**THE REPORT AND SOULUTIONs FOR MINIST LAB.**

***Lab 1’ Requirment: The given pyhton code is use to predict number 5 in the MINIST data set, student have to transfer some code to predict number 7.***

\*Solution for Lab 1: **Binary Classification Setup**

***Step 1: We chose number 7 instead number 5***

In the original code:

y\_train\_5 = (y\_train == 5)

y\_test\_5 = (y\_test == 5)

I rewrite and implement it as follow:

y\_train\_7 = (y\_train == 7)

y\_test\_7 = (y\_test == 7)

***Step 2: Updating the valuable “some\_digit” to predict image of number 7***

From: some\_digit = X[0]

To: some\_digit = X[15]

***Step 3: Retrain the model with lable of number 7 to predict the result***

sgd\_clf = SGDClassifier(random\_state=42)

sgd\_clf.fit(X\_train, y\_train\_7)

sgd\_clf.predict([some\_digit])

**Step 4: *We also update all the variable*** *y\_train\_5 or y\_test\_5* ***to become*** *y\_train\_7 or y\_test\_7*

**Finally, the result is:**

**A screenshot of a computer code

AI-generated content may be incorrect.**

***\*\*Performance Evaluation***

* **Used:**
  + Confusion matrix to observe true/false positives and negatives.

A screenshot of a computer code

AI-generated content may be incorrect.

* + Precision, Recall, and F1-score to evaluate the classifier’s performance.

A screenshot of a computer

AI-generated content may be incorrect.

* + Precision-Recall curve

A green and blue line graph

AI-generated content may be incorrect.

* + ROC curve to analyze the precision-recall tradeoff and overall classification ability.

A blue line graph with a dotted line

AI-generated content may be incorrect.

***\*\*\*Alternative Models***

* Implemented a RandomForestClassifier as an alternative model.
* Compared prediction and probability outputs with SGDClassifier.

***\*\*\*\*Scaling and Improved Performance***

* Scaled the training data using StandardScaler.
* Observed a significant accuracy boost after scaling.