## Task 2

### Dataset Overview and Preprocessing

Task 2 assignment is focusing on the objectives and importance of multi-label classification on multiple digits from images, where each image contains several digits.The dataset for this project consists of images each labeled with three digits.

### Data Loading and Initial Exploration

To load and explore the dataset, we define the load\_images\_from\_folder function, which reads images from subdirectories, resizes them, converts them to grayscale, and assigns labels based on subdirectory names. Directories for training, validation, and test datasets are specified to organize the data. We print the contents of these directories to verify their structure. We load the datasets into arrays of images and corresponding labels. This process prepares the dataset for further processing and model training.

### Preprocessing Steps:

Preprocessing steps taken to prepare the data for modeling. This includes reshaping and normalizing pixel values, One-hot encoding of labels, and creation of a custom data generator for efficient data handling during model training.

### Challenges:

The challenges during preprocessing involved reshaping the images to fit the model's input requirements, normalizing pixel values to the rnge [0, 1], and encoding the labels for a multi-label setup using one-hot encoding. We faced several challenges when we need to choose which activation function and which loss function we need to use.

### Methodology and Techniques:

CNN models were developed, comparing Categorical Crossentropy with Binary Crossentropy as loss functions and Sigmoid versus Softmax as activation functions. Techniques like early stopping, model checkpointing, and TensorBoard logging were used to enhance training performance and avoid overfitting. Hyperparameter tuning was performed to optimize learning rates, batch sizes, and model architecture.

### Model Comparison and Selection

Binary crossentropy as the loss function outperformed categorical crossentropy, showing lower loss and higher accuracy for multi-label classification, demonstrating its suitability for this task. While recommendations suggest Sigmoid for multi-label and Softmax for multi-class, the use of Softmax achieved higher accuracy in this context.

### Hyperparameter Tuning

To optimize a model's performance, we adjust model parameters like learning rate and batch size to find the most effective settings for the specific problem. We use techniques like K-fold cross-validation to more accurately evaluate model performance across different subsets of data, ensuring robustness and reliability. We also implement batch normalization in the model to normalize the inputs for each mini-batch, thereby reducing internal covariate shift and stabilizing training.

### Results and Discussion:

Initial models using Sigmoid activation and categorical crossentropy achieved high accuracy (around 95%) and loss (around 1.00), indicating no underfitting or overfitting

Test loss: 1.0049712657928467

Test accuracy for digit\_0: 0.9526000022888184

Test accuracy for digit\_1: 0.9549999833106995

Test accuracy for digit\_2: 0.9416499733924866

Switching to Sigmoid activation and binary crossentropy reduced losses (around 0.3), but lowered accuracy. The final evaluation on the test dataset showed accuracies above 90% for all digits.

Test loss: 0.3763981759548187

Test accuracy for digit\_0: 0.907800018787384

Test accuracy for digit\_1: 0.9002500176429749

Test accuracy for digit\_2: 0.9164999723434448

The model using Softmax activation and binary crossentropy showed the best performance with around 98% accuracy and lower losses (around 0.07). After hyperparameter tuning, the model demonstrated improvement and better performance. Final testing on unseen data confirmed the model's high accuracy:

Test loss: 0.07641927897930145

Test accuracy for digit\_0: 0.9811499714851379

Test accuracy for digit\_1: 0.9810500144958496

Test accuracy for digit\_2: 0.9817000031471252

There is high accuracy on the test set which the model has never seen before, that means the model can predict on the new unseen data. The model regularized well and trained appropriately, no overfitting nor underfitting.

### Conclusion

The project successfully demonstrated the application of CNNs to multi-label image-based digit classification, with Softmax activation and binary crossentropy showing the highest performance over Sigmoid activation and categorical crossentropy as loss function.

This is showing how importance in selecting appropriate loss functions and optimizing model architectures. Future work could explore more complex architectures, additional regularization techniques, and further hyperparameter tuning to enhance model performance and robustness, potentially experimenting with different convolutional layer configurations to improve focus on relevant parts of the image.

### Ethical, Legal and Social Considerations in AI

When developing AI models for multi-label image classification, such as with triple MNIST digits, it's important to address ethical, legal, and social issues. Data protection laws like GDPR or CCPA must be adhered to, ensuring that any personal data is handled with strict privacy measures.   
  
The ownership of the AI model, its training data, and its outputs need to carefully consider to ensure responsible and beneficial AI deployment.