

Advanced-Data Analytics-MCA472A

ETE-3 Data Analytics Project

Assignment Details: Individual Assignment- Experiential Learning

Submission Dead Line: 15 August 2024

This assignment aims to develop a comprehensive data analytics project using a real-world dataset. You will be required to preprocess the data, implement various algorithms, analyze the results, and compare the performance of the algorithms using suitable metrics. Additionally, you will document your findings in a research paper/ blog post (Medium)/ or an article for a magazine/ newspaper.

- Identify the domain and decide on the problem statement
- Collect or select a relevant real-world dataset. The dataset should be domain-specific and substantial enough to perform meaningful analysis.
- Clean and preprocess the data to handle missing values and outliers and perform necessary transformations.
- Implement at least five machine learning / deep learning algorithms or Integrated/Hybrid/Novel on the collected or selected dataset.
- Analyze the results of the algorithms using at least five suitable metrics (e.g., accuracy, precision, recall, F1 score, ROC-AUC, etc.).
- Compare the performance of the implemented algorithms based on the chosen metrics.
- Discuss the results, insights, and conclusions drawn from the analysis. Highlight any integrated, hybrid, or novel approaches used.
- Write a research paper detailing your approach, methodology, results, and conclusions./ Create a blog post on Medium summarizing your project and key findings./ Write an article suitable for publication in a newspaper or magazine.

Submission:

1. **Code:** Provide a GitHub repository link containing your code, dataset, and other relevant files.
2. **Outcome:**
 - a. **Research Paper:** Submit a PDF of your research paper.
 - b. **Blog Post:** Provide a link to your Medium blog post.
 - c. **Article:** Submit a PDF of the article for newspaper or magazine publication.

Additional Notes:

- Ensure your project is original and all sources are correctly cited.
- Follow ethical guidelines for data collection and usage.

Evaluation Rubrics -ETE - Component -3		
Rubrics	Rubrics	Marks
Individual (50 Marks)	R1-Topic Selection	02 Marks
	R2-Data set Collection/Selection	10 Marks
	R3-Preprocessing	05 Marks
	R4- Implementation	15 Marks
	R5-Comparison	08 Marks
	R6-Outcome	10 Marks