Project Milestone 7: Final Report and Presentation

Introduction

This project investigates data mining of demographic data to create one or more classification models capable of accurately identifying individuals who qualify for a specific service offering. The details of the project requirements are described in the **Project Outline** document. The project outline underscores the use of CRISP-DM as a structured approach to guide you through the data science project, emphasizing its importance in the context of real-world data analysis and modelling.

Outline

The project is divided into seven (7) milestones, where at each stage some deliverables are to be produced in terms of a series of reports that describe the project plan, the work carried out during the iterative process of data preparation, modelling, evaluation, and deployment.

The CRISP-DM methodology breaks down a data mining project life cycle into six phases with each phase consisting of many secondary tasks. This milestone addresses the *Deployment Phase* of CRISP-DM.

Deployment

Deployment is the sixth and final phase of the CRISP-DM methodology and comes after the evaluation phase. The models created during the modelling phase are not particularly useful unless the results are available for use by the intended audience. The solutions generated in the modelling phase are only valuable if they are put into action. The objective of the deployment phase is to take the solutions generated in the modelling phase and put them in an environment where they can be used to drive business value. This phase is critical to the success of the data mining process because without proper deployment, the business benefits of the project are not realised despite all the efforts that went into the project.

Tasks

This milestone consists of the following tasks that should be submitted as an assignment on given due date. The focus of this milestone are the last two tasks of the Deployment Phase.

Produce Final Report

- Summarise the project and your experiences.
- Produce a final report (in PDF or document format) detailing all project steps, key findings, ethical implications and recommendations for improvement.
- Include all the previous deliverables and summary of the results.
- Include a section on ethical analysis, detailing the ethical challenges encountered when handling sensitive information and how these were addressed.
- Prepare a final demo and presentation of the data mining model and results.

Review Project

- Summarise important experiences made during the project (what went well, what could have been better, and what improvements can be made in the future)
- Discuss your overall learning experience (what have you learned, how relevant are the knowledge and skills gained from this project, how can the projects be handled in future)

Ensure that all your code is well-organized according to the CRISP-DM steps, thoroughly documented, uses appropriate data science libraries and tools, runs without errors, produces correct results, and is optimized for efficiency and performance.

Deliverables:

- Final Project Report (in PDF format).
- Presentation (not more than 10 slides of content)
- Final project files and code (R Markdown) developed throughout the project.
- If the code is hosted on a cloud repository or version control platform like GitHub, GitLab, BitBucket and SourceForge, also provide the link to the code repository as part of the report.
- Project files and/or code for any other tools used during this assignment.

Grading Criteria:

Criteria	Weight
Introduction	5
Business Objectives	5
Data Mining Goals	5
Project Team Organisation	5
Data Exploration and Quality Assessment	10
Data Preparation	10
Modelling	10
Model Evaluation	10
Deployment and Maintenance	10
Ethical Considerations	5
Conclusion and Reflection	5
Project Files and Code	20
TOTAL	**100**

Additional Information

- This is a group assignment. Please continue with your project groups. Remember that a group must have at least 2 members but may not exceed four people.
- All work must be original. Copying another group's work will not be tolerated.
- Includes names of all group members on the Cover Page.
- Submit your project electronically on Moodle (BC Connect) before the due date.
- All writing must be correctly cited and referenced.
- **Plagiarism is a serious offence**. Belgium Campus uses software that can scan for plagiarism and a student caught doing this will get 0 for this assignment.
- No mark will be awarded if the assignment is not uploaded via BC Connect.
- Late assignments will not be accepted; missing the deadline is an automatic 0.