**Project Report**

**On**

# **Deep Learning for Detecting Pneumonia from X-ray Images**



**Submitted by**

**Shruti Mukherjee**

**Roll No - 1828203**

**School of Computer Engineering , KIIT Deemed to be University,**

**Bhubaneswar, India**

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Topic No** | **Topics** | **Page no** |
| 1 | Abstract | 3 |
| 2 | Introduction | 4 |
| 3 | About Pneumonia | 5 |
| 4 | About Data | 7 |
| 5 | Related Work | 8 |
| 6 | Knowledge about software | 9 |
| 7 | Pneumonia Detection System | 10 |
| 8 | Results and Discussion | 17 |
| 9 | Future Work | 24 |
| 10 | Conclusion | 25 |
| 11 | References | 26 |

1. **Abstract:**

Pneumonia causes the death of around 700,000 children every year and affects 7% of the global population. Chest X-rays are primarily used for the diagnosis of this disease. However, even for a trained radiologist, it is a challenging task to examine chest X-rays. There is a need to improve the diagnosis accuracy. In this work, an efficient model for the detection of pneumonia trained on digital chest X-ray images is proposed, which could aid the radiologists in their decision making process. A novel approach based on a weighted classifier is introduced, which combines the weighted predictions from the state-of-the-art deep learning models such as ResNet18, Xception, InceptionV3, DenseNet121, and MobileNetV3 in an optimal way. This approach is a supervised learning approach in which the network predicts the result based on the quality of the dataset used. Transfer learning is used to fine-tune the deep learning models to obtain higher training and validation accuracy. Partial data augmentation techniques are employed to increase the training dataset in a balanced way. The proposed weighted classifier is able to outperform all the individual models. Finally, the model is evaluated, not only in terms of test accuracy, but also in the AUC score. The final proposed weighted classifier model is able to achieve a test accuracy of 98.43% and an AUC score of 99.76 on the unseen data from the Guangzhou Women and Children’s Medical Center pneumonia dataset. Hence, the proposed model can be used for a quick diagnosis of pneumonia and can aid the radiologists in the diagnosis process.

****Keywords:****pneumonia, chest X-ray images, convolution neural network (CNN), deep learning, transfer learning, computer-aided diagnostics

1. **Introduction:**

Pneumonia is an acute respiratory infection that affects the lungs. It is a fatal illness in which the air sacs get filled with pus and other liquid. There are mainly two types of pneumonia: bacterial and viral. Generally, it is observed that bacterial pneumonia causes more acute symptoms. The most significant difference between bacterial and viral pneumonia is the treatment. Treatment of bacterial pneumonia is done using antibiotic therapy, while viral pneumonia will usually get better on its own. It is a prevalent disease all across the globe. Its principal cause includes a high level of pollution. Pneumonia is ranked eight in the list of the top 10 causes of death in the United States. Due to pneumonia, every year, 3.7 lakh children die in India, which constitutes a total of fifty percent of the pneumonia deaths that occur in India. The disease frequently goes overlooked and untreated until it has reached a fatal point, especially in the case of old patients. It is the single largest cause of death in children (especially under the age of five) worldwide. According to the WHO, “Every year, it kills an estimated 1.4 million children under the age of five years, accounting for 18% of all deaths of children under five years old worldwide. Pneumonia affects children and families everywhere but is most prevalent in South Asia and sub-Saharan Africa. Children can be protected from pneumonia. It can be prevented with simple interventions and treated with low-cost, low-tech medication and care”. Therefore, there is an urgent need to do research and development on computer-aided diagnosis so that the pneumonia-related mortality, especially in children, can be reduced.

One of the following tests can be done for pneumonia diagnosis: chest X-rays, CT of the lungs, ultrasound of the chest, needle biopsy of the lung, and MRI of the chest. Currently, chest X-rays are one of the best methods for the detection of pneumonia. X-ray imaging is preferred over CT imaging because CT imaging typically takes considerably more time than X-ray imaging, and sufficient high-quality CT scanners may not be available in many underdeveloped regions. In contrast, X-rays are the most common and widely available diagnostic imaging technique, playing a crucial role in clinical care and epidemiological studies. There are several regions across the globe where there is a scarce availability of practiced healthcare workers and radiologists whose prediction on such diseases matter greatly. Computer-aided diagnosis using artificial intelligence based solutions is becoming increasingly popular these days. This facility can be made available to a large population at a minimal cost. Another issue with this disease is that sometimes, the features that describe the very existence of the disease often get mixed with other diseases, and hence, radiologists find it challenging to diagnose this disease. Deep learning techniques solve all these problems, and their accuracy in the prediction of the disease is the same and sometimes even greater than an average radiologist. Among the deep learning techniques, convolutional neural networks (CNNs) have shown great promise in image classification and segmentation and therefore are widely adopted by the research community. Biomedical image diagnosis that uses the techniques of deep learning and computer vision has proven to be very helpful to provide a quick and accurate diagnosis of the disease that matches the accuracy of a reliable radiologist. Currently, deep learning based methods cannot replace trained clinicians in medical diagnosis, and they aim to supplement clinical decision making. In this paper, a model is presented based on the applications of deep learning and convolutional neural networks that are capable of classifying automatically that the patient has pneumonia or not. The proposed methodology uses a deep transfer learning algorithm that extracts the features from the X-ray image that describes the presence of disease automatically and reports whether it is a case of pneumonia.

Physicians often use chest X-rays to quickly and cheaply diagnose disease associated with the area. However, it is much more diffificult to make clinical diagnoses with chest X-rays than with other imaging modalities such as CT or MRI. With computer-aided diagnosis, physicians can make chest X-ray diagnoses more quickly and accurately. Pneumonia is often diagnosed with chest X-Rays and kills around 50,000 people each year. With computeraided diagnosis of pneumonia specififically, physicians can more accurately and effificiently diagnose the disease. In this project, we hope to train a model using the dataset described below to help physicians in making diagnoses of pneumonia in chest X-Rays. Our problem is thus a binary classifification where the inputs are chest X-ray images and the output is one of two classes: pneumonia or non-pneumonia.

1. **ABOUT PNEUMONIA**

1. Pneumonia is a common infection of lunge effecting mostly the microscopic air sacs. In people with pneumonia, the air sace are filled with fluid or pus, hindering the gas exchange process, resulting in difficulty breathing and cough reflex.

Other symptoms may include:-

a. Chest pain

b. Fever

c. Chills

d. Confusion

2. Various organisms can cause pneumonia like-

a. Bacterial pneumonia is the most common

b. Streptococcus pneumoniae

c. Viral pneumonia is more common in young children.

3.Pneumonia commonly starts as an infection of the upper respiratory tract through a cold or flu, which then spreads to the lungs. The most common transmission route is through inhalation of contaminated aerosol droplets and aspiration of oral bacteria into the lungs.

4.The setting in which pneumonia develops is an important information as it helps to identify the source of the causative agent and hence the treatment approach. Generally, community acquired pneumonia is less dangerous than healthcare associated, hospital acquired, or ventilator associated pneumonia. This is because infections acquired outside health care facilities are less likely to involve multidrug resistant bacteria. Patients who are in hospitals are also more likely to have other health problems and weakened immune system, thus they are less able to fight the disease.

5. Pneumonia is often diagnosed based on physical exams and a chest X-ray. In children diagnosis is primarily based on rapid respiratory rate, a cough, the presence of lower chest wall indrawing, and through the level of consciousness. In adults vital signs and presence of chest crackles are checked.

