Assignment 2:

1) **Performance Comparison: Discuss the test accuracy and loss of both models:**

MLP results:

313/313 ━━━━━━━━━━━━━━━━━━━━ **0s** 1ms/step - accuracy: 0.4710 - loss: 1.5196

[1.520288109779358, 0.4722000062465668]

By the results we see above, the MLP model correctly predicted the labels for approximately 47.1% of the test samples, our model is correct in less than half of our tests, which is not an optimal result.

This loss value is also reported during this evaluation and represents the model's performance in terms of how well its predictions align with the true labels. In the MLP model we got a loss value of 1.5196 , which also suggests that the model is not performing optimally.

CNN results:

10/10 ━━━━━━━━━━━━━━━━━━━━ **8s** 767ms/step - accuracy: 0.7254 - loss: 0.8206

[0.8248381018638611, 0.72079998254776]

By the results we see above, the CNN model correctly predicted the labels for approximately 72.1% of the test samples, indicating that our model is performing relatively well, as it successfully classifies more than two-thirds of the tests.

The loss value is also reported during this evaluation and represents the model's performance in terms of how well its predictions align with the true labels. In the CNN model, we got a loss value of 0.8248, which suggests that the model is performing better than the previous model, as it indicates a lower prediction error. Overall, the CNN shows a significant improvement over the MLP in both accuracy and loss.

2) **Model Complexity: Explain why CNNs are generally better suited for image classification tasks compared to MLPs, focusing on the spatial hierarchies learned by CNNs:**

CNN models are generally better suited for image classification tasks than MLP models because of their ability to learn spatial hierarchies effectively. CNNs use convolutional layers that apply filters to small regions of an image, allowing the model to capture local patterns such as edges, textures, and shapes. This local connectivity enables the CNN to maintain the spatial structure of the image, which is crucial for recognizing complex objects. In contrast, MLPs treat the input images as flat vectors, losing the spatial relationships that are essential for interpreting visual data. As a result, CNNs can build hierarchical feature representations, where lower layers learn basic features and deeper layers combine these to recognize more complex structures, leading to better classification performance.

3) **Training Time: Compare the training time and resource usage of both models.**

The CNN model took approximately 4 minutes and 25 seconds (265 seconds) to train, compared to 1 minute and 1 second (61 seconds) for the MLP model, indicating that the CNN required about 4.34 times longer. This longer training time is expected due to the complexity of CNN model, which involve multiple convolutional layers. I used TPU for both as my runtime resource but because CNN models require more memory and are more complex than MLP models, the training time of MLP models are much faster.

4) **Overfitting and Regularization: Discuss any signs of overfitting and how regularization techniques like dropout help mitigate it:**

A sign for overfitting is when the model performs well on the training dataset, but not on the test dataset. It usually happens because the model "memorizes" the training dataset but when dealing with unseen data, such as the test data, the model will not perform well. One of the most common ways to deal with this problem is using dropout layers.

During training dropout layers choose randomly a set of units from the preceding layer and set their output to 0. Doing this prevents the model from being overly dependent on only a few units, and instead spread the "knowledge" of the model across the entire network, which results in better performance on unseen data.