0917451 AI and Machine Learning (Spring 2022)

Quiz 1B

الاسم: رقم التسجيل: رقم الشعبة الأصلية:

==========

Instructions: Time **15** minutes. Open book and notes exam. No electronics. Please answer all problems in the space provided and limit your answer to the space provided. No questions are allowed.

P1. Complete the following code to find and print the model's RMSE on the **train set**.

```
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error

train_set, test_set = train_test_split(housing, test_size=0.3)
X_train = train_set.drop("y", axis=1)
y_train = train_set["y"].copy()
X_test = test_set.drop("y", axis=1)
y_test = test_set["y"].copy()

…   # some code is omitted

forest_reg = RandomForestRegressor(random_state=42)
forest_reg.fit(X_train_prepared, y_train)

housing_predictions = forest_reg.predict(X_train_prepared)

mse = mean_squared_error(y_train, housing_predictions)
print("RMSE = ", np.sqrt(mse))
```

P2. The following code is used in the MNIST classification problem. What is the main purpose of this code and how many training jobs it includes?

```
from sklearn.model_selection import GridSearchCV

param_grid = [{'weights': ["uniform", "distance"],
               'n_neighbors': [3, 4, 5, 6]}]

knn_clf = KNeighborsClassifier()
grid_search = GridSearchCV(knn_clf, param_grid, cv=3)
grid_search.fit(X_train, y_train)
```

This is grid search to find best hyper-parameters for the KNN classifier.

It fits 3 folds for each of the 2 × 4 = 8 candidates, totaling 24 fits.

<Good Luck>