Big Data Course

Capstone Project   
Action Plan

For students (instructor’s review required)

ⓒ2023 SAMSUNG. All rights reserved.

Samsung Electronics Corporate Citizenship Office holds the copyright of this document.

This document is a literary property protected by copyright law so reprint and reproduction without permission are prohibited.

To use this document other than the curriculum of Samsung Innovation Campus, you must receive written consent from copyright holder.

|  |  |
| --- | --- |
| Course | Big Data Course |
| Team Name | No.1 |
| Team Leader/  Members | Nguyễn Phúc Cường/   Lê Thị Ngọc Ánh, Phạm Hùng Cường, Hà Văn Hậu, Trịnh Ngọc Khiêm, Văn Thanh Lân |
| Project Title | Building and Optimizing Machine Learning Models |
| Goal |  |
| - Successfully build three models to evaluate customer shopping habits  - Optimizing models to obtain the highest f1-score | |
| Abstract |  |
| Capture the main features of the dataset, preprocessing and tranform data to build the model with the highest f1-score. Compare the result and assess the best model | |
| Method |  |
| - Data analysis and selection  - Data and tranform data  - Models picking  - Build and train models  - Evaluate f1-score | |

|  |  |
| --- | --- |
| Data |  |
| - The data was acquired from instructor’s resource  - The processed data was used to build models | |
| Expected  Outcome |  |
| Implement selected models without bugs. The expected f1-score is over 0.85 and have the consistency. All members can learn how to construct models and optimize it | |
| Role by  Member |  |
| - Hậu, Khiêm: Build and optimize Random Forest Classifier model.  - Ánh, Hùng Cường: Build and optimize Support Vector Machine model.  - Lân, Phúc Cường: Build and optimize Logistic Regression model. | |

|  |  |
| --- | --- |
| Schedule  Summary |  |
| The schedule was detaily described in the WBS file. | |
| Comment &  Assessment |  |
| <Comment and assessment **by the instructor.**> | |