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 ASSIGNMENT

PROGRAM: *BACHELOR OF SCIENCE IN DATA SCIENCE*

YEAR: 2   SEMESTER: 1

COURSE NAME: *Data mining*

15TH SEPTEMBER

2025

In-class Assignment

Assignment Instructions

Outliers in Data Science

Definition and importance: Research the definition of an outlier in data science and its importance in data analysis. Explore potential sources of outliers and their impact on data validity.

Types of outliers: Investigate different types of outliers. Provide examples for each type.

Methods for detecting outliers: Study various visualization and statistical methods for outlier detection.

Impact of outliers: Analyze how outliers affect statistical measures like mean, variance, and correlation. Examine their impact on the performance of machine learning models.

Strategies for handling outliers: Research strategies for handling outliers. Consider factors influencing the decision-making process.

Univariate Data Relationships

Definition and characteristics: Research the definition and key characteristics of univariate data. Differentiate it from bivariate and multivariate data and explore different types of univariate data.

Methods for visualizing univariate data: Investigate graphical methods for visualizing univariate data. Analyze insights gained from each type.

Measures of central tendency and variability: Study measures of central tendency and variability in univariate analysis.

Descriptive and inferential methods: Differentiate between descriptive and inferential methods in univariate analysis and explore statistical tests for inferential analysis.

Bivariate Data Relationships

Definition and characteristics: Research the definition of bivariate data. Explore the concepts of explanatory and outcome variables.

Types of bivariate analysis: Investigate different types of bivariate analysis, including dependence and interdependence analysis. Understand the concepts of causation, covariation, and correlation.

Methods for analyzing bivariate data: Study techniques for analyzing relationships between qualitative nominal variables and quantitative variables. Understand the use of contingency tables, correlation analysis, and linear regression.

Graphical methods for Bivariate Data: Explore common graphical methods for visualizing bivariate data, such as scatter plots, box plots, and mosaic plots.

Multivariate Data Relationships

Definition and characteristics: Research the definition of multivariate data and its differences from univariate and bivariate data. Understand the importance of relationships among multiple variables.

Dependence vs. interdependence techniques: Differentiate between dependence and interdependence techniques in multivariate analysis. Provide examples for each type of technique.

Multivariate analysis techniques: Study techniques for analyzing relationships between these variables. Explain the purpose of each technique and provide examples of applications.

Visualization methods for multivariate data: Explore methods for visualizing multivariate data.