**2.2 SYMPTOMS**

The symptoms of Pneumonia can be very severe from person to person. Here are some of the most common examples:-

a. Cough

b. Fever and Chills

c. Shortness of breath

d. Fatigue

e. Cyanosis

f. Diaphoresis

**2.3 CURE**

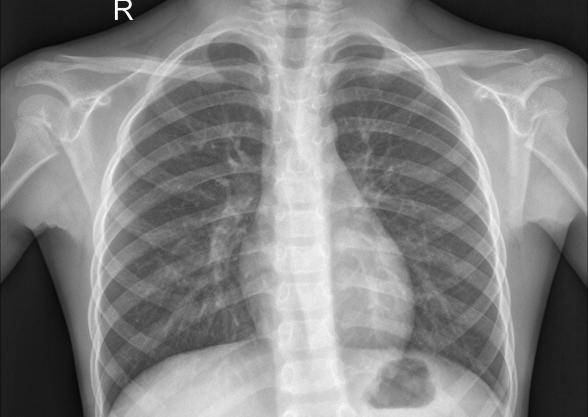
Bacterial pneumonia is treated with antibiotics. The dose and type of antibiotics depends on the age of the patient, health conditions of the patient and how the infection was acquired. Viral pneumonia caused by influenza viruses may be treated with antiviral drugs. In severe cases with breathing difficulty and other health problems hospitalization may be required.

1. **Data:**

We use a dataset compiled by the NIH which contains 112,120 chest X-ray images from 30,805 unique patients . The dataset is available from Kaggle. Each image in the dataset is labeled with one or more diagnoses (“Pneumonia”, “Fibrosis”, “Mass”, etc), or “No fifinding” if the patient was healthy. These labels were inferred through natural language processing by mining disease classifification from the associated radiological reports, and are estimated to be at least 90% accurate. For the sake of this project, we follow past approaches and treat the labels as ground truth for the purpose of classifification. For this project, we focus on binary classifification, attempting to classify a particular X-Ray as having pneumonia or not. There is a strong class imbalance in the dataset, with only about 1% of images labeled as having pneumonia. For this reason, we us AUC as our error metric, rather than accuracy. Images from the NIH dataset are 1024x1024. To begin, we resized each image using an anti-aliasing fifilter.

As explained below, our Logistic Regression baseline uses 32x32 resolution. Our Deep Learning model uses 224x224 resolution. We also standardize the data so that each feature (each pixel) has zero mean and approximately unit variance. There are multiple images from many of the patients,

and in order to ensure that our models do not see data from the same patient across training and test, we separate the data by patient before splitting into training, validation, and test sets. We begin with a brief exploration of the data using unsupervised techniques. For the sake of data exploration, we explore 500 random samples due to resource constraints.



1. **Related Work:**

Deep learning based methods are already being used in various fields. Different authors have already proposed several biomedical image detection techniques. M.I.Razaak discussed the challenges and the future of medical image processing. Much work has already been done for the detection of numerous diseases by using deep learning based techniques, as stated by Dinggang Shen. Andre presented a deep learning model for dermatologist-level classification of skin cancer, and F.Milletari also proposed a methodology for the depiction of prostrate in MRI volumes using CNN. Grewal used the technique of deep learning for brain hemorrhage detection in CT scans, and Varun proposed a method for detecting diabetic retinopathy in retinal fundus photographs. Y. Bar also discussed chest pathology detection by the techniques based on deep learning. Methods regarding the examination of the detection of disease by chest X-ray have also been worked on earlier by performing various examination techniques. The chest X-ray images are passed through the evaluation process of scan line optimization such that it eliminates all the other body parts to avoid any error in diagnosis. The algorithm was described by S. Hermann. Nasrullah et al. used two deep three-dimensional (3D) customized mixed link network (CMixNet) architectures for lung nodule detection and classification. Yao L et al. combined DenseNet and long-short term memory networks (LSTM) to exploit the dependencies between abnormalities. Several authors also have worked on pneumonia classification. Khatri et al. proposed the use of EMD (earth mover’s distance) to identify infected pneumonia lungs from normal non-infected lungs. Rahib et al. and Okeke et al. used a CNN model for pneumonia classification. Some researchers have shown assuring results such as Cohen et al. and Rajaraman et al. Rajaraman et al. tried to explain the performance of customized CNNs to detect pneumonia and further differentiate between bacterial and viral types in pediatric CXRs. Sirazitdinov et al. used a region based convolutional neural network for segmenting the pulmonary images along with image augmentation for pneumonia identification. Lakhani and Sundaram used the AlexNet and GoogLeNet neural networks with data augmentation and without any pre-training to obtain an area under the curve (AUC) of 0.94–0.95. Rajpurkar et al. used CheXNeXt, a very deep CNN with 121 layers, to detect 14 different pathologies, including pneumonia, in frontal-view chest X-rays. A localization approach based on pre-trained DenseNet-121, along with feature extraction, was used to identify 14 thoracic diseases in . Saraiva et al. , Ayan et al., and Rahman et al. used deep learning based methods for pneumonia classification. Xiao et al. proposed a novel multi-scale heterogeneous three-dimensional (3D) convolutional neural network (MSH-CNN) based on chest computed tomography (CT) images. Xu et al. used a hierarchical convolutional neural network (CNN) structure and a novel loss function, sin-loss, for pneumonia detection. Jaiswal et al. used Mask-RCNN, utilizing both global and local features for pulmonary image segmentation, with dropout and L2 regularization, for pneumonia identification. Jung et al. used a 3D deep CNN (3D DCNN), which had shortcut connections. Vikash et al. combined the outputs of different neural networks and reached the final prediction using majority voting. None of the above-mentioned approaches except that of Vikash et al. tried to combine predictions from different neural networks.

The main contribution is a weighted classifier that integrates five deep learning models. The weights for each model are based on each model’s performance on the testing dataset.

**6.Knowledge about Software:**

**INTRODUCTION ABOUT ANACONDA AND JUPYTER**

Anaconda is nothing but a bunch of popular python packages. And a packet manager called conda(similar to pip). this python package are very popular in data science communities. Some of the popular packages are numpy, scipy, jupyter, nltk, scikit-learn etc. anaconda consist several python libraries. A light weight version of anaconda is also available called mini conda.in addition anaconda supplant their own package called conda. It is very efficient than PIP. Jupyter is a interpreter that is based on browser that help you to work on python and R. anaconda consist jupyter libraries. You can consider jupyter as a notebook which is digital that provides you an chance to execute commands draw charts and takes notes. Data scientist used this as on prior basis. This is very helpful tool if you are learning python and R. jupyter is much better than shell. Jupyter is amazing tool for the analytical work where you could show your code in “modules” adding common formatting option between modules and include of formatted output of modules and generate the graph in well suited manner in other modules code. Jupyter assure reproducibility in other’s work. Therefore if someone come back after few months then by seeing the code he/she will easily get what someone has tried to do. And can exactly tell which code run which conclusion and visualization.

**7 PNEUMONIA DETECTION SYSTEM**

**7.1 MACHINE LEARNING**

Machine learning (ML) is the study of computer algorithms that automatically improves through experience by the use of data. It is a part of artificial intelligence. Machine learning algorithms is build on a model based on sample data, known as the **training data**, in order to make predictions or decisions without being clearly programmed to do required tasks. Machine learning algorithms have a wide variety of applications, such as email filtering and computer vision, where it is difficult or impracticable to develop conventional algorithm methods to perform the required tasks.

A subdivision of machine learning is closely related to computational statistics, which uses computers to focus on making predictions; but not all machine learning is statistical or arithmetical learning. Mathematical optimization study delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, which focuses on exploratory data analysis using unsupervised learning. In its application across business problems, machine learning is also referred to as predictive analytic.

Machine learning methods are basically divided into three broad categories, depending on the nature of the ‘feedback’ or ‘signal’ which are available to the learning system:

* **Supervised learning**

The computer is introduced with example inputs and their desired outputs, given by a ‘teacher’, and the goal is to learn a general rule that maps inputs to the outputs.

A support-vector machine or SVM is a supervised learning model that divides the respective data into regions separated by a linear boundary. Here, the linear boundary separates the black circles from the white.

