Course: Programming Techniques

Project

In this project, students will implement a solution to address a real-world problem. Even if the project follows a data science approach, the primary focus should be developing a solution rather than solely on the results. The data used for the project must be based on real-world data sources.

1. Project Deliverables:

- Report (students are required to submit the report in both .docx and PDF formats)
- Jupyter Notebook/Code
- Dataset(s) (including the source of these dataset(s))
- Students may also create a web application using Flask to present their results.

Please be vigilant about plagiarism, as ISEG Rules will be rigorously enforced.

2. Report Structure

The report should have the following structure:

I.   Introduction

In the introduction, students should set the context for the project. It is crucial to identify the overarching problem students intend to solve and the primary objective of the empirical work.

II.   Literature Review

The group should identify and summarize a select group of papers that can provide insights into similar work done by other authors. These papers should support the group's approach or decisions.

III.   Method

The method is the strategy to achieve the objectives defined in the introduction. If you propose a solution, you may base your method on the Design Science Approach (Aparicio et al., 2023). If you are adopting a data science approach, you should follow the CRISP-DM methodology (Costa & Aparicio, 2020, 2021).

IV.   Results

This section should outline what was accomplished and present key tables with results or a solution description.

V.   Discussion

The discussion involves comparing the proposed results with other findings presented in the literature review.

VI.   Conclusions

What was the primary purpose of the project? What conclusions can be drawn from both the literature and the empirical work?

References

Aparicio, J. T., Aparicio, M., & Costa, C. J. (2023). "Design Science in Information Systems and Computing." In Proceedings of the International Conference on Information Technology and Applications: ICITA 2022 (pp. 409-419). Singapore: Springer Nature Singapore.
Costa, C. J., & Aparicio, J. T. (2020). "POST-DS: A methodology to boost data science." In the 15th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.
Costa, C. J., & Aparicio, J. T. (2021). "A Methodology to Boost Data Science in the Context of COVID-19." In Advances in Parallel & Distributed Processing, and Applications: Proceedings from PDPTA'20, CSC'20, MSV'20, and GCC'20 (pp. 65-75). Springer International Publishing.

3. Jupyter Notebook/Code

The report should comprehensively detail the steps taken in the data analysis. The team is encouraged to explore techniques beyond those covered in class, but proficiency in these techniques is expected.

4. Datasets

Collecting original, interesting, and valuable data is challenging and highly valued in this project. Students must identify their data sources and provide a well-documented account of the data collection process. It is preferable not to select datasets already used in published papers to ensure a good grade.

5. Web App

Using Flask Creating a web application using Flask is not mandatory but can be considered as an additional feature.

6. Presentation Dates:

Scheduled according to the course timetable.