Section 6: AI Tools & LLMs

Your Name

April 10, 2025

Option A: Prompt Engineering

I used Claude to help me with creating machine learning models for the student performance dataset. Below are the prompts I used and the responses received.

My Prompt

I shared the StudentsPerformance.csv dataset with Claude and asked for assistance with the Python/ML task:

```
see for task 1 I have to do this Tasks:
1. Load the dataset and perform data cleaning.
2. Perform EDA using visualizations (use seaborn, matplotlib).
3. Predict whether a student will pass or fail based on scores and other features.
4. Use Logistic Regression or Random Forest for classification.
5. Evaluate the model using accuracy, confusion matrix, F1-score.
6. Add comments/markdown to explain the steps.
```

Claude's Response

Claude provided a detailed, step-by-step implementation of the machine learning pipeline, including the following key components:

[Insert screenshot of Step 1: Import Libraries here]

Figure 1: Step 1 - Import Libraries

[Insert screenshot of Step 2: Load and Examine the Dataset here]

Figure 2: Step 2 - Load and Examine the Dataset

For the model building part, Claude provided comprehensive code for both Logistic Regression and Random Forest classifiers:

[Insert screenshot of Step 7: Model Building - Logistic Regression (Part 1) here]

Figure 3: Step 7 - Model Building - Logistic Regression (Part 1)

[Insert screenshot of Step 7-8: Logistic Regression (Part 2) and Random Forest (Part 1) here]

Figure 4: Steps 7-8 - Logistic Regression (Part 2) and Random Forest (Part 1)

[Insert screenshot of Step 8-9: Random Forest (Part 2) and Model Comparison (Part 1) here]

Figure 5: Steps 8-9 - Random Forest (Part 2) and Model Comparison (Part 1)

[Insert screenshot of Step 9: Model Comparison and Feature Importance here]

Figure 6: Step 9 - Model Comparison and Feature Importance

My Reflection

Claude provided well-structured, comprehensive code for implementing both Logistic Regression and Random Forest models, complete with proper evaluation metrics. The AI did an excellent job explaining each step through comments and organizing the code in a logical sequence. I had to modify the code slightly to adapt to my specific dataset structure and column names. I also enhanced some visualizations to make them more informative for my presentation. Overall, the AI-generated code served as an excellent starting point that saved significant development time while ensuring all required evaluation metrics were properly implemented.