

# **CLASSIFICATION OF CHEST X-RAY IMAGES DIAGNOSED WITH PNEUMONIA USING DEEP LEARNING MODELS**

Amir Khoeilar

## **Abstract**

Pneumonia is a lung infection that can range from mild to so severe that you have to go to the hospital. It happens when an infection causes the air sacs in your lungs (your doctor will call them alveoli) to fill with fluid or pus. That can make it hard for you to breathe in enough oxygen to reach your bloodstream.

Your symptoms can vary depending on what's causing your pneumonia, your age, and your overall health. They usually develop over several days. Common pneumonia symptoms include:

1. Chest pain when you breathe or cough
2. Cough that produces phlegm or mucus
3. Fatigue and loss of appetite
4. Fever, sweating, and chills
5. Nausea, vomiting, and diarrhea
6. Shortness of breath

Using deep learning models, I'm going to be analyzing images provided by the Kaggle website of patients with normal chest x-ray and also chest x-ray images diagnosed with Pneumonia. These models will eventually help the medical field to increase the accuracy and also speed of diagnosis of potential patients and take the necessary future steps.

## **Design**

I began with simple EDA in order to have a better understanding of the data provided by the Kaggle website. Loading the data from directories to the Jupiter notebook was quite difficult, harder to get the correct format.

As of the first step, I ran a Simple PCA logistic regression model on the data. As for my deep learning models, I started with convolutional neural network models and performed a baseline model on the data with reducing the size of the images and then implementing 4 hidden layers. Besides the baseline model, I ran 2 other CNN models in order to get higher result on accuracy and also get a better understanding of what the relation between numbers of epoch, batch size and all the numbers calculated by the `.summary()` function. Finally I ran a MobileNet model using transfer learning for comparison.

## Data

The data provided by the Kaggle website had 5600 Xray images separated in 2 directories of "Normal" and "Pneumonia". I also found another extra dataset to test my models on consisted of 600 images as well. With the use of pandas, seaborn and also scikit-learn, Keras, and Tensorflow libraries I calculated different metrics for each models in order to find the best performing deep learning model.

## Algorithms

### *Feature Engineering*

1. Regularization ( Image Augmentation , early stopping)
2. Transfer Learning ( MobileNetV2)
3. Use of extra test data.

### *Model Evaluation and Selection*

In the conclusion, using deep learning models and finding the best performing model by comparison of the F1 metric, we can provide the medical field with a diagnose machine learning model that would not only increase accuracy and speed of chest X-rays of Pneumonia but also decrease the error human error.

### Tools

- numpy, pandas
- Seaborn, matplotlib
- Scikit-learn , Keras, tensorflow

## Communication

There is going to be the code for the data operations and also PDF slides of powerpoint presentation available on my GitHub account.

[https://github.com/rezxoiki/Deep\\_Learning\\_Chest\\_Xray](https://github.com/rezxoiki/Deep_Learning_Chest_Xray)