Pagalanan, Jon Enrico B.

BSCS-3B

Instructions: Write a 1–2 page essay reflecting on your project experience.

CSST – 102 “My First End-to-End ML Experience”

In this project, I learned how to build a complete machine learning pipeline from start to finish. I began by working with two datasets, one from 2023 and one from 2024, that contained different factors affecting the Happiness Score of countries. During this stage, I discovered that cleaning and preparing data is one of the most important parts of any project. I learned how to handle missing values, remove outliers, and combine datasets properly before training any models. I also realized that choosing the right target variable and features has a big impact on how well a model performs.

After preparing the data, I trained two machine learning models: Linear Regression and Decision Tree Regressor. This helped me understand how different algorithms can produce different results even when given the same data. I learned how to evaluate their performance using Mean Squared Error (MSE) and R^2 Score, which measure how accurate the models are. I also used visualizations such as scatter plots, heatmaps, and bar charts to see patterns and relationships more clearly. These visuals made it easier to explain my results and compare both models.

The most challenging part of the project was handling missing or inconsistent data. At first, my models gave errors because of NaN values and non-numeric columns. I solved this by filling missing numeric values with their mean and filling categorical ones with the mode. I also removed outliers using the Interquartile Range (IQR) method. Another challenge was understanding why one model performed better than the other. After testing and reading more about the algorithms, I learned that the Linear Regression model worked better because the data had mostly linear relationships.

Overall, this project gave me practical experience in how data becomes a working machine learning model. I learned how to clean, train, evaluate, and visualize data effectively. These skills can be applied in real-life situations, such as predicting housing prices, forecasting business trends, or analyzing customer satisfaction. This project not only improved my technical skills but also taught me the importance of patience, problem-solving, and clear thinking in data science.