NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

SBA4105

FACULTY OF APPLIED SCIENCES

DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH

SBA4105: CASE STUDIES IN BUSINESS ANALYTICS

DECEMBER 2024: EXAMINATION

Time : 3 hours

Candidates should attempt **ANY FIVE** questions (20 marks each).

CAREFULLY READ THE FOLLOWING CASE STUDIES BEFORE ATTEMPTING CORRESPONDING QUESTIONS.

A1. Call for consultants to Carry Out an End of Term Evaluation

Zimbabwe Environmental Law Association (ZELA), a public interest environmental law organisation that seeks to promote environmental justice, sustainable and equitable use of natural resources, democracy and good governance in the natural resources and environment sector, is seeking the services of Consultant, that will help the organisation review its strategic programmes and produce strategic reports, as well an achievements paper, through funding support from the Embassy of Sweden (Sida). The period under review for the Strategic programmes is 2023 to 2024. In addition, the Consultants will be required to evaluate the primary project that was funded by (Sida) and produce an end of project evaluation report. The project's goal was to strengthen community-level governance frameworks for improved food security, climate resilience and environmental sustainability in Binga, Gutu, and Gwanda districts in Zimbabwe, with two specific programming themes, Climate Change and Wildlife. The consultant(s) will work closely with the Organization's Monitoring, Evaluation, Accountability and Learning (MEAL) department in executing the assignment.

Objectives

-To establish the impact made by ZELA , for the 2020 to 2024 period, and key achievements of the organisation, including knowledge management, visibility and resource mobilisation strides, in line with its strategic plan.

-To gather key stakeholders feedback (including marginalised communities representatives, partners, duty bearers and staff), on the impact of the Sida funded project that was implemented by ZELA in Gutu district, Binga district and Gwanda district.

-To establish recommendations for sustaining the impact generated by ZELA under the 2020 and 2024 Strategic plan to strategically position for future work.

(a) State any one source of dataset to solve the problem at hand [1]

(b) Explain how you can determine the extent to which the programme has achieved its expected results at output and outcome levels after collecting data? [9]

(c) Explain any two visualisations you can include on your dashboard [4]

(d) Describe and explain how you communicate your results to the stakeholders [2]

(e) State any two tactics you can implement to create value and enhance the experience for your client [2]

(f) You are expected to implement structured problem solving methodology. Outline the steps. [6]

(g) What type of business consultancy is required? Justify your answer [2]

A2. Harare City Council Smart City Implementation Plan

Employing 6,000 people, Harare City Council delivers housing, water and transport services to 1.2 million citizens across the Zimbabwean capital. To keep the city moving, the council's traffic control center (TCC) works together with local transport operators to manage an extensive network of roads, tramways and bus lanes. Using operational data from the TCC, the council's roads and traffic department is responsible for predicting Harare's future transport requirements, and developing effective strategies to meet them. Harare City Council has a wide array of technology at its disposal. Traffic detecting sensors, rain gaugesand closed-circuit television (CCIV) cameras collect data from across Harare, and each of the city's 1,000 ZUPCO buses transmits a GPS update every 20 seconds.

In the past only a small proportion of this Big Data was available to controllers at Harare's TCC, reducing their ability to identify, anticipate and address the causes of traffic congestion. As Eng. Thusa, Head of Technical Services-Roads and Traffic Department at Harare City Council, explains: Previously, our TCC systems only offered a narrow window on the overall status of our transport network, for example, controllers could only view the status of individual bus routes. Our legacy systems were also unable to monitor the geospatial location of Harare's bus fleet, which further complicated the traffic control process. Because we couldn't see the whole transport network in real time, it was very difficult to identify traffic congestion in its early stages. This meant that the causes of delays had often moved on by the time our TCC operators were able to select the appropriate CCTV feed, making it hard to determine and mitigate the factors causing congestion. Harare City Council wanted to ease traffic congestion across Harare.

To achieve this, the council needed to find a way to integrate, process and visualize large amounts of structured and unstructured data from its network of sensor arrays-all in real time. To help develop a smarter approach to traffic control, Harare City Council entered into a research partnership with RIPTECH CS Research-Zimbabwe. Eng. Shoniwa, Research Manager-Smarter Urban Dynamics at RIPTECH CS Research,

comments: Smarter Cities are cities with the tools to extract actionable insights from massive amounts of constantly changing data, and deliver those insights instantly to decision-makers.

(a) Is there a strong case to make, that all large cities in Zimbabwe should adopt and start using big data analytics and related information technologies to help ease the traffic problems? Justify your answer. [5]

(b) How can big data analytics be used in policy making at national level using a bottom up approach, starting with local councils implementing start cities going up to ministries implementing smart government? [5]

(c) Design risk management model associated with implementation of smart cities. [10]

A3. Customer Churn Prediction for Financial Institutions Using Deep Learning Artificial Neural Networks in Zimbabwe

According to Ahmed et al (2017), customer churn prediction models done through machine learning techniques gained massive popularity in just a few decades ago. Similarly, so as the trend in the financial industry where predictions have been done to predict dissatisfied customers who were likely to go for other service providers primarily using traditional machine learning techniques such as Support Vector Machine (SMV) and Decision Trees (DTs). Globally, Farquad, Vadlamani, Ravi, Bapi and Raju (2014) states that the Germany N26 Bank and American Express from America successfully implemented customer churn prediction. The American Express now relies on sophisticated predictive modelling that forecasts and prevents customer churn. This was achieved through the analysis of past customer transactions hence identification of customer accounts that are likely going to close and take preventive action to mitigate the problem. Amuda and Adeyemo (2019), argues that the cost of retaining existing customers is lower than the cost of acquiring new customers when they did a customer churn prediction using multi-layer perceptron for financial institutions in Nigeria. In the African continent, International Financial Corporation (IFC) World Bank (2018) states that Greenfield Microfinance in Sub-Sahara Africa is a business model to advanced financial inclusion which successfully implemented churn predictions. The following was its customer churn rate in the year of implementation.

