**Student Name:** Marion Jepkorir  
 **Course:** Software Development – AI & Machine Learning  
 **Institution:** Power Learn Project Academy  
 **Assignment Title:** AI Tools: Theory, Practice & Ethics  
 **Submission Date:** [15 june]

## **🔹 Part 1: Theoretical Understanding (40%)**

**1. Differentiate between TensorFlow and PyTorch.** *TensorFlow is more production-ready with strong deployment support, especially on mobile and web. PyTorch is more flexible and preferred for research because of its dynamic computation graph.*

**2. What are the primary uses of Jupyter Notebook in AI projects?** *Jupyter Notebook is used for writing, testing, and visualizing code in real-time. It supports data analysis, model building, and easy presentation in a single document.*

**3. Mention two key advantages of spaCy in NLP tasks.**

* *Fast and efficient for real-time NLP.*
* *Built-in support for Named Entity Recognition (NER) and syntactic parsing.*

**4. Compare Scikit-learn and TensorFlow. When should you use each?** *Use Scikit-learn for quick classical machine learning tasks (e.g., decision trees, logistic regression). Use TensorFlow for deep learning models like CNNs or RNNs.*

## **🔹 Part 2: Practical Implementation (50%)**

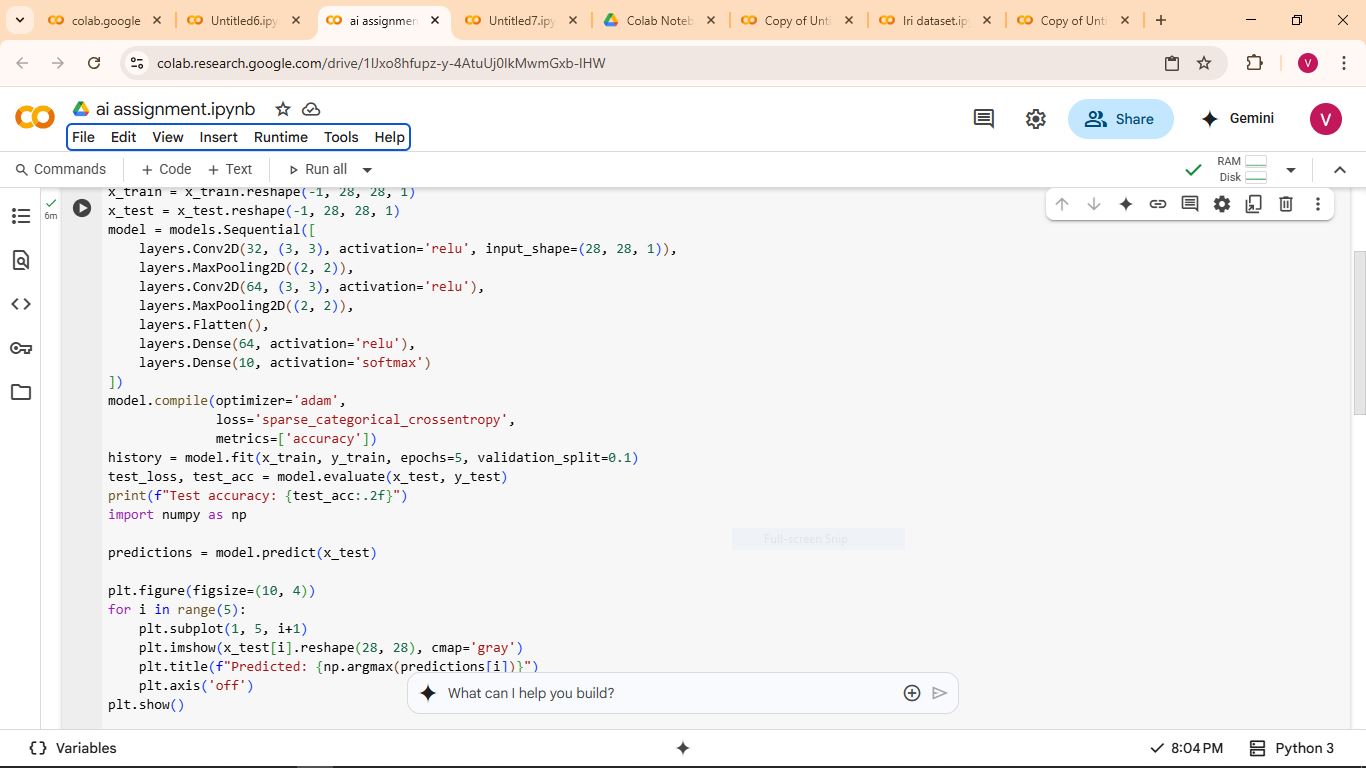
### **✅ Task 1: Classical Machine Learning with Scikit-learn**

