Assignment 3

Q1. Develop a committee of deep learning models on the Fashion MNIST dataset and compare their combined performance against individual models, using various metrics including confusion matrices and classification reports (70 points).

Dataset
Fashion MNIST: 60,000 training images and 10,000 testing images, each 28x28 pixels, categorized into 10 clothing classes.

Tasks

Data Exploration & Preparation:
Exploration: Visualize a sample of images from each class in the Fashion MNIST dataset.
Normalization: Scale the images to have pixel values between [0,1].
Data Formatting: Convert labels to one-hot encoded vectors. Prepare data suitable for CNNs.

Model Development:
Model 1: Design a shallow neural network. Explain the choice of layers and activations.
Model 2: Construct a basic convolutional neural network (CNN).
Model 3: Build a deeper CNN, ensuring it's architecturally different from Model 2.

Training, Validation & Model Evaluation:
Train each model, recording loss and accuracy metrics over epochs.
Use validation data to monitor performance and avoid overfitting.
Save the best weights for each model(optional).
Assess each model's performance on the test data, reporting accuracy.

Performance Metrics & Analysis:
Generate and visualize confusion matrices for each model and the committee.
Generate classification reports, comparing precision, recall, and F1-scores. Comment on the strengths and weaknesses of each model, focusing on which classes are most often misclassified.

Report Writing:
Introduction: Briefly explain the dataset and the objective.
Methodology: Elaborate on the design choices for each model and justify.
Results: Tabulate and visualize metrics for easy comparison.
Analysis: Discuss any improvements or setbacks observed with the committee approach. Suggest potential reasons for the observed behavior of each model.
Conclusion: Summarize findings and propose future improvements or experiments.

Extra Task:

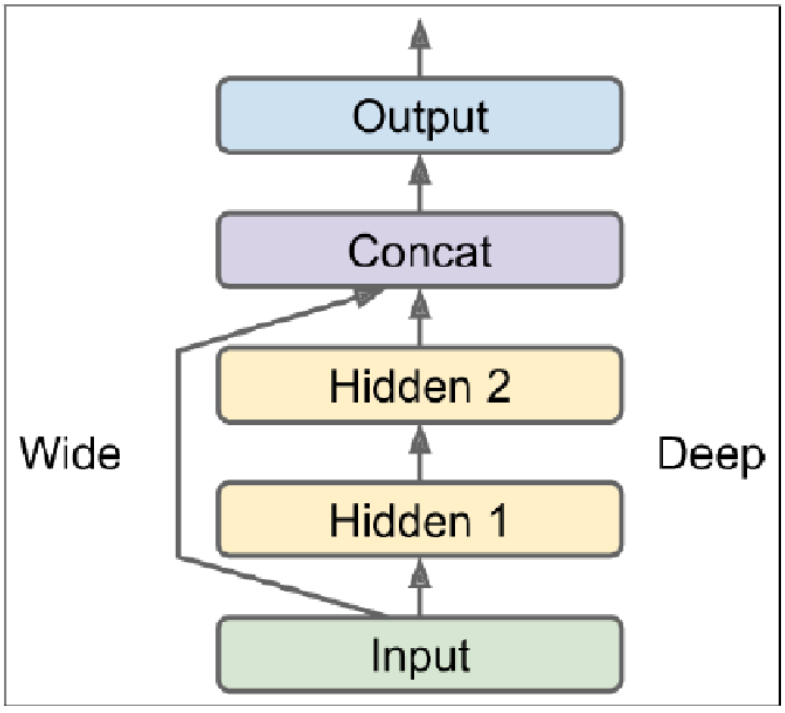
Implement techniques to enhance model performance: data augmentation, dropout, batch normalization, etc. Explore ensemble techniques beyond simple averaging, e.g., weighted averaging based on validation performance.

- Deliverables

Comprehensive code (preferably in Jupyter notebook format) with comments.
A structured report covering all tasks.
Visual aids: Confusion matrix heatmaps, sample predictions, training curves.

Q2. Please use the Fashion MNIST dataset in Q1 to implement the following network along with batch normalization and dropout layers. (30 Points)

```
input = keras.layers.Input(shape=X_train.shape[1:])
hidden1 = keras.layers.Dense(30, activation="relu")(input)
hidden2 = keras.layers.Dense(30, activation="relu")(hidden1)
concat = keras.layers.Concatenate()([input, hidden2])
output = keras.layers.Dense(1)(concat)
model = keras.models.Model(inputs=[input], outputs=[output])
```



.

- Deliverables

Comprehensive code (preferably in Jupyter notebook)
A structured report with Confusion matrix heatmaps, sample predictions, training curves and classification reports.