This led to a 43 percent of adults in this continent to have access of formal financial services in comparison with a 23 percent as per the year 2011. It also states that its subsidiaries in South Africa managed to incorporate customer churn prediction which led to a successful avoidance of customer attrition. In Zimbambe, customer churn prediction has been done for mobile network operators by Kusotera, Chimire and Mapuwei (2018) in their study titled An Analysis of Customer Churn for Mobile Network Operators in Zimbabwe. Upon observation no bank has engaged in customer attrition predictions yet. This study aims at developing a customer churn prediction model using a case of one of the leading financial institutions in Zimbabwe.

According to Amuda and Adeyemo (2019) tones and tones of data is being generated as a result of technological advancement hence data has become oil of the twenty first century however oil is just useless unless it is refined into fuel. Several techniques and methods were introduced for hidden meaningful and salient information to be extracted from this data. The techniques include traditional machine learning and deep learning artificial neural networks. Customer churn is a critical challenge for corporates as this entails loss of customers to competitors. An in advance prediction of customer churning behaviour is therefore crucial for high valuable insight extraction in order to retain and maximise a corporates customer base. In Zimbabwe, many if not all financial institutions have not yet implemented customer churn prediction hence failure to realise the benefits of customer retention at the expense of acquiring new customers. Therefore the purpose of this study is to develop a deep learning customer churn prediction model using customer past transactional data.

(a) What is the main objective from the case study [1]

(b) List any three specific objectives to for the highlighted aim and explain corresponding methods to achieve them. [9]

(c) State any four feature engineering techniques you will apply to enhance the dataset? [4]

(d) Explain how you will ensure higher model accuracy for churn customer prediction [6]

A4. Fake News Detection Using Machine Learning Approaches

The fake news on social media and various other media is wide spreading and is a matter of serious concern due to its ability to cause a lot of social and national damage with destructive impacts. Fake News contains misleading information that could be checked. This maintains lie about a certain statistic in a country or exaggerated cost of certain services for a country. which may arise unrest for some countries like in Arabic spring. There are organizations, like the House of Commons and the Crosscheck project, trying to deal with issues as confirming authors are accountable. However, their scope is so limited because they depend on human manual detection, in a globe with millions of articles either removed or being published every minute, this cannot be accountable or feasible manually. A solution could be, by the development of a system to provide a credible automated index scoring, or rating for credibility of different publishers, and news context.

A methodology to create a model that will detect if an article is authentic or fake based on its words, phrases, sources and titles, by applying supervised machine learning algorithms on an annotated (labeled) dataset, that are manually classified and guaranteed is proposed. Then, feature selection methods are applied to experiment and choose the best fit features to obtain the highest precision, according to confusion matrix results. We propose to create the model using different classification algorithms. The product model will test the unseen data, the results will be plotted, and accordingly, the product will be a model that detects and classifies fake articles and can be

used and integrated with any system for future use.

Suggestions has been made to start with an analysis of the research related to fake news detection and explores the traditional machine learning models to choose the best, in order to create a model of a product with supervised machine learning algorithm, that can classify fake news as true or false, by using tools like python scikit-learn, NLP for textual analysis. This process will result in feature extraction and vectorization; we propose using Python scikit-learn library to perform tokenization and feature extraction of text data, because this library contains useful tools like Count Vectorizer and Tiff Vectorizer. Then, we will perform feature selection methods, to experiment and choose the best fit features to obtain the highest precision, according to confusion matrix results.

(a) Explain the type of data to be collected to build the fake news classification model? [2]

(b) State any two libraries will you use to solve the problem at hand. [2]

(c) State the steps you will take to prepare data before building news classification model. [6]

(d) How can you extract detailed information on fake news with negative sentiments. [5]

(e) Briefly, explain how you will use data analytics techniques for the development of fake news detection model? [5]

A5. Minimum Spanning Tree for a Networking Company

A networking company is planning to connect several offices across different cities using a fiber-optic cable. The cost of laying the cable between any two cities is known. The company wants to minimize the total cost of connecting all cities, ensuring that there is a path between every pair of cities without forming any loops. The cities and the cost of connecting them are represented by an undirected, weighted graph where:

- Each city is a node.
- Each edge between nodes has a weight, representing the cost of laying the cable between those two cities.

Table 1: Cities and the costs of connecting them

City 1	City2	Cost (Weight)
A	B	4
A	C	6
B	C	6
B	D	3
C	D	8
C	E	5
D	E	7

The goal is to find the Minimum Spanning Tree (MST) for this graph, minimizing the cost to connect all cities.

(a) Why is it important to implement an algorithm to perform the above task? [5]

(b) Which algorithm is more appropriate for solving this problem? Justify your answer. [3]

(c) State two ways you can use to determine the complexity of this algorithm. [2]

(d) Write down an algorithm, to find the Minimum Spanning Tree (MST) for the given network to minimise the cost to connect all the cities. [10]

END OF QUESTION PAPER