Executive Summary: Analysis of Fashion-MNIST Classification

Introduction

This research investigates how to use ANN and CNN to classify images in Fashion-MNIST database. The dataset has 70,000 black and white images of size 28×28, from ten clothing categories such as shoes, shirts, and hats. The purpose was to build, run and rate two neural net models that could arrange suitable the images, and also the effect of building choices and training procedures on their success.

Model Overview

The ARTIFICIAL NEURAL NETWORK (ANN) was made by creating a speech feature extraction and a completely connected multi-layer perceptron that had two hidden layers of 256 and 128 neurons respectively. The presence of high number of parameters can cause overfitting. For this re saliency ReLU was used as an activation function and a dropout rate of 0.3. The second architecture was a convolutional neural network (CNN) that contained two convolutional layers of 32 and 64 filters of kernel size 3×3. After each convolutional layer, ReLU was applied. Learning where a max pooling layer of 2×2 was provided after each convolutional layer, and a dense layer of 128 neurons. The architecture of each mode was trained for 25 epochs on GPU environment. The Adam optimizer with learning rate of 0.001 was used with the function CrossEntropyLoss..

Key Findings

CNN beat the ANN by a large margin with a test accuracy of 92% vs the ANN 87%. The ANN could detect some simple patterns of strength but it also found it hard to tell the difference between visually similar groups – for example Coats and Shirts. Instead, the CNN was good at knowing about where things are in images and that let it generalize better which got it to learn better and keep on doing it during training. Accuracy scores on both systems showed that the CNN was better with a Precision of 0.92, a Recall of 0.91 and an F1-score of 0.91. When looking at what types of errors each system made the diagonal pattern was very strong showing that most errors it did make were actually predictions to the correct classes.

Innovations and Experimental Approach

A variety of optimization approaches improved both the efficiency and stability of training:

- Relying on Data Augmentation techniques such as random rotations and flips enhanced robustness to variation at the input side.
- Applying the dropout and normalization methods lowered the risk of overfitting and aided consistent convergence.
- The Adam optimizer facilitated adaptable learning rates and effective gradient descent.
- To evaluate learning patterns and the dependability of classifications, visualization tools like accuracy and loss graphs, along with confusion matrices, were employed.

Conclusions and Insights

This research results show good understanding of how neural nets work in visual sort. Convolutional nets (CNNs) were best in tracking space and avoiding mistakes and that made the results clearer and more accurate. These results emphasize how important the depth of the model, how the model is adjusted, and how the model is set up to give good results. Next steps might include looking at new Net types like ResNet and VGG, adding in transfer learning, and working on tools to make it easier to analyze using Grad-CAM. To finish, this research shows a high level of accuracy in theory, strong checks for the results, and an organized way to check the neural nets.

Kirtan Nitin Patel
(P2952845)