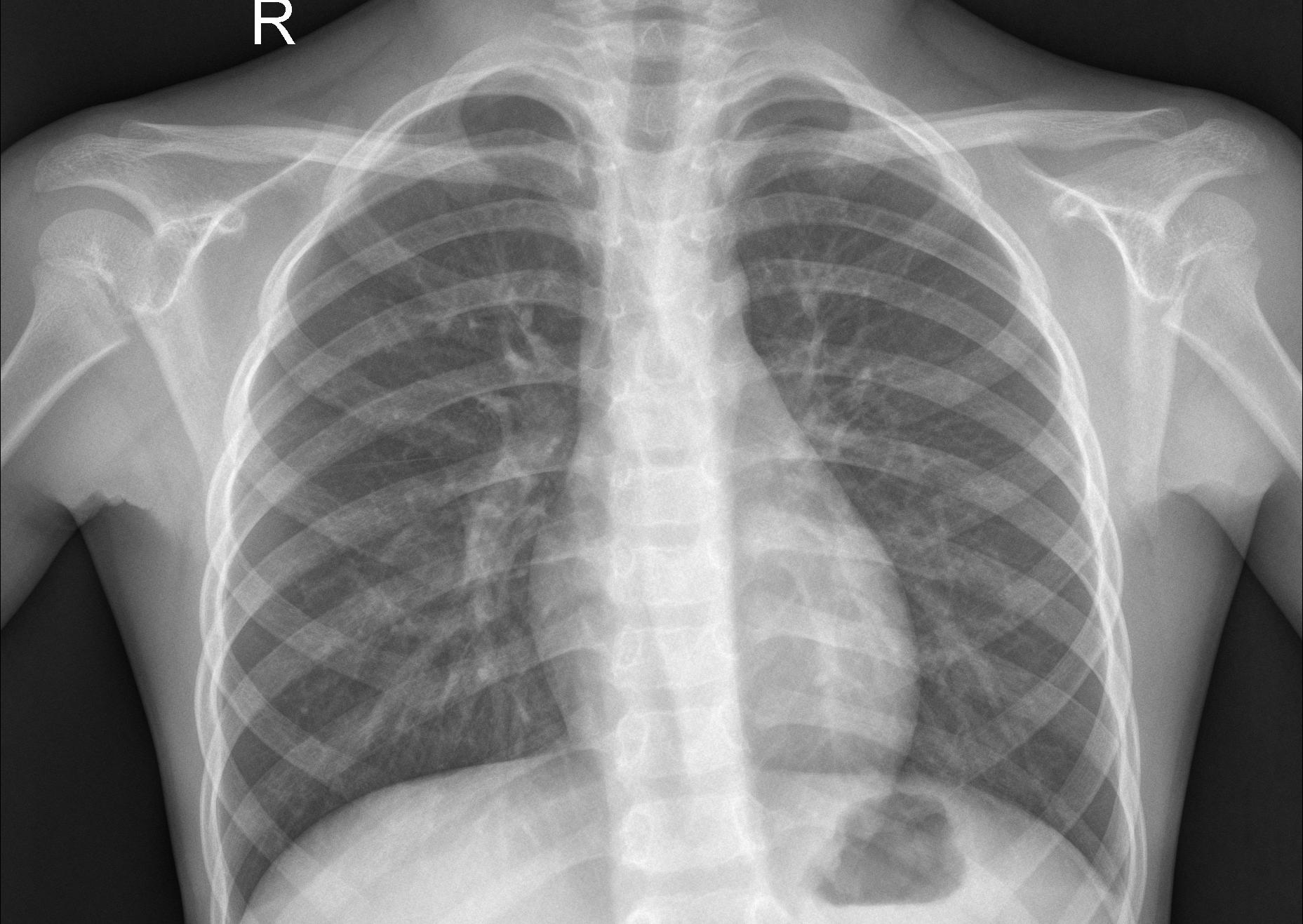
**Pneumonia Detection using Chest X-ray Image**

This documentation provides an overview of the machine learning code for pneumonia detection using chest X-ray images. The code utilizes various deep learning models, including Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and MobileNet V2, for pneumonia detection. The code is written in Python using the TensorFlow and Keras libraries.

