

DATA ANALYTICS FOR ARTIFICIAL INTELLIGENCE

Module Code: H9DAI

Module Level: Level 9

Programme

MSCAI_JAN24

Lecturer

Dr. Anu Sahni

Continuous Assessment

Datasets

Will be provided

Questions

1. Data characteristics (10 points)

Present the data statistics and characteristics in your dataset:

- Data population (3 points)
- Data attributes (4 points)
- Types of attributes (4 points)

2. Data Manipulation (20 points)

Apply 2 data manipulation techniques on your dataset.

- Technique dealing with Missing Values / Categorical Data where appropriate (10 points)
- Detecting and Removing Outliers (10 points)

3. Data Preprocessing (30 point)

Apply 3 data preprocessing techniques on your data:

- Scaling / Normalisation techniques where appropriate (10 points)

- Calculate feature correlations in your data (10 points)
- Apply 1 feature selection technique from the list below (10 points)
 - Univariate Feature Selection
 - Tree-based Feature Selection
 - Greedy Feature Selection

4. Data Visualisation (30 points)

Visualise your data using 3 graphs from the table below



- 1 graph from the top group: histogram, pie chart, stacked column chart, venn diagram (10 points)
- 1 graph from the middle group: scatter chart, line chart, multidimension column chart (10 points)
- 1 graph from the bottom group: heat map, bubble chart (10 points)

5. Data Summary (10 points)

Summarise the statistical analysis of your data and propose one or more machine learning algorithms that potentially perform well on your data.