

# Proposal of a subject:

## Building a handwritten digit recognition system using MNIST dataset

Hoai Thuan TRAN

### Problem

Handwritten digit recognition is a common problem in a field of Optical Character Recognition (OCR), and it has many practical applications, such as digitizing documents, automating data entry, or processing handwritten forms, such as invoices or surveys. In this project, we will develop a system that can recognize handwritten digits from images or digits written on a screen by using machine learning technology. This system can be extended to perform some basic calculations based on the recognized digits.

### The aim of this project are:

- Build a system that can accurately recognize handwritten digits (0–9) from input images.
- Convert handwritten input into machine-readable format and perform basic calculations on them such as: addition, subtraction, multiplication and division.

### Some main tasks:

- **Data collection:** In this project, you use the MNIST dataset, which consists of 60,000 training images and 10,000 testing images of handwritten digits. Each image is 28x28 pixels and labeled with the correct digit.
- **Image Preprocessing:** Normalize the image data (e.g., convert pixel values to a scale from 0 to 1), resize all images to the same dimensions, and apply noise reduction to improve recognition accuracy.
- **Model Development:** We will study how to build and train a model using an advanced technique that separates different classes of digits in machine learning which is *Support Vector Machine (SVM)* algorithm.
- **Evaluation:** Test the model's performance using the test dataset and assess its accuracy using metrics like *accuracy*, *precision*, *recall*, and *F1 score*.
- **Application:** Implement a user interface where users can upload or draw digits on screen, and the system will return the predicted digit based on the trained model. Your system should also be able to work with basic calculations.