



PAMIBIA UNIVERSITY
OF SCIENCE AND TECHNOLOGY

FACULTY OF COMPUTING AND INFORMATICS

DEPARTMENT OF INFORMATICS

QUALIFICATION: Bachelor of Informatics Honours (with specialisations in Web Informatics and Business Informatics)	
QUALIFICATION CODE: 08BIFH/08BIHB	COURSE LEVEL: NQF LEVEL 8
COURSE: Data Science and Analytics	COURSE CODE: DSA821S
DATE: JANUARY 2023	SESSION: 2
DURATION: 2 Hours	MARKS: 60

SECOND OPPORTUNITY/SUPPLEMENTARY EXAMINATION QUESTION PAPER	
EXAMINER(S):	DR LAMECK MBANGULA AMUGONGO
MODERATOR (S):	MS EMILIA SHIKEENGA

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: Regression**[12]**

- 1) An Autohaus in Windhoek wanted to investigate how the price of one of its car models depreciates with age. The research department at the company took a sample of eight cars of this model and collected the following information on the ages (in years) and prices (in hundreds of Namibian dollars) of these cars.

Age	8	3	6	9	2	5	6	2
Price	38	220	95	33	267	134	112	245

- a. Find the regression line with price as a dependent variable and age as an independent variable. **[3]**
- b. Give a brief interpretation of the values of a and b calculated in part b. **[3]**
- c. Predict the price of a 7-year-old car of this model. **[3]**
- d. Estimate the price of an 18-year-old car of this model. **[3]**

Question 2: Association analysis

[15]

The following incomplete table summarises supermarket transaction data, where **Boerewors** refers to the transactions containing Boerewors, \sim **Boerewors** refers to the transactions that do not contain Boerewors, **pies** refer to the transactions containing pies, and \sim **pies** refers to the transactions that do not contain pies.

	Boerewors		\sim Boerewors	Total
Pies	2000		500	
\sim Pies	1000			
Total				5000

- a) Complete the table. [5]
- b) Assume that the association rule "**Boerewors** \Rightarrow **Pies**" is mined. Given a minimum support threshold of 25% and a minimum confidence of 50%, is this association rule strong? Support your answer with calculations. [5]
- c) Based on the given data, is the purchase of Boerewors independent of the purchase of Pies? If not, what kind of correlation relationship exists between the two? Show the **Lift** calculation. [5]

Question 3: Association analysis

[10]

A table has five transactions. Let the minimum support (min sup) = 60% and min confidence (conf) = 70%.

ItemID	Items_bought
F100	{Pork, Beans, Mutton, Beef, Salad}
F101	{Mutton, Beef, Milk, Salad, Lamb}
F102	{Lamb, Beef, Mutton, Salad}
F103	{Rice, Pap, Mutton, Beef, Lamb}
F104	{Rice, Pap, Pap, Beef, Salad, Beans}
F105	{Beans, Mutton, Beef, Pork, Lamb, Rice, Pap}

- 1) Find all frequent item sets using Apriori algorithm. [10]

Question 4: Classification**[13]**

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

- a) Complete the confusion matrix. [4]
 b) Compute the misclassification rate. [4]

2. Consider the following 3-class confusion matrix:

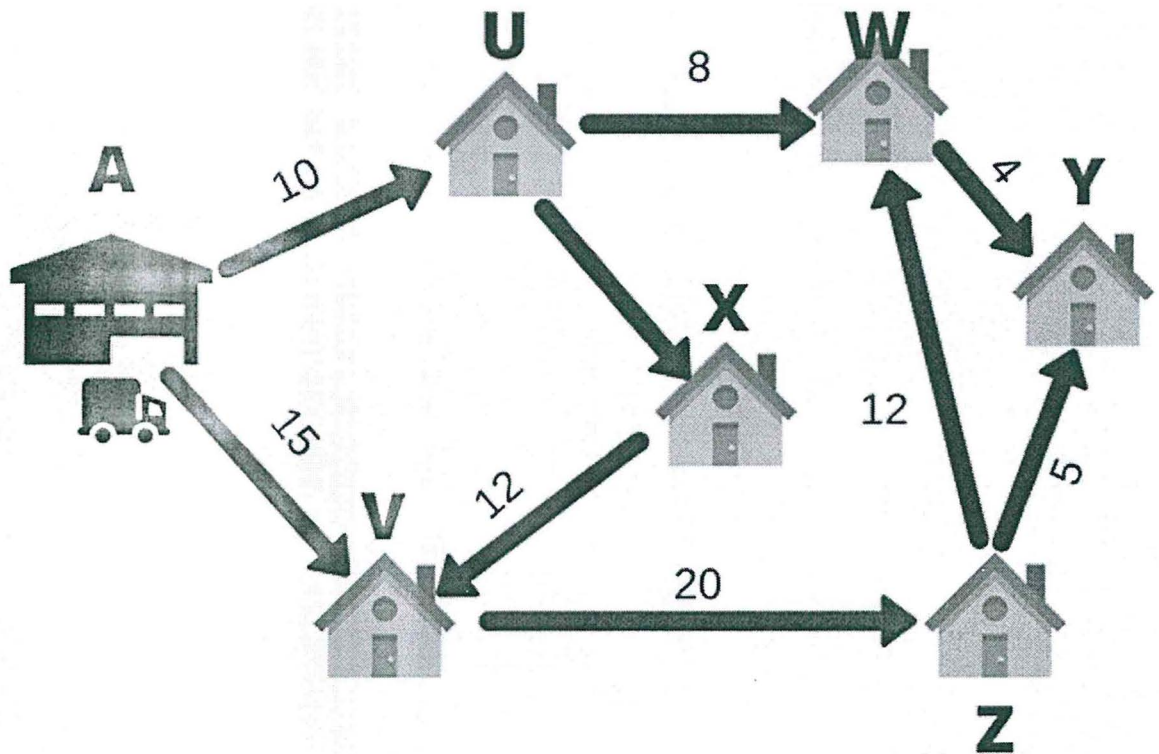
		Predicted		
		A	B	C
Actual	A	25	5	2
	B	3	32	4
	C	1	0	15

- 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]

Question 5: Linear optimisation

[10]

A new logistic company, has 6 packages to deliver in a day. The warehouse is located at point A. The 6 delivery destinations are given by U, V, W, X, Y, and Z. The numbers on the lines indicate the distance between the cities. To save on fuel and time the delivery person wants to take the shortest route.



- a) Compute different routes for going to all the 6 destinations and then come up with the shortest route.

END OF EXAM



NAMIBIA
UNIVERSITY
OF SCIENCE AND
TECHNOLOGY

P/Bag 13388
Windhoek
NAMIBIA

2022 -10- 18

M. J. ...

FACULTY OF COMPUTING & INFORMATICS
DEPARTMENT: INFORMATICS