Supervised learning algorithms helps to build a mathematical model of a set of data that contains both the inputs and their desired outputs. The data is known as training data, and they consist of a set of training examples. Each training example has one or more inputs and the desired output,which is also known as a supervisory signal. In the mathematical model, each training example is represented by an array or a vector, sometimes called a feature vector, and the training data is represented by a matrix. Through repetitive optimization of an objective function, supervised learning algorithms learn a function which can be used to predict the output associated with the new inputs . An ideal function will allow the algorithm to correctly determine the output for inputs that were not a part of the training data. An algorithm will improve the accuracy of its output or predictions over time it is said to have learned to perform that particular task.

Different types of supervised learning algorithms includes:

* Active learning
* Classification
* Regression
* **Unsupervised learning**

There are no labels given to the learning algorithm, leaving it on its own to find the structure in its input. Unsupervised learning can be a goal in itself (discovering hidden patterns in the data) or a medium towards the end (feature learning).

Unsupervised learning algorithms takes a set of data that contains only inputs, and find the structure in the data, like grouping or clustering of the data points. The algorithms, therefore, learn from the test data that has not been labeled, classified or categorized. Instead of acknowledging the feedback, unsupervised learning algorithms identify similarities in the data and reaction based on the presence or absence of such similarities in each new piece of the data. A central application of unsupervised learning is in the field of density estimation in statistics, such as finding the probability density function.Though unsupervised learning encloses other domains involving summarizing and explaining the data features.

Cluster analysis is the task of converting a set of observations into subsets known as clusters so that observations within the same cluster are similar according to one or more predesignated criteria, while observations drawn out from the different clusters are dissimilar. Different clustering techniques make different assumptions on the structure of the data, often defined by some similarity metric and evaluated, for example, by internal denseness, or if their is similarity between members of the same cluster, and separation, the difference between the clusters. Other methods are based on the estimated density and graph connectivity.

* **Semi-supervised Learning**

Semi-supervised learning falls between the unsupervised learning (without any labeled training data) and the supervised learning (with completely labeled training data). Some of the training examples have missing training labels, yet many machine-learning researchers have found that unlabeled data, when used in conjunction with a small amount of labeled data, can produce appreciable improvement in learning accuracy.

* **Reinforcement Learning**

Reinforcement learning is a subset of machine learning which is concerned with how software agents ought shall take actions in environment in order to maximize some conception of cumulative reward. Due to its widespread, the field is studied in many other ways and formats, such as game theory, control theory, operations research, information theory, simulation-based optimization, multi-agent systems, swarm intelligence, statistics and genetic algorithms. In machine learning, the whole environment is typically represented as a Markov Decision Process (MDP). Most of the reinforcement learning algorithms uses dynamic programming techniques. Reinforcement learning algorithms don’t assume knowledge of an exact mathematical model of the MDP, and are used when exact models are improbable. Reinforcement learning algorithms are used in self-determining vehicles or learning to play game against a human opponent.

* **AI and Machine Learning**

A lot of sources continue to assert that machine learning remains a sub-field of AI. The main conflict point is whether all of ML is part of AI, as this would mean that anyone who is using ML could easily claim they are using AI. Others have the view that not all of ML is part of AI but only an 'intelligent' subset of ML is a part of AI.

According to the ‘The Book of Why’ by Judea Pearl, ML learns and predicts on the basis of yielding observations, whereas AI implies an agent interacting with the environment to learn and how to take actions that maximize its chance of achieving its goals successfully.

* **Data Mining**

Machine learning and Data Mining often make use of the same methods and overlay significantly, but we know that machine learning focuses on prediction, based on known properties that are learned from the training data, whereas data mining focuses on the revelation of (previously) unknown properties in the data (this is the step where we analyse the knowledge discovery in databases). Data mining uses many machine learning methods, but having different goals; on the other hand, machine learning also uses data mining methods as ‘unsupervised learning’ or as pre-processing step to improve the learner accuracy.

* **Optimization**

Machine learning also has a familiar relationship with optimization, many learning problems are drawn up as minimization of some loss functions on a training set of examples. Loss functions express the disparity between the predictions of those model being trained and the actual problem instances.

* **Generalization**

The difference between optimization and machine learning arises from the goal of generalization: while optimization algorithms can cut down the loss on a training set, machine learning is concerned with minimizing the loss on unseen samples. Characterizing the generalization of various learning algorithms is a topic of conversation of current research, especially for deep learning algorithms.

* **Statistics**

Machine learning and statistics are closely related fields in terms of methods, but are different in their principal goal: statistics draws population inferences from a sample, while machine learning finds generalized predictive patterns.

**Transfer Learning**

Transfer learning is a part of machine learning in which once we train the model, the model learns from the training and the output. Next time when we give a similar model as the previous one , we do not need to train the model again it would work by itself using things it has learned during the training of previous model.

## 7.2 PROGRAMMING LANGUAGES FOR MACHINE LEARNING

**There are various programming languages for machine learning like:**

1. ****Python****

Python is high level programming language. It is a powerful and versatile programming language. In python there is no need to compile the code because an interpreter is used which takes care of the testing and debugging at a very high speed. An open source library is available for python. Since it is a very popular programming language therefore it can be used in web development, software development, system scripting. It works easily on any platform such as R, Raspberry, windows etc. syntax of python is similar as to English language which helps the programmer to write the less lines of code as compared to other programming languages. The most updated version of python is python3. It is the updated version of python2 which is quite popular.

**It’s** popularity is due to the increased development and use of deep learning frameworks that are available for this language recently, including [TensorFlow](https://www.tensorflow.org/), [PyTorch](http://pytorch.org/), and [Keras](https://keras.io/). As a language that has readable syntax and the ability to be used as a scripting language, Python proves to be powerful and very straightforward both for pre-processing data and working with the data directly. The [scikit-learn](http://scikit-learn.org/stable/) machine learning library is built on top of several existing Python packages that the Python developers are already familiar with, namely [NumPy](http://www.numpy.org/), [SciPy](https://www.scipy.org/), and [Matplotlib](https://matplotlib.org/).

1. ****Java****

It is widely used in enterprise programming, and is generally used by front-end desktop application developers who are also working on machine learning at the enterprise level. Among machine learning libraries for Java are deep learning an open-source and distributed deep-learning library written for both Java and Scala ; [MALLET](http://mallet.cs.umass.edu/) (****MA****chine ****L****earning for ****L****anguag****E**** ****T****oolkit) allows for machine learning applications on text, which includes natural language processing, topic modeling, document classification, and clustering; and [Weka](http://www.cs.waikato.ac.nz/ml/weka/index.html), which is a collection of machine learning algorithms used for data mining tasks.

****c. C++****

It is the language of choice for machine learning and artificial intelligence in game or robot applications (including robot locomotion). Embedded computing hardware developers and electronics engineers are more likely to favor C++ or C in machine learning applications due to their proficiency and level of control in the language. Some machine learning libraries you can use with C++ include the scalable [mlpack](http://www.mlpack.org/), [Dlib](http://dlib.net/ml.html) offering wide-ranging machine learning algorithms, and the modular and open-source [Shark](http://image.diku.dk/shark/).

Deep learning is another subset of machine learning, which utilizes a hierarchical level of artificial neural networks to carry out various processes of machine learning. The artificial neural networks are built like a human brain, with neuron nodes connected together like a web. While traditional programs build analysis with data in a linear way, the hierarchical function of deep learning systems helps machines to process data with a unpredictable approach.

**DEEP LEARNING**

Deep learning is another subset of machine learning, which utilizes a hierarchical level of artificial neural networks to carry out various processes of machine learning. The artificial neural networks are built like a human brain, with neuron nodes connected together like a web. While traditional programs build analysis with data in a linear way, the hierarchical function of deep learning systems helps machines to process data with a unpredictable approach.

Deep Learning is a sub field of machine learning . It contains the algorithms related or inspired to the structure and function of the brain called artificial neural networks. One of the  benefit of deep learning is their ability to perform automatic feature extraction from raw data, known as feature learning. Feature learning or representation learning is a set of techniques that learn a feature and applies it. Feature learning is divided into 2 categories that are supervised and unsupervised feature learning.