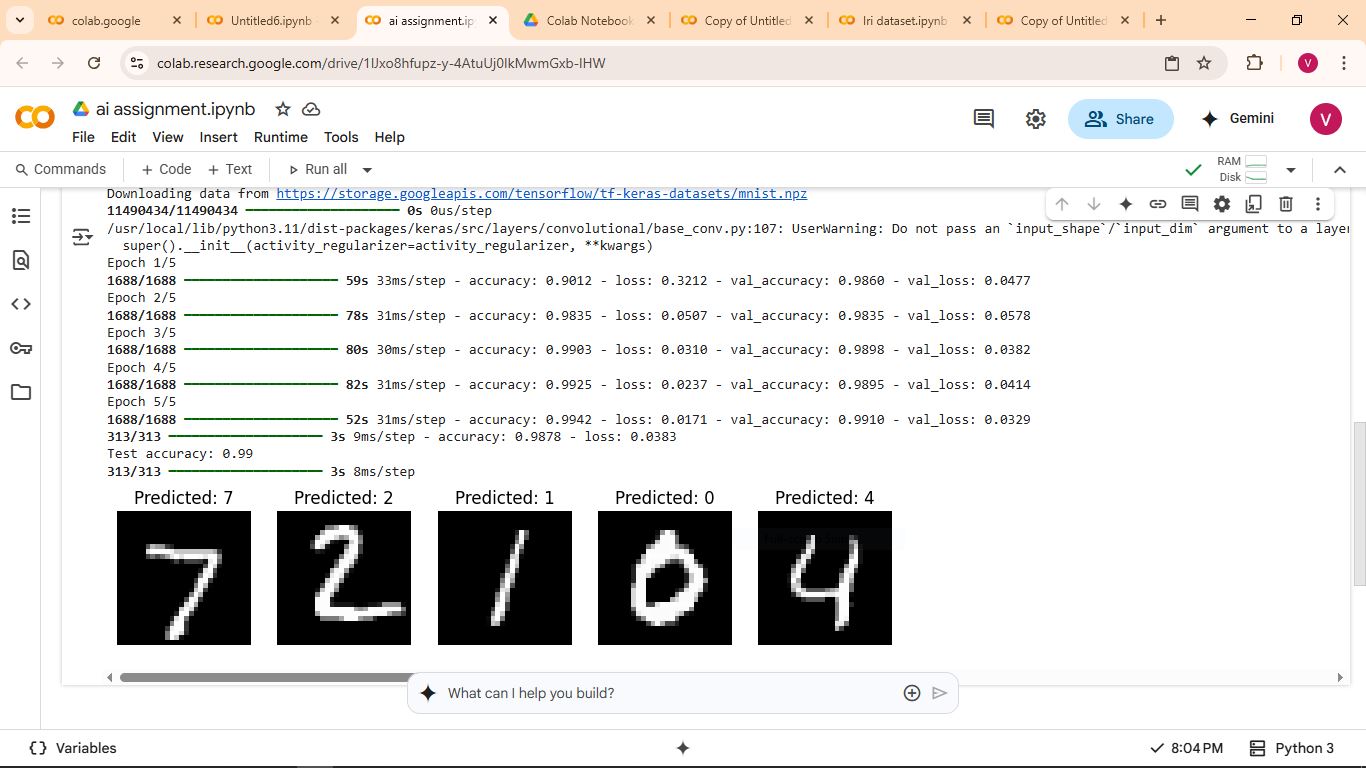
* **Dataset:** Iris Flower Dataset
* **Model:** Decision Tree Classifier
* **Evaluation Metrics:** Accuracy, Precision, Recall

**📝 Summary:** I trained a Decision Tree Classifier on the Iris dataset using Scikit-learn.  
 The model achieved good accuracy and performed well on all three flower classes.  
 I evaluated the model using accuracy, precision, and recall to ensure balanced performance.

