Project-2 : images dataset using Convolutional Neural Networks (CNN)

In this project, each group will choose **one image dataset** from Kaggle and build a complete **Convolutional Neural Network (CNN)** model to classify images.
The goal is to practice the full machine learning workflow from data collection to evaluation and reporting.

Project Steps:

1. Data Collection

- Download a dataset from **Kaggle**.
- Make sure it contains images with labels (classes).

---

2. Data Preprocessing

- Convert images to the same size (e.g., 28x28 or 32x32).
- Normalize pixel values (e.g., divide by 255).
- Optionally add **data augmentation** such as:
    - Rotation
    - Flipping
    - Zoom
    - Shift
      (This helps improve accuracy.)

3. Split Data

- Divide the dataset into:
    - **Training data**
    - **Testing data**
- Optionally include **validation data**.

---

4. Model Building

- Build a CNN model using layers such as:
    - Convolution layers
    - Activation functions (ReLU, Softmax)
    - Pooling layers (MaxPooling)
    - Fully connected (Dense) layers
- Students are free to choose the architecture.

---

5. Training & Evaluation

- Train the model on training data.
- Test the model using testing data.

- Calculate performance metrics such as:
    - Accuracy
    - Loss
    - Confusion matrix

6. Model Improvements

Try to **improve model performance** by using techniques like:

- Regularization
- Dropout
- Changing the number of layers
- Adjusting learning rate
- Data augmentation
- Optimization tuning

(Students should try at least 2 improvement techniques)

Examples for datasets : Choose any dataset you want .

1. Fashion-MNIST

Description: Gray-scale images of clothing items.
Image size: 28x28
Number of classes: 10

Classes:

1.T-shirt/top, 2. Trouser, 3. Pullover, 4. Dress, 5. Coat, 6. Sandal, 7. Shirt, 8. Sneaker, 9. Bag, 10. Ankle boot

---

2. MNIST Handwritten Digits

Description: Gray-scale handwritten digits.
Image size: 28x28
Number of classes: 10 (Digits from **0 to 9)**

3. CIFAR-10

Description: Colored (RGB) natural images.
Image size: 32x32
Number of classes: 10

Classes:

1.Airplane, 2. Automobile, 3. Bird, 4. Cat, 5. Deer, 6. Dog, 7. Frog, 8. Horse, 9. Ship, 10. Truck

---

4. CIFAR-100

Description: Colored (RGB) natural images.
Image size: 32x32
Number of classes: 100

Examples of classes:
apple, aquarium fish, baby, bear, bicycle, bottle, bus, camel, cup, chair, food containers, vehicles, large animals, etc.

---

5. Chest X-Ray: Pneumonia Detection

Description: Medical X-ray images of chest.
Task: Binary classification

Classes:

1. Normal

2. Pneumonia

6. Brain Tumor MRI Classification

Description: MRI brain scan images to detect tumors.
Number of classes: Usually 4

Classes:

1. Glioma tumor
2. Meningioma tumor
3. Pituitary tumor
4. No tumor

(Some versions may have 3 classes, but the common dataset contains 4.)

---

7. FER-2013 (Facial Expression Recognition)

Description: Facial emotion images.
Image size: 48x48 gray-scale
Number of classes: 7

Classes:

1. Angry, 2. Disgust, 3. Fear, 4. Happy, 5. Sad, 6. Surprise, 7. Neutral

---

Final Project Rules

1. **Group Size**
    - Each group must have **5 to 6 students**.
2. **Submission Format**
    - **Python Notebook**: Each group must submit a **Jupyter Notebook (.ipynb)** containing all code, organized and well-commented.
    - **Reports**:
        - **2 detailed reports**: one for Project 1 and one for Project 2
        - Format: **PDF**
        - Each report should include:
            - Dataset description
            - Preprocessing steps
            - Model architecture
            - Training and testing steps
            - Evaluation metrics (accuracy, confusion matrix, loss curves, etc.)
            - Improvements and experiments
            - Observations and conclusions

3. **Deadline**
    - All notebooks and PDF reports must be submitted by:
      Monday, 22-12-2025 at 11:59 PM

4. **Project Discussion**
    - Discussion and presentations will take place on:
      Tuesday, 23-12-2026
    - Each group will present their project, results, and improvements.

5. **General Rules**
    - Each group must submit **one notebook and one set of reports per project**.
    - Late submissions may not be accepted.
    - Plagiarism is strictly prohibited.
    - All code in the notebook must be **well-commented and clear**.

Good luck!