* In Supervised feature learning the features are learned with defined input data. Example:- Multi layer Perceptron
* In unsupervised feature learning the features are learned with undefined input data. Example:- dictionary learning

**CNN ( Convolutional Neural Networks)**

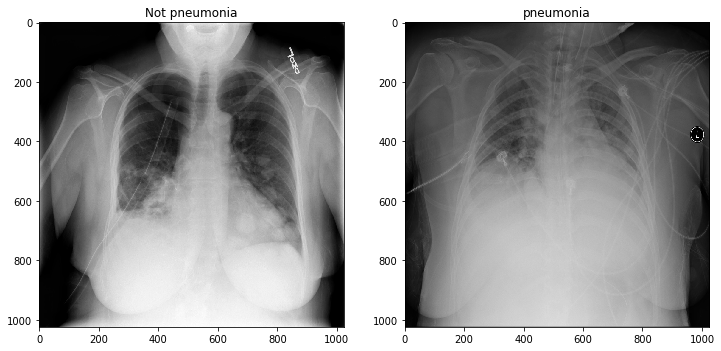
Its is a deep learning algorithm which is able to take images as input and can diferentiate them from each other.

Result and discussion:

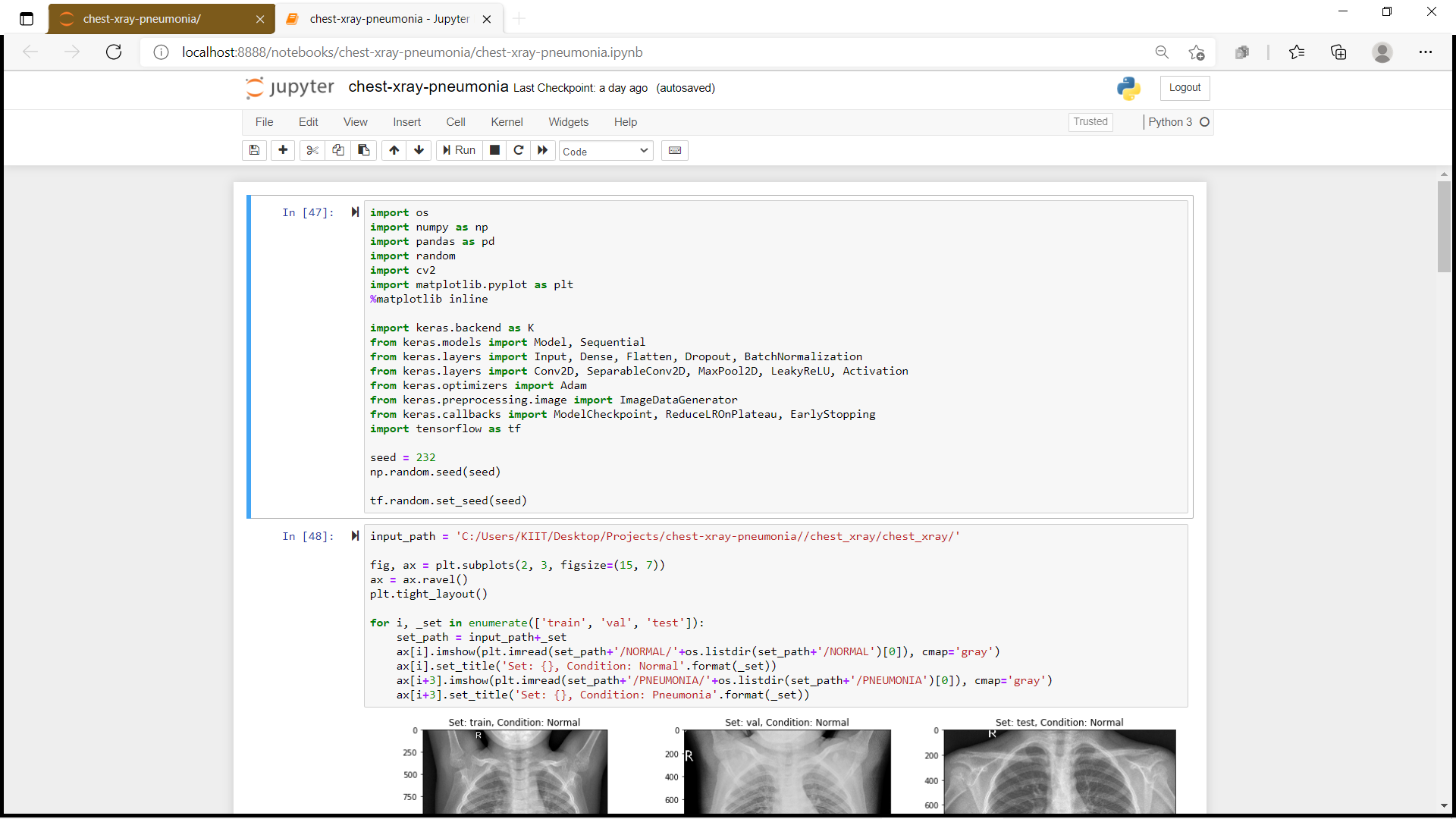
The risk of pneumonia is immense for many, especially in developing nations where billions face energy poverty and rely on polluting forms of energy. The WHO estimates that over 4 million premature deaths occur annually from household air pollution-related diseases including pneumonia. Over 150 million people get infected with pneumonia on an annual basis especially children under 5 years old. In such regions, the problem can be further aggravated due to the dearth of medical resources and personnel. For example, in Africa’s 57 nations, a gap of 2.3 million doctors and nurses exists. For these populations, accurate and fast diagnosis means everything. It can guarantee timely access to treatment and save much needed time and money for those already experiencing poverty.

**Problem Statement**

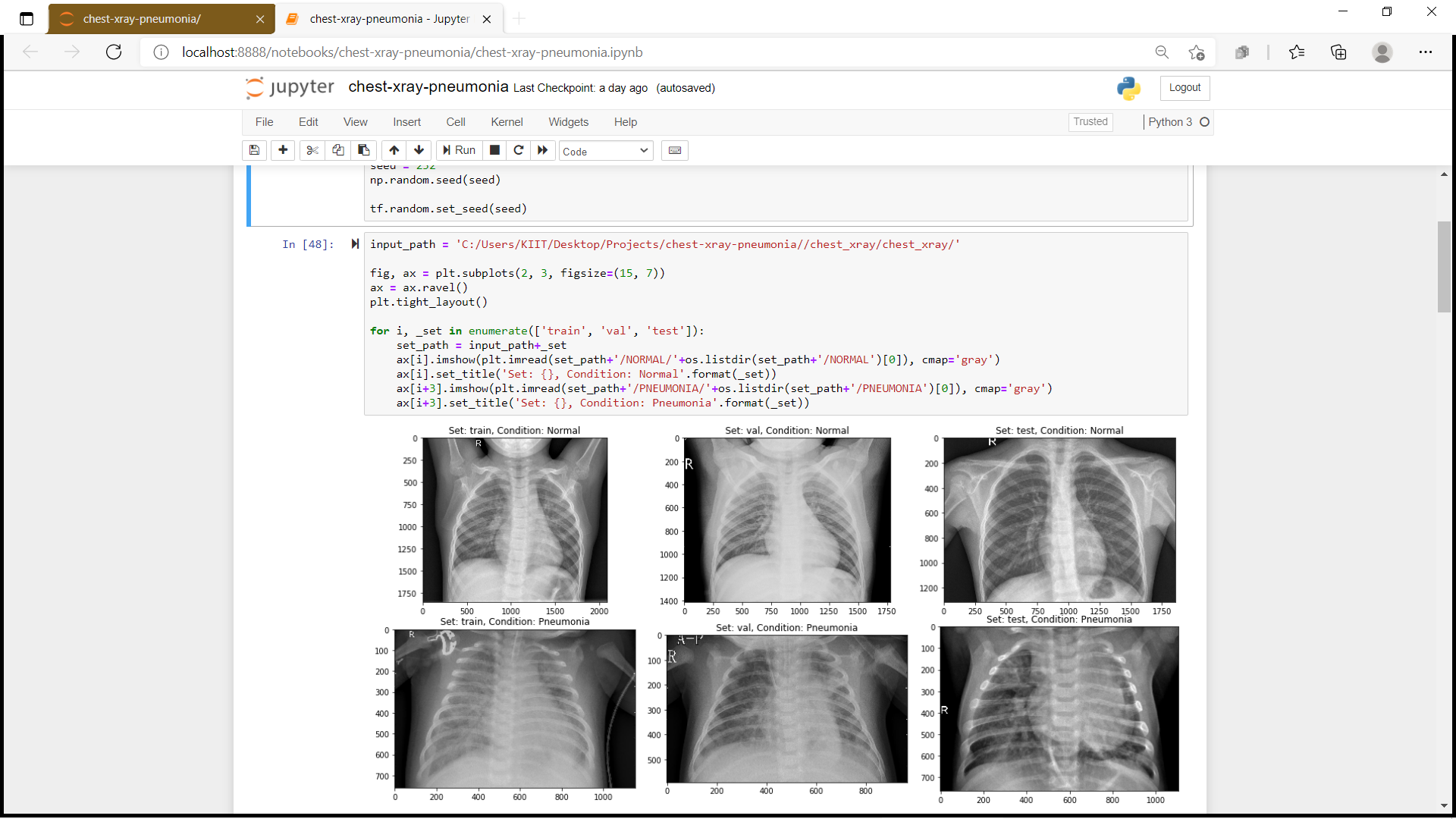
Build an algorithm to automatically identify whether a patient is suffering from pneumonia or not by looking at chest X-ray images. The algorithm had to be extremely accurate because lives of people is at stake.



