Emmanuel ABABIO

ITAI 1371

Journal Report ML 1371

Learning Experience Analysis - Module 2

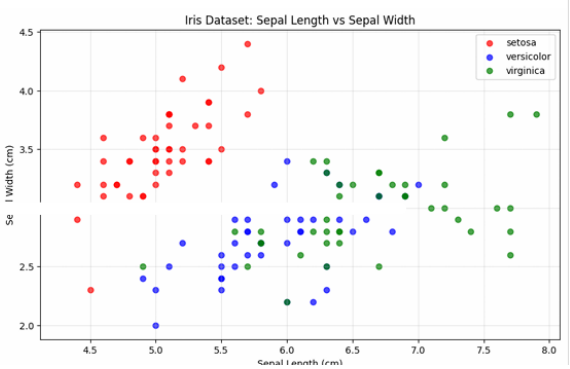
In this module, I began to understand how the various tools in machine learning interact with one another. At first, I thought programs like Jupyter, Google Colab, and VS Code were just different ways to type code. But while working through the exercises, I realized they also change how I think and work. Jupyter and Colab provided me with instant results, making it easier to test ideas step by step. VS Code felt more professional and showed me how real projects are built and organized. This made me understand that the environment I use shapes the way I solve problems.

The Python libraries also helped me see things differently. Pandas was confusing in the beginning, especially the indexing. But then I saw it is like Excel, only with more power. One line of code can do what would take many steps in Excel, and that made me appreciate how efficient it is. NumPy was also a big step forward. When I saw how arrays and vectorization work, I understood why they save time compared to writing loops. It showed me how machine learning needs speed and structure when working with large datasets.

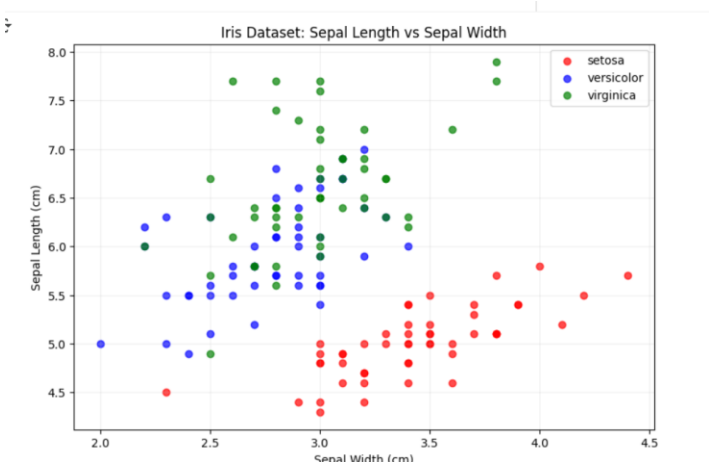
When working on the scatter plot, I noticed that one species could have three different widths for a single length. To make the code run, I had to change line 2 into `plt.figure(figsize=(10,6))`, which helped me display the graph correctly. Looking at the data, I found that the **Setosa** species had a minimum width of 3.0 and a maximum of 4.4. This made them smaller than the two other species. The **Versicolor** and **Virginica** species looked similar, each with a length-to-width ratio of about 5.5 to 6.8. They also seemed only slightly smaller than the Setosa. From this exercise, I learned some key lines of code for creating graphics and how to import data from a database. Even though I still need more practice with data manipulation and building graphs, I feel more confident about the process.

When trying out different changes in Part 3, I inverted the place of the width data with the length, and vice versa. I saw that the original graph turned upside down. This simple change showed me how the position of data on the axes can completely change the way we read and understand a graph. It made me realize how important it is to organize data carefully so that the visualization makes sense.

Original graph:



Modified graph:



At first, I thought using GitHub and Markdown was just extra work. But then I realized they are important because they make my projects clear and easy to share. Markdown lets me explain my work in the same notebook as my code, which makes it easier to understand. GitHub gives me a way to store my code, update it, and share it, instead of keeping files all over my computer. These skills feel important for working with others and also for tracking my own progress.

Overall, this module showed me that learning machine learning is not just about running algorithms. It is about building habits with the right tools, organizing my work, and understanding how each step in the process affects the next one. Data cleaning, visualization, and libraries like pandas and NumPy all connect to how accurate and useful a model can be. Even though I still need more training, this module helped me see how the tools work together and gave me more confidence to keep improving.