Professional training in AI II Project :
Hand-written Digit recognizer using the MNIST DATASET

BY

Ahmed Alaa 221006208,

Amr el shamy 221006429 ,

Mostafa helmy 221002048

Problem Definition

This project is centered around training 3 distinct models(SVC , ANN , and CNN) to be able to detect and recognize digits from 0 to 9 , the models were trained using the infamous MNIST dataset which contains around 70,000 (60,000 for training and 10,000 for testing) different images of handwritten digits. The user would then interact with the models through a streamlit application which allows them to draw a digit on a canvas and choose which model they want to predict their handwritten digit.

Data Pre-processing

The data was split into training and testing , in which 60000 were training and the rest were testing. The x_train and x_test were then normalized by dividing them by 255 in order to have a range of values from 0-1 which is much more efficient and easier to handle for the models due to the decreased noise and consistent range. The x_train and x_test were then flattened meaning that the 2D arrays which were 28*28 is now 784 , this is done due to the neural networks needing a 1D input, and to increase the model performance as it would be easier to train on a 1D array than a 2D array. For the CNN , the y_train and y_test are encoded into categories to make it easier for the model to train.

Summary of Results

Three models — Support Vector Classifier (SVC), Artificial Neural Network (ANN), and Convolutional Neural Network (CNN) — were trained on the MNIST handwritten digits dataset to recognize digits from 0–9.

Model	Accuracy	Training Time	Notes
SVC (RBF kernel, C=10, γ=0.01)	98.3%	Long	Performs well but not optimized for image data.
ANN (3 hidden layers: 512–256–128, ReLU)	98.15%	Moderate	Learns global features but lacks spatial awareness.

CNN (Conv2D + MaxPooling)	98.64%	Short	Best performance – captures spatial patterns effectively.

Model Comparison

CNN achieved the highest accuracy (98.64%) and generalized best on unseen data. SVM also performed well (98.3%) but was computationally expensive. ANN achieved good accuracy (98.15%) but lacked CNN's spatial understanding.

Why CNN’s are better for image processing problems

CNN’s have the ability to learn local edges in images and image patterns , as well as the fact that it reduces parameters using convolutional filters , its ability to recognize shifted/altered patterns , and finally the fact that cnn’s are not sensitive noise. All these reasons justify why CNNs are superior when it comes to image processing problems.

Conclusion

Among the three models, CNN achieved the best performance and generalization ability. SVM showed strong accuracy but was slower, while ANN provided a balanced performance. Overall, CNNs are the most effective model for image recognition tasks due to their ability to extract and combine spatial features efficiently.