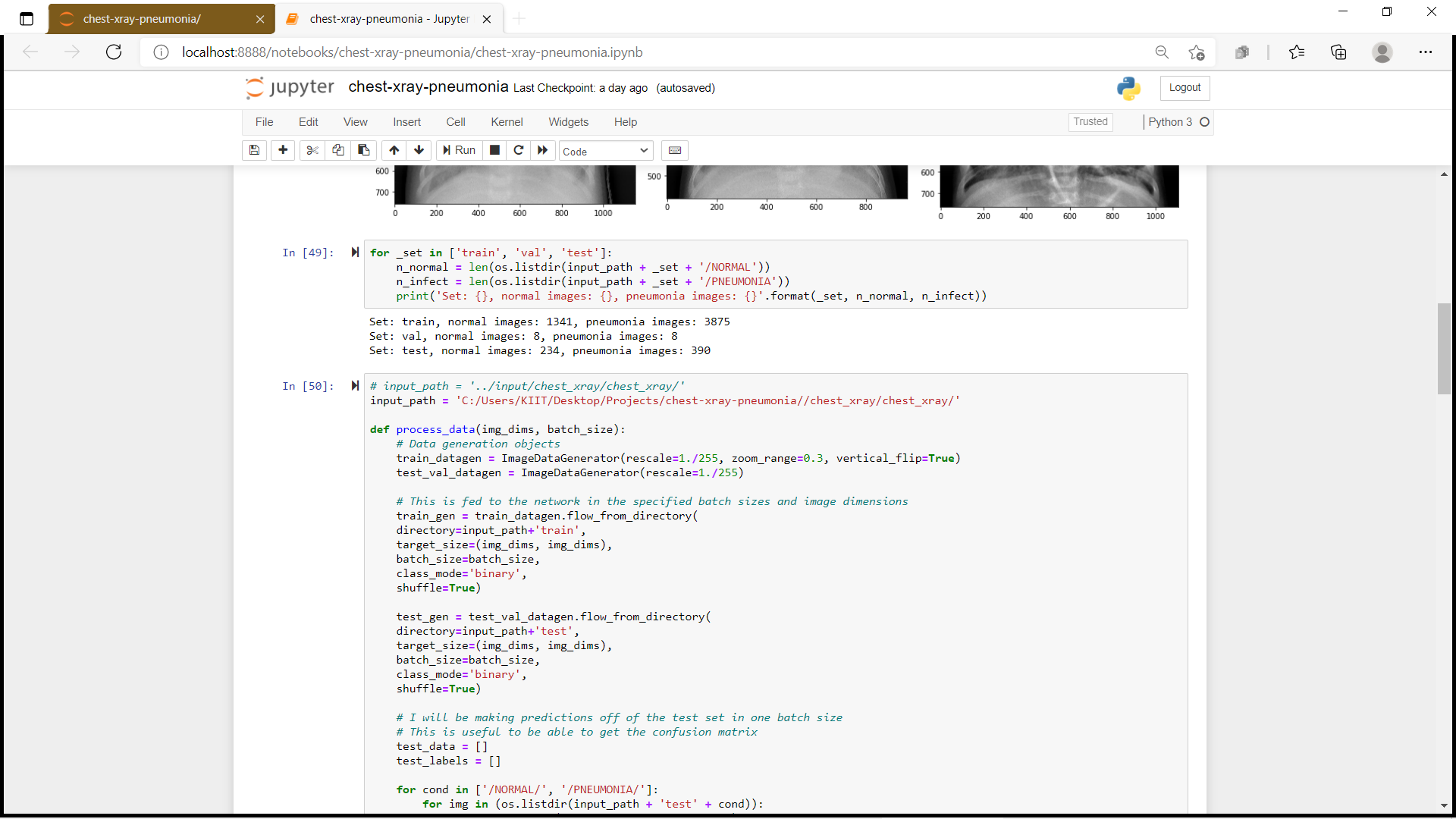
**Results and Discussion:**



Next I displayed some normal and pneumonia images to just have a look at how much different they look from the naked eye. Well not much!



Then I split the data-set into three sets — train, validation and test sets.