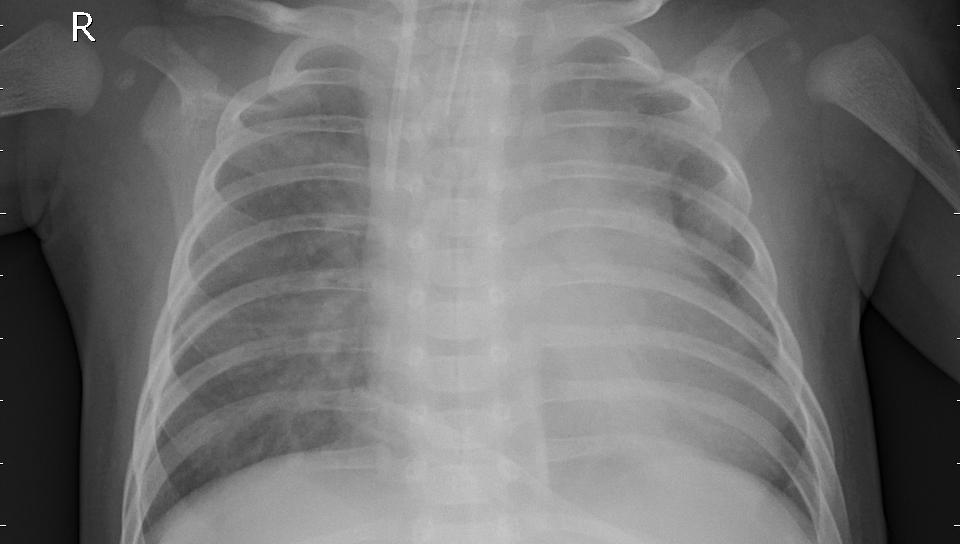
**Dataset**

The dataset used in this code consists of two classes: 'pneu' (for pneumonia cases) and 'healthy' (for healthy cases). The dataset is divided into two dictionaries: 'pneu\_images\_dict' containing the images for each class, and 'pneu\_labels\_dict' containing the corresponding labels for each class.

Example image of healthy datatset:



Example image of pneumonia dataset:



**Data Preprocessing**

The images are read using OpenCV library and resized to a fixed size of 224x224 pixels using the 'cv2.resize' function. The resized images are then appended to a list 'X', and their corresponding labels are appended to another list 'y', using the 'pneu\_labels\_dict' dictionary.

The dataset is then split into training and testing sets using the 'train\_test\_split' function from the 'sklearn.model\_selection' module. The images in the training set are further scaled by dividing by 255 to normalize the pixel values.

**Artificial Neural Network (ANN) Model**

The ANN model consists of several dense layers with varying number of units and activation functions, including 'relu' and 'softmax'. The model is compiled using the 'adam' optimizer, 'sparse\_categorical\_crossentropy' loss function, and 'accuracy' metric. The training data (X\_train and y\_train) is then used to train the ANN model for 5 epochs.

**Convolutional Neural Network (CNN) Models**

Two CNN models are implemented: one without data augmentation and dropout, and another with data augmentation and dropout for regularization.

**CNN (Without Data Augmentation and Dropout)**

The first CNN model consists of several convolutional and max pooling layers, followed by flatten and dense layers with 'relu' activation functions. The model is compiled using the 'adam' optimizer, 'SparseCategoricalCrossentropy' loss function, and 'accuracy' metric. The training data (X\_train\_scaled and y\_train) is then used to train the CNN model for 16 epochs.

**CNN (With Data Augmentation and Dropout)**

The second CNN model includes data augmentation techniques, such as random flip, rotation, and zoom, using the 'layers.experimental.preprocessing' module. The model also includes dropout layers for regularization. The training data (X\_train\_scaled and y\_train) is augmented using the 'data\_augmentation' sequential model, and then used to train the CNN model for 40 epochs.

**CNN (With Dropout Only)**

The third CNN model includes dropout layers for regularization, without any data augmentation. The training data (X\_train\_scaled and y\_train) is used to train the CNN model for 20 epochs.

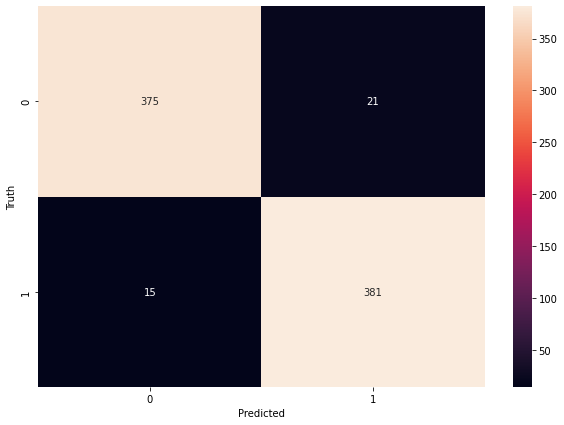
**MobileNet V2 Model**

The MobileNet V2 model is a pre-trained model that is loaded using the 'hub.KerasLayer' function from the TensorFlow Hub. The model is then used as a feature extractor, and additional layers are added on top of it for classification. The training data (X\_train\_scaled and y\_train) is used to train the MobileNet V2 model for 10 epochs.

**Model Evaluation**

The accuracy of each model is calculated using the 'accuracy' metric during training. The accuracy of the ANN model is 0.4865, the accuracy of the CNN model without data augmentation and dropout is 0.9806, the accuracy of the CNN model with data augmentation and dropout is 0.9499, the accuracy of the CNN model with regularisation only is 0.9844,the accuracy of cnn model with Mobile net architecture is 0.9895.

**Confusion matrix for mobile net architecture:**



**Conclusion:**

The implemented ML model for predicting Pneumonia using X-Ray scans of patients shows promising accuracy results. Further optimization and fine-tuning of the model parameters may be necessary to achieve higher accuracy. Additionally, thorough validation and testing on diverse datasets are recommended to ensure the model's reliability and generalization capability.