UGANDA CHRISTIAN UNIVERSITY
A Centre of Excellence in the Heart of Africa

FACULTY OF ENGINEERING, DESIGN AND TECHNOLOGY
DEPARTMENT OF COMPUTING AND TECHNOLOGY
ADVENT 2025 SEMESTER **TEST**

PROGRAM: *[BSDS, BSCS]*

YEAR: 3 SEMESTER: 1

COURSE CODE: *[*DSC3108*]*

COURSE NAME: *[BIG DATA MINING AND ANALYTICS]*

EXAMINATION TYPE: *PROJECT-BASED TEST*

PROJECT DURATION: OCTOBER 2025

TIME ALLOWED: *[3 DAYS]*

Examination Instructions

1. The general Uganda Christian University examination guidelines and academic & financial policies apply to this examination. Violating any of the policies by the student automatically makes this examination attempt void, even if you have completed and submitted the answer booklet.
2. This test consists of a project to be executed in [3] days.
   i. Assessment of the project shall be based on five milestones, evaluated in the duration of the project. Each milestone has an assigned 20 marks as stated below.
   ii. To achieve all the milestones, students should submit their jupyter notebooks as links to github or kaggle under the students' personal accounts.
   iii. Students should also submit a powerpoint presentation on Moodle (To be presented on 16th October 2025).
3. Every student has a responsibility to prove their contribution towards every milestone, and marks will be awarded to every student individually.

Project-based assessment guidelines

S/N	Milestone Description	Maximum Marks
1	**MILESTONE ONE** [Ability to display datasets using meaningful visuals]	**20 %**
2	**MILESTONE TWO** [Ability to explore datasets and discern emerging patterns i.e. descriptive analytics]	**20 %**
3	**MILESTONE THREE** [Ability to respond and make comprehensive interpretations that can analytically respond to the pertinent research questions]	**20 %**
4	**MILESTONE FOUR** [Ability to generate data mining models that can make robust descriptions and predictions]	**20 %**
5	**MILESTONE FIVE** [Data Science communication]	**20 %**
	TOTAL MARKS	**100 %**

SECTION A: MILESTONE 3 [20 MARKS]

DATASET DESCRIPTION (*Question3.pdf*)

A review article titled "Building natural language processing tools for Runyakitara" (https://doi.org/10.1515/applirev-2020-2004)

1. Write a short essay describing the insights you derived from the Review article. **[10 MARKS]**

2. What recommendations would you make to the ethnic leaders of the Runyakitara speakers regarding the use of NLP models for language preservation. **[10 MARKS]**

SECTION B: MILESTONE 1,2,4 [60 MARKS]

Natural Language Processing Project: Sentiment Analysis in the Finance/ Health/ Education Domains

Context:
Generate a project on Sentiment Analysis targetted at any of the three domains of finance, health or education. You will be evaluated on your ability to achieve the following goals.

1. Data sourcing: Text data or social media posts with associated metadata and key words associated with the domain of interest. **[5 MARKS]**

2. Generation of research questions: Create guiding research questions or objectives key to the data you have sourced. **[10 MARKS]**

3. Data preprocessing and EDA: Illustrate your skills in transforming obtained data using the appropriate python libraries for NLP, and ability to answer posed research questions **[15 MARKS]**

4. Sentiment classification: Group the sentiments using an appropriate classification model. **[15 MARKS]**

5. Model evaluation: Test your model using the appropriate accuracy metrics. **[15 MARKS]**

SECTION C: MILESTONE 5 [20 MARKS]

Data Science Communication:

Context:
Illustrate your communication skills.

1. Generate powerpoint slides on the following
    a. Background of Sentiment Analysis. **[10 MARKS]**

    b. The summary of the project you worked on in Section B. **[10 MARKS]**

~END OF PROJECT TEST~