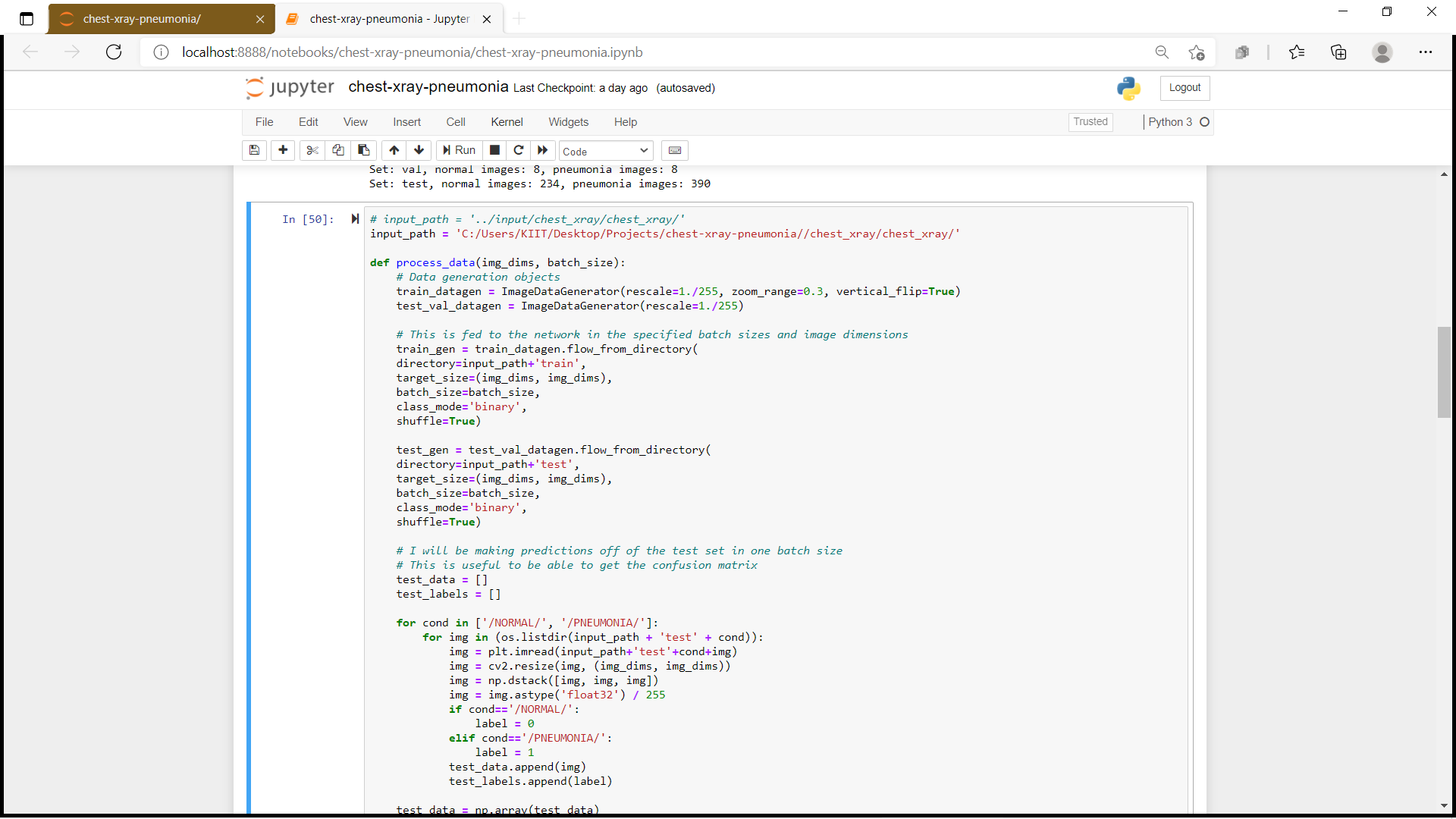


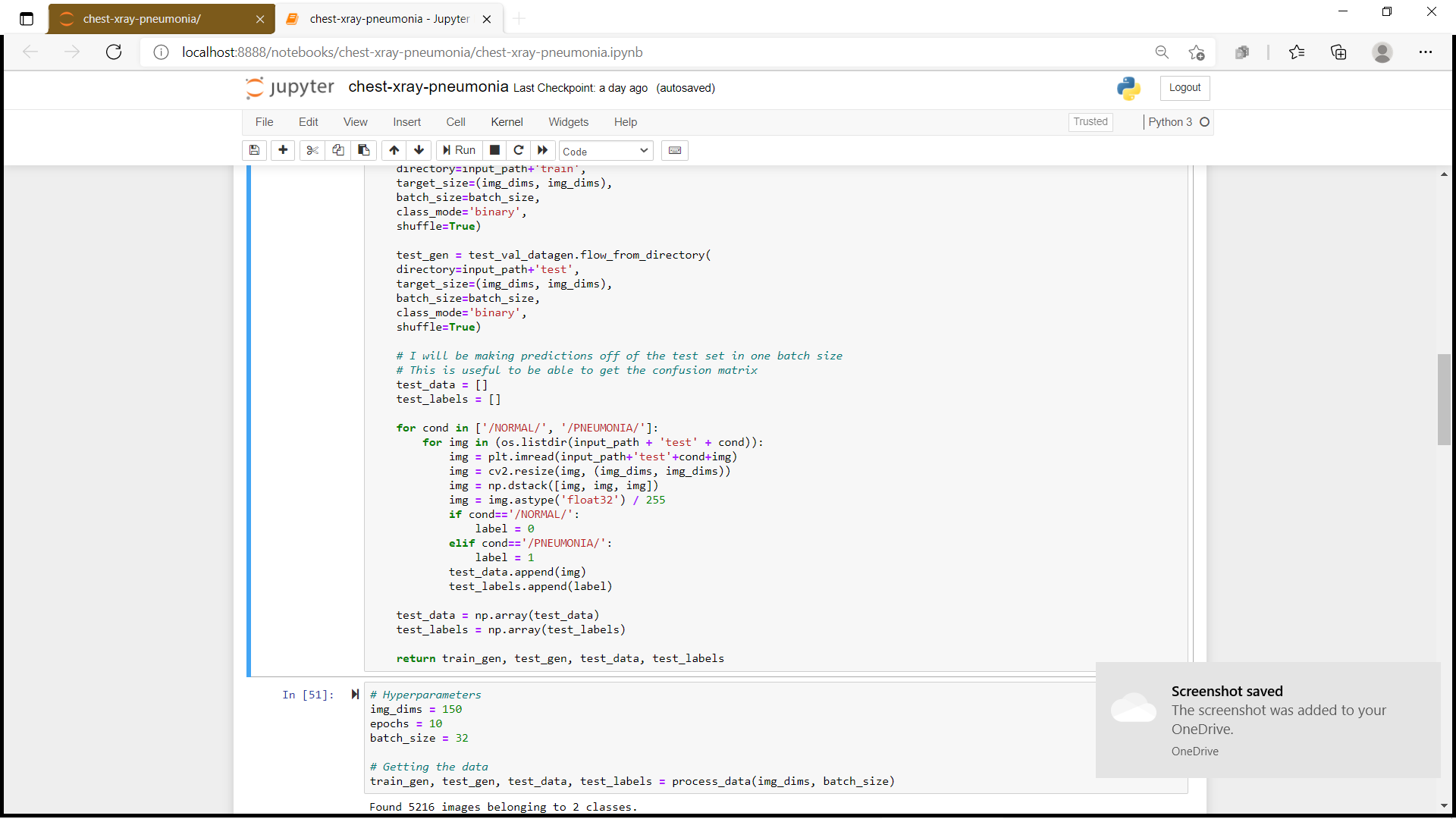
Next I wrote a function in which I did some data augmentation, fed the training and test set images to the network. Also I created labels for the images.

The practice of data augmentationis an effective way to increase the size of the training set. Augmenting the training examples allow the network to “see” more diversified, but still representative, data points during training.

