Learning Problem:
The goal is to train a neural network model to accurately classify images of clothing into 10 categories (t-shirt, trousers, pullover, dress, coat, sandal, shirt, sneaker, bag, ankle boot). The training data consists of 28×28 pixel grayscale images from the Fashion MNIST dataset.

Model Architecture:
The model uses a simple sequential architecture with two dense layers. The input layer flattens the 28×28 images into 784 pixel values. This is connected to a hidden layer with 128 neurons and ReLU activation. More neurons could potentially model more complex patterns, but too many may lead to overfitting. 128 was chosen by experimentation to balance model capacity with avoiding overfitting.

The output layer has 10 neurons with softmax activation, one for each clothing category. This turns the outputs into probability-like values summing to 1, suitable for the categorical cross-entropy loss function used during training.

Additional Design Decisions:
The Adam optimization algorithm was used for fast, adaptive learning rates during training. L2 kernel regularization was added to further combat overfitting, with a strength of 0.001 chosen by testing values between 0.0001-0.1.

The data was split 80/20 into training and validation sets. The model was trained for 10 epochs with a batch size of 32. This combination of epochs and batch size was chosen by experimentation to allow robust model convergence without overfitting.

Experimental Results:
The model achieved 87% accuracy on the test set. The confusion matrix shows some difficulty distinguishing between certain categories like coats vs shirts, but overall the performance is very good, indicating the model has learned robust features for generalizing across clothing types.

The difference between training and validation accuracy is small, suggesting little overfitting. This shows the regularization and early stopping mitigated overfitting.

Conclusions:
The high test accuracy indicates the model architecture has sufficient complexity for this task. The Adam optimizer and regularizers enabled efficient training without much overfitting. For further improvement, data augmentation techniques could generate more training examples to handle edge cases. Overall this model provides a solid baseline for classifying clothing from images.