There are three neural network designs:
ANN, CNN, and ResNet-20 that were used in this project to perform the Fashion-MNIST classification task. The main objective was the study of the impact of model architecture and depth on image classification accuracy generalisation and robustness on their implementation with the help of TensorFlow. The Fashion-MNIST dataset has 60, 000 training grayscale images and 10, 000 test grayscale images in 10 fashion categories, including T-shirt, coat, pullover and sneaker etc.

Images are 28x28 and hence a small but difficult dataset because there can be visual areas of similarity among some categories. It can be used as a powerful baseline to compare the deep learning architectures and to comprehend the trade-off between complexity and performance of the model. Model Design and Implementation. The ANN model was created as a benchmark based on many fully connected (dense) layers with ReLU LeakyRelu and Prelu activations and dropout regularisation to decrease overfitting. The ANN with its no spatial feature extraction and simplicity approach gave a good test accuracy of 90, which shows that well tuned dense networks can be competitive when used with normalised and small scale data of images.

The CNN model added convolutional and max average and global pooling to the network in order to learn local spatial hierarchies and texture patterns. Generalisation and training stability was enhanced by using batch normalisation and dropout. The CNN had a test result of 92, which is a definite lead on the ANN, as it uses feature maps, where edges, textures, and shapes are directly computed out of image pixels.

ResNet-20 model as an extension of CNN used residual connexions, which enabled the gradient to flow effectively to add deeper layers and avoid degradation as it progresses in training. The ResNet-20 implemented via the functional API of TensorFlow trainable well and did not have a vanishing gradient problem; the model had the highest test accuracy of 92.8, which is the highest of all three models.

This shows the usefulness of residual learning even with relatively complicated datasets such as Fashion-MNIST. Training and Evaluation Each of the models was run in TensorFlow/Keras and trained minimum with 20-30 epochs of Adam optimizer and categorical cross-entropy loss. The data were brought to a [0,1] scale and broken down into training, validation, and test sets. To improve convergence and avoid overfitting, early termination and learning-rate scheduling were used.

Test accuracy, confusion matrices and training/validation loss curve were used to measure the performance. ResNet-20 model exhibited the smoothing convergence and more balanced predictions and had less confusion of similar appearances of the classes like shirts and coats. Key Findings and Conclusion The results show a steady improvement in generalisation and accuracy with the increasing model depth and architecture complexity.

The ANN gave a good background, whereas the CNN and ResNet-20 gave greater accuracy with the spatial feature aspects and residual mapping. The study states with 90, 92, and 92.8 accuracies that modern deep learning structures are by far superior to simple dense networks in performing tasks of visual recognition.

To sum up, this project emphasises the usefulness of TensorFlow in building and training neural networks and explains how architectural improvements, in particular, residual learning, can improve performance and stability. It could be extended further in the future with more sophisticated data augmentation like CutMix or MixUp and some experimentation with pretrained models to do even better transfer learning.