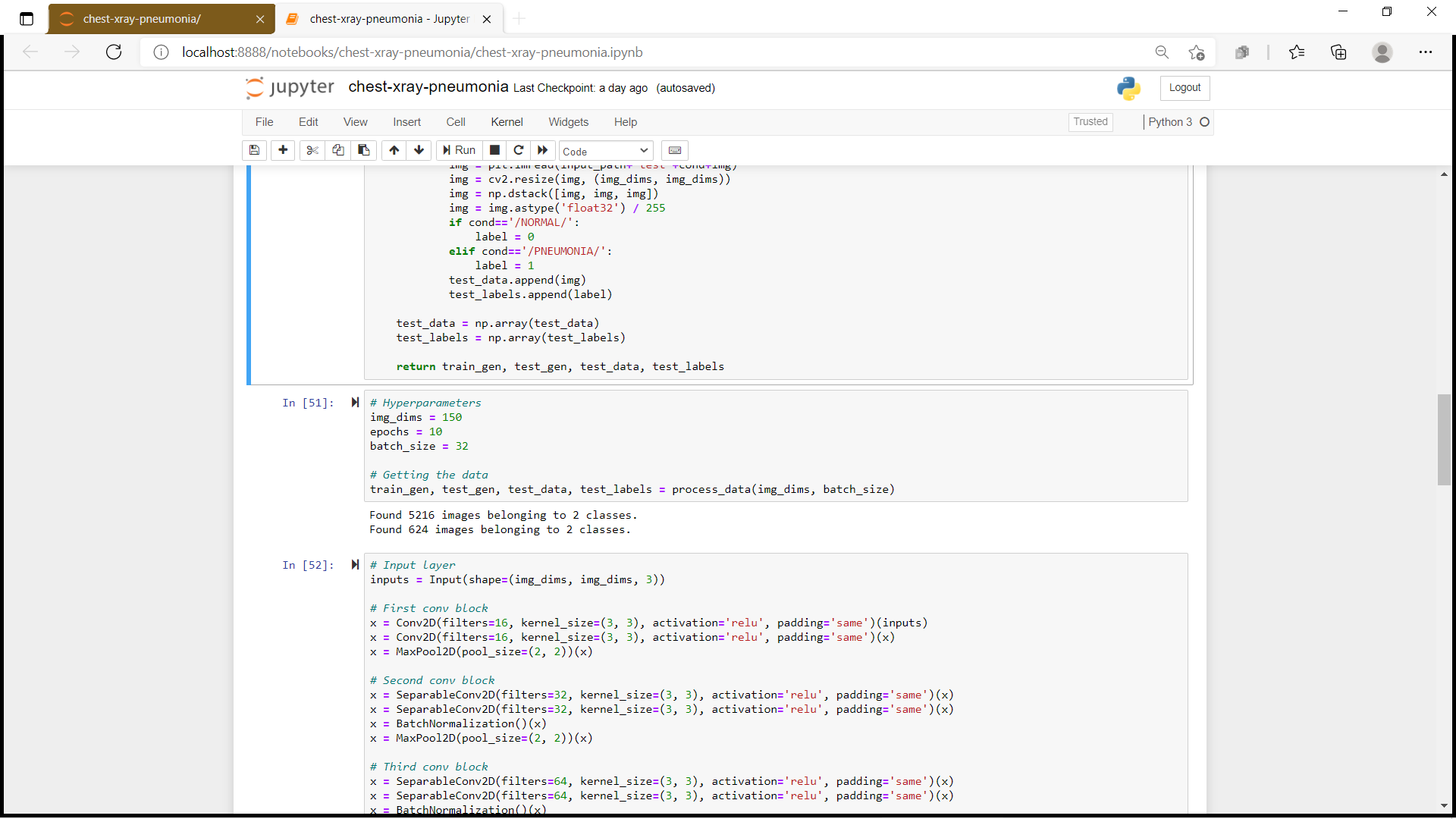
Then I defined a couple of data generators: one for training data, and the other for validation data. A data generatoris capable of loading the required amount of data (a mini batch of images) directly from the source folder, convert them into *training data*(fed to the model) and *training targets*(a vector of attributes — the supervision signal).

For my experiments, I usually set the batch\_size = 64. In general a value between 32 and 128 should work well. Usually you should increase/decrease the batch size according to computational resources and model’s performances.





After that I defined some constants for later usage.

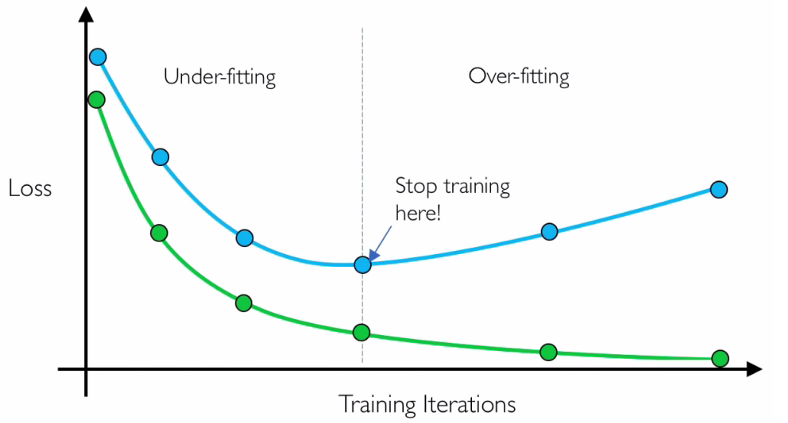


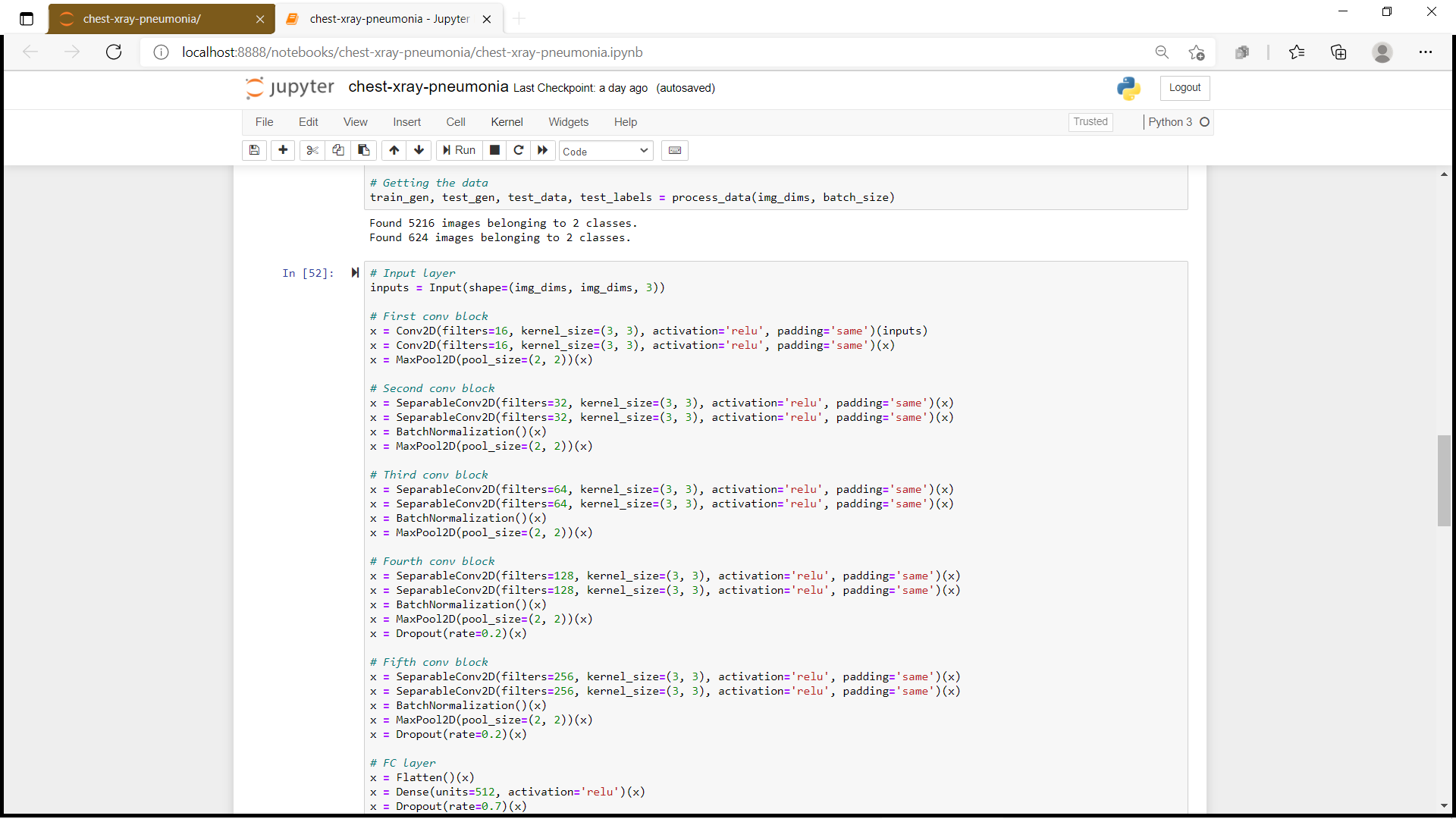
The next step was to build the model. This can be described in the following 5 steps.

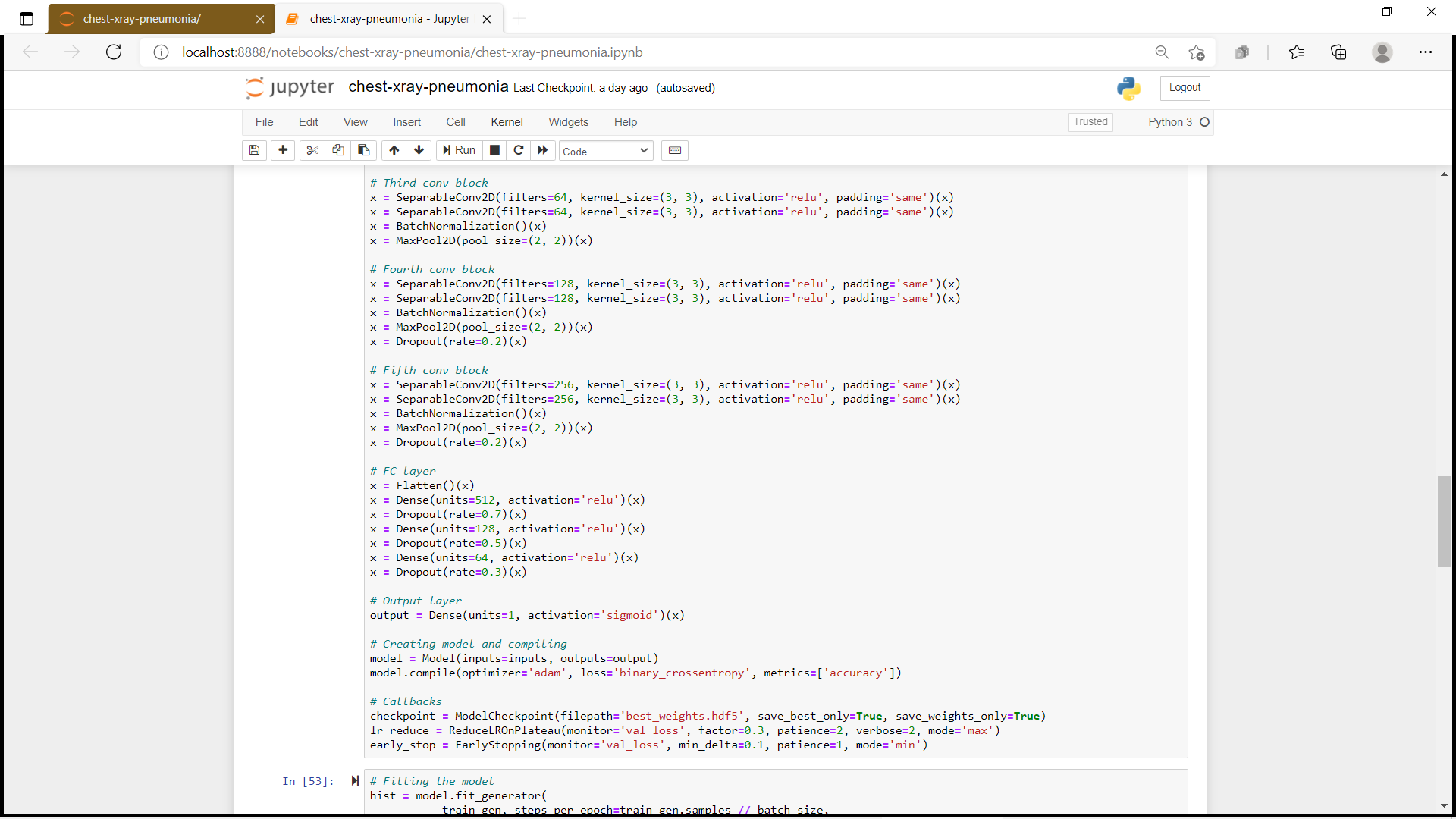
1. I used five convolutional blocks comprised of convolutional layer, max-pooling and batch-normalization.
2. On top of it I used a flatten layer and followed it by four fully connected layers.
3. Also in between I have used dropouts to reduce over-fitting.
4. Activation function was Relu throughout except for the last layer where it was Sigmoid as this is a binary classification problem.
5. I have used Adam as the optimizer and cross-entropy as the loss.

Before training the model is useful to define one or more callbacks. Pretty handy one, are: ModelCheckpoint and EarlyStopping.

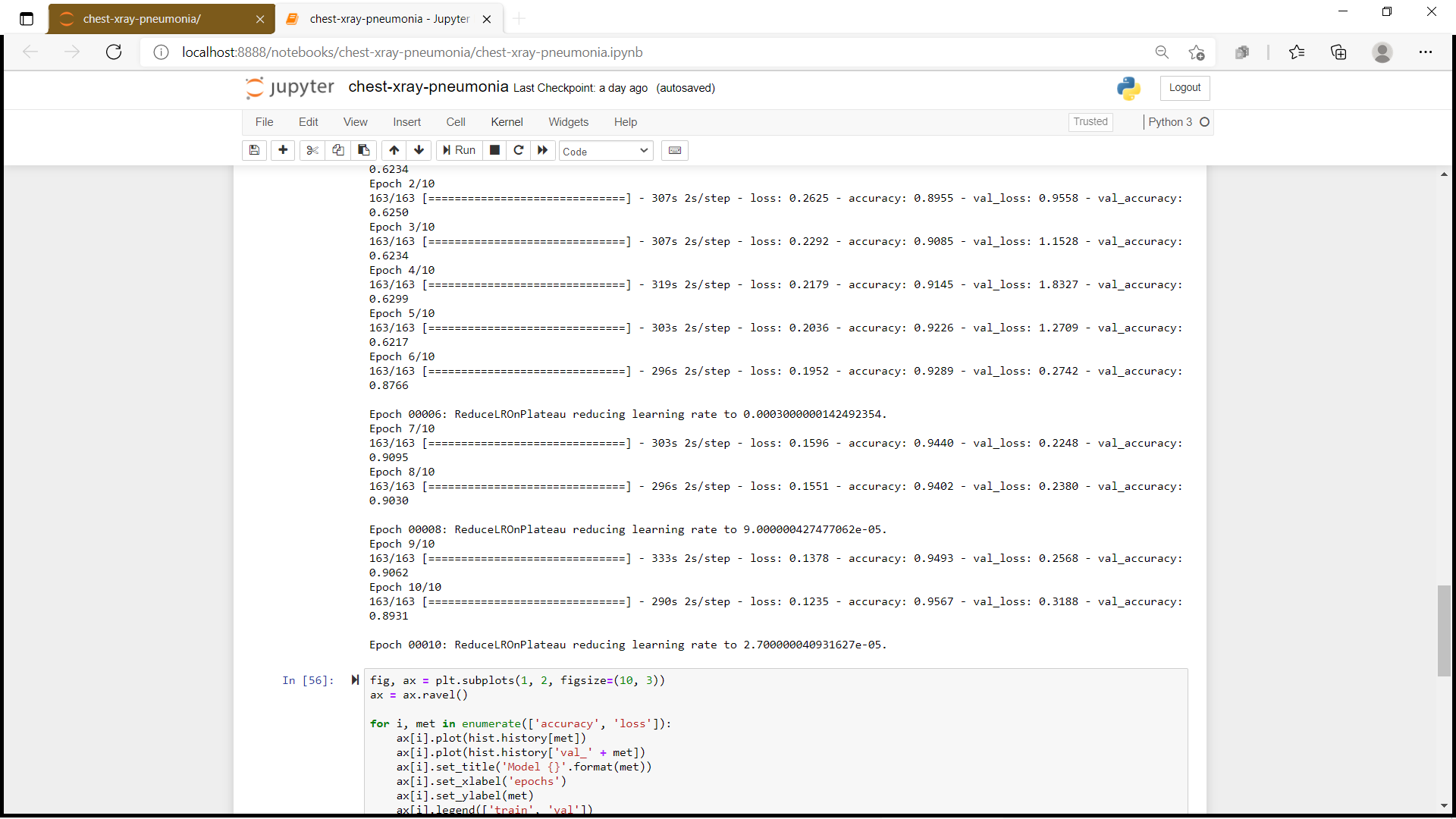
