A Neural Network Model for Fashion-MNIST Image Classification

This project examines how the Artificial Neural Networks (ANN) and the Convolutional Neural Networks (CNN) can be used to classify images as they are presented with the use of the Fashion-MNIST dataset which includes 70,000 grayscale images (28×28 pixels) depicting 10 categories of clothing pieces like shirts, trousers, and sneakers. It was proposed to design, train and test both of the models with the same data and training parameters and to compare their performance and investigate how a variant of architecture influences learning and accuracy.

To preprocess the data, first, the dataset was loaded through the TensorFlow/Keras and split it into training, validation, and testing chunks, normalized values of pixels between 0 and 1, and one-hot coded the labels. Sample images were also visualized to ensure that the dataset was loaded and available to be modelled.

The model used was ANN, which was a two-hidden-layers simple feed-forward network, with 256 and 128 neurons (ReLU activation), dropout regularization (0.2) and a softmax output layer. It was also trained with Adam optimizer and categorical cross-entropy loss. Although the ANN gave average performance (somewhere between 87 and 87 percent test accuracy), it was only able to distinguish similar classes that were visually close like shirts and coats because of its inability to learn the spatial information.

CNN model is a better and more efficient model that produced spatial and local features of images made using convolutional and pooling layers. It had two convolutional layers (32 and 64 filter), max-pooling, dense (128 neurons), dropout (0.25) neurons, and softmax output. The CNN trained under the same conditions showed better performance with a test accuracy of about 91% and convergence rate as it is able to identify visual data faster.

Assessments measures like accuracy curves as well as confusion matrices and images of as erroneously classified images indicated that in all stages, the CNN performed better than the ANN. It had less misclassifications and extrapolated more onto unseen data. Nevertheless, the two models had difficulties when working with grayscale pictures and similar items.

The paper is concluded by revealing that CNN networks are better in the image classification tasks than fully connected ANNs due to exploitation of spatial arrangements and local patterns. Additional enhancements might be provided by data augmentation, more sophisticated architectures (e.g. VGG or ResNet), and hyperparameter optimization to achieve an additional increase in the robustness and quality of the model. Generally, this project offered a positive experience of practical exposure to data preprocessing, model training, testing, and learning the effects of neural network architectures on their performance in the realm of deep learning.