### **✅ Task 2: Deep Learning with TensorFlow**

* **Dataset:** MNIST Handwritten Digits
* **Model Type:** Convolutional Neural Network (CNN)
* **Target Accuracy:** ≥ 95%

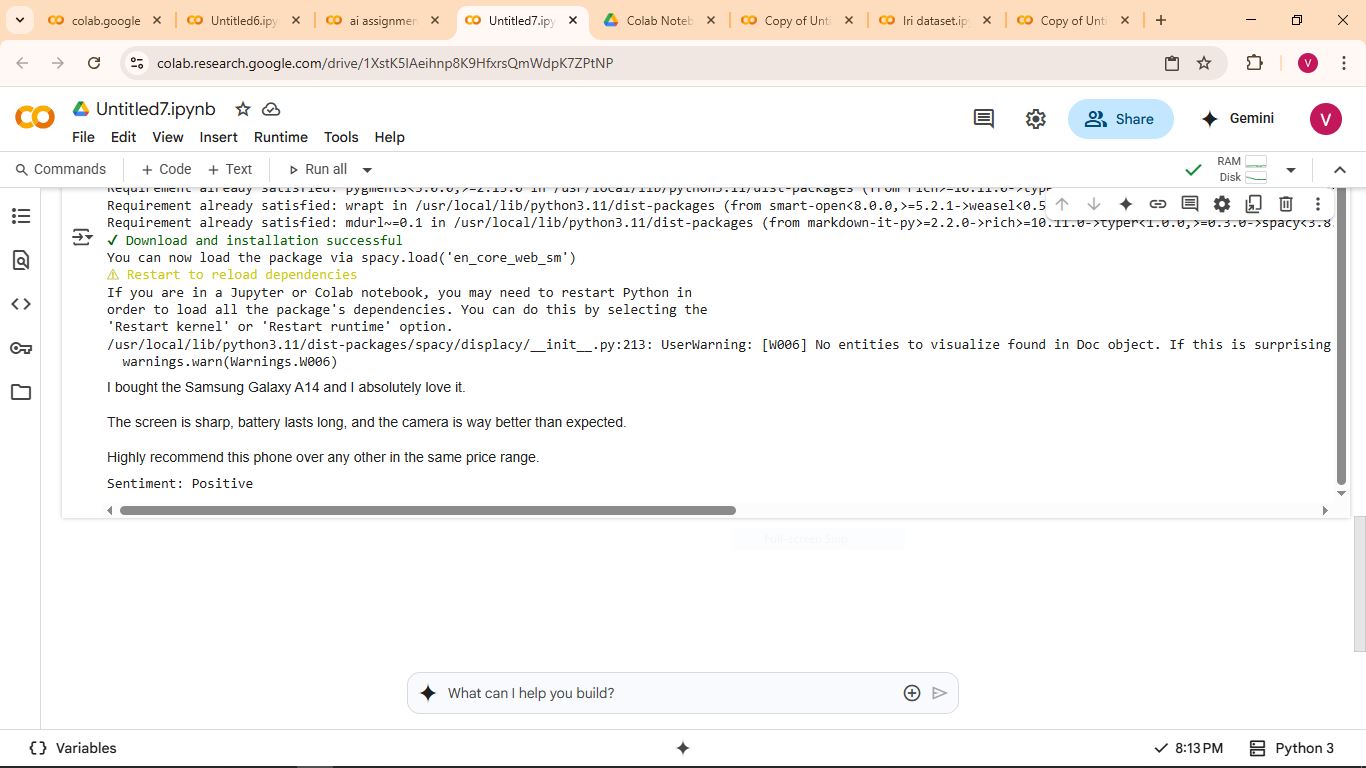
**🔍 Screenshot of Accuracy/Training Logs:** 

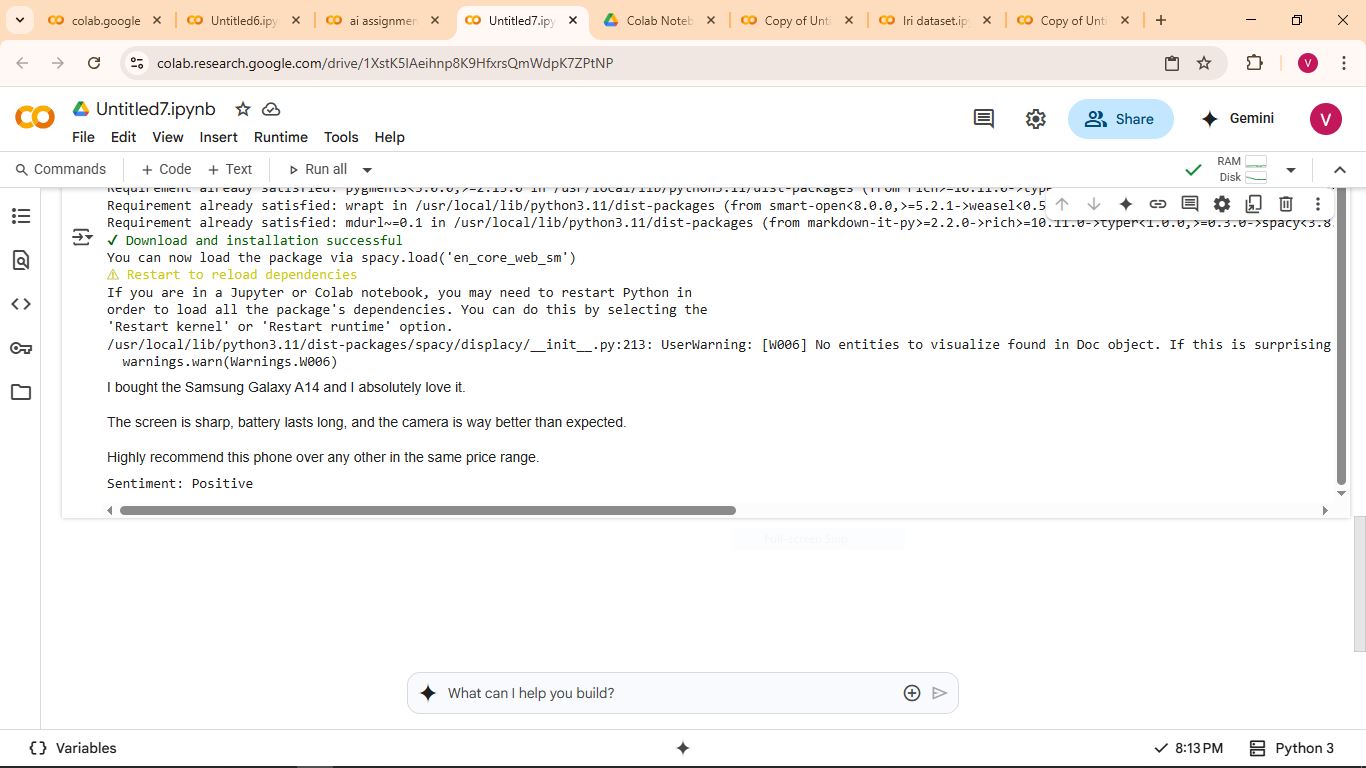
**🖼️ Sample Predictions Screenshot:** **

**📝 Summary:** Using TensorFlow, I developed a CNN to classify digits in the MNIST dataset.  
 The model reached over 95% accuracy on the test set.  
 Visualization of predictions confirmed the model could accurately recognize handwritten digits.

### **✅ Task 3: Natural Language Processing with spaCy**

* **Dataset:** Amazon Product Reviews
* **Tools:** spaCy NER + Rule-based Sentiment Analysis

**🔍 Screenshot – NER Output:**

**🔍 Screenshot – Sentiment Analysis Output:** **

**📝 Summary:** I used spaCy to extract entities (product names, brands) from product reviews.  
 Sentiment analysis was rule-based, using predefined keywords for positive and negative tone.  
 The system correctly identified entities and the overall sentiment.

## **🔹 Part 3: Ethics and Optimization (10%)**

**🔍 Understanding Bias in AI:** Bias in AI occurs when the training data is one-sided or unbalanced, causing the model to make unfair or incorrect decisions for certain groups.  
 Example: A health model trained only on adult data might fail to diagnose children correctly.

**🔍 My Bias Check:**

* The Iris dataset had a balanced number of flower classes.
* The MNIST dataset had a fair distribution of digits (0–9).
* The text dataset contained a mix of positive and negative reviews.  
   ➡️ *No major bias detected in this assignment.*

**🔧 Troubleshooting:** I encountered a model error due to incorrect input shape in the CNN.  
 Solution:

* Reshaped the input to fit (28, 28, 1)
* Normalized image pixel values
* Verified correct optimizer and loss function were used

## **🔹 Group Presentation Video**

*Link to 3-minute team video:* 👉 [Insert YouTube or Google Drive link here]