g(z)[1-g(z)]$
  - b. Derivative of a hyperbolic tangent activation  $k(z)$  function is  $1-(k(z))^2$
  - c. Derivative of a leaky RELU Activation function  $h(z)$  is 1
  - d. Derivative of RELU activation function  $l(z)$  is 0 for  $z < 0$
- i) Choose the correct option for residuals in Linear regression? **[1]**
  - a. Residuals are horizontal offset, and the sum of residuals varies between  $[0,1]$
  - b. Residuals are horizontal offset, and the sum of residuals can be unity.
  - c. Residuals are vertical offset, and the sum of residuals is always unity.
  - d. Residuals are vertical offset, and the sum of residuals is always zero.
- j) Which of the following are correct related to the Confusion Matrix? **[2]**



- a. Confusion matrix is always a square matrix
  - b. Confusion matrix is a way to judge our classification model
  - c. Diagonal entries in a confusion matrix may be zero or non-zero
  - d. Confusion matrix is a symmetric matrix
- k) Which of the following statements are correct for Support Vector Machines (SVM)? [1]
- a. A support vector machine is a machine learning algorithm that analyses data for both classification and regression analysis.
  - b. SVM is an unsupervised learning method.
  - c. An SVM finds the hyperplane which is having the largest margin value.
  - d. SVMs are used in text categorization, image classification recognition, etc.
- l) Which is not a deep learning method: [1]
- a. Learning rate Decay.
  - b. Dropout.
  - c. Training from scratch.
  - d. Bootstrapping.
  - e. Transfer Learning.
- m) If we have a date column in your dataset, then how will you perform Feature Engineering using Python? **Hint:** A date column, has lots of important features such as: day of the week, day of the month, day of the quarter, and day of the year etc. [4]



**Question 2: Apriori algorithm**

**[16]**

A table has five transactions. Let the minimum support (min sup) = 60% and min confidence (conf) = 80%.

ItemID	Items_bought
F100	{Bread, Egg, Milk, Butter, Honey, Sugar}
F101	{Cereal, Egg, Milk, Butter, Honey, Sugar}
F102	{Bread, Bacon, Butter, Honey}
F103	{Bread, Jam, Cookie, Butter, Sugar}
F104	{Cookie, Egg, Egg, Butter, Cucumber, Honey}

- a) Find all frequent item sets using Apriori algorithm. **[10]**
- b) List all the strong association rules **(with support and confidence)**. **[6]**

**Question 3: Classification**

**[16]**

1. The table below illustrates the prediction for a model to predict Bankruptcy. Based the test set, calculate the evaluation measures.

No	Target	Prediction	No	Target	Prediction	No	Target	Prediction
1	Bankruptcy	Bankruptcy	8	No Bankruptcy	No Bankruptcy	15	Bankruptcy	Bankruptcy
2	Bankruptcy	Bankruptcy	9	Bankruptcy	Bankruptcy	16	No Bankruptcy	No Bankruptcy
3	Bankruptcy	Bankruptcy	10	Bankruptcy	Bankruptcy	17	Bankruptcy	No Bankruptcy
4	Bankruptcy	Bankruptcy	11	Bankruptcy	Bankruptcy	18	No Bankruptcy	Bankruptcy
5	No Bankruptcy	No Bankruptcy	12	No Bankruptcy	No Bankruptcy	19	No Bankruptcy	No Bankruptcy
6	Bankruptcy	Bankruptcy	13	Bankruptcy	Bankruptcy	20	No Bankruptcy	No Bankruptcy
7	No Bankruptcy	No Bankruptcy	14	No Bankruptcy	No Bankruptcy	21	No Bankruptcy	No Bankruptcy





- a) Complete the confusion matrix. [4]
- b) Compute the misclassification rate. [2]
- c) Compute F1-measure [4]

2. Consider the following 3-class confusion matrix:

	Predicted			
Actual		A	B	C
	A	25	5	2
	B	3	32	4
	C	1	0	15

- a) What is the overall accuracy? [2]
- b) What can you say about Recall and Sensitivity? [2]
- c) What is the precision for class A? [1]
- d) What is the specificity of class C? [1]

**Question 4: Linear Optimisation [10]**

Pick n Pay Oshakati during the festive season combines two products rice and potato to form a gift pack which must weigh 5 kg. At least 2 kg of rice and not more than 4 kg of potato should be used. The net profit contribution to the Pick n Pay is Namibian dollars 5 per kg for Rice and N\$ 6 per kg for potato. Formulate LP Model to find the optimal factor mix.

- a) Formulate the objective function. [3]
- b) Formulate Constraints. [3]
- c) Non-negative constraints. [1]
- d) Summarise the optimization problem. [3]

**END OF EXAM**



NAMIBIA  
UNIVERSITY  
OF SCIENCE AND  
TECHNOLOGY

P/Dag 13300  
Windhoek  
NAMIBIA

2022 -10- 18

*M. B. Lewis*

FACULTY OF COMPUTING & INFORMATICS  
DEPARTMENT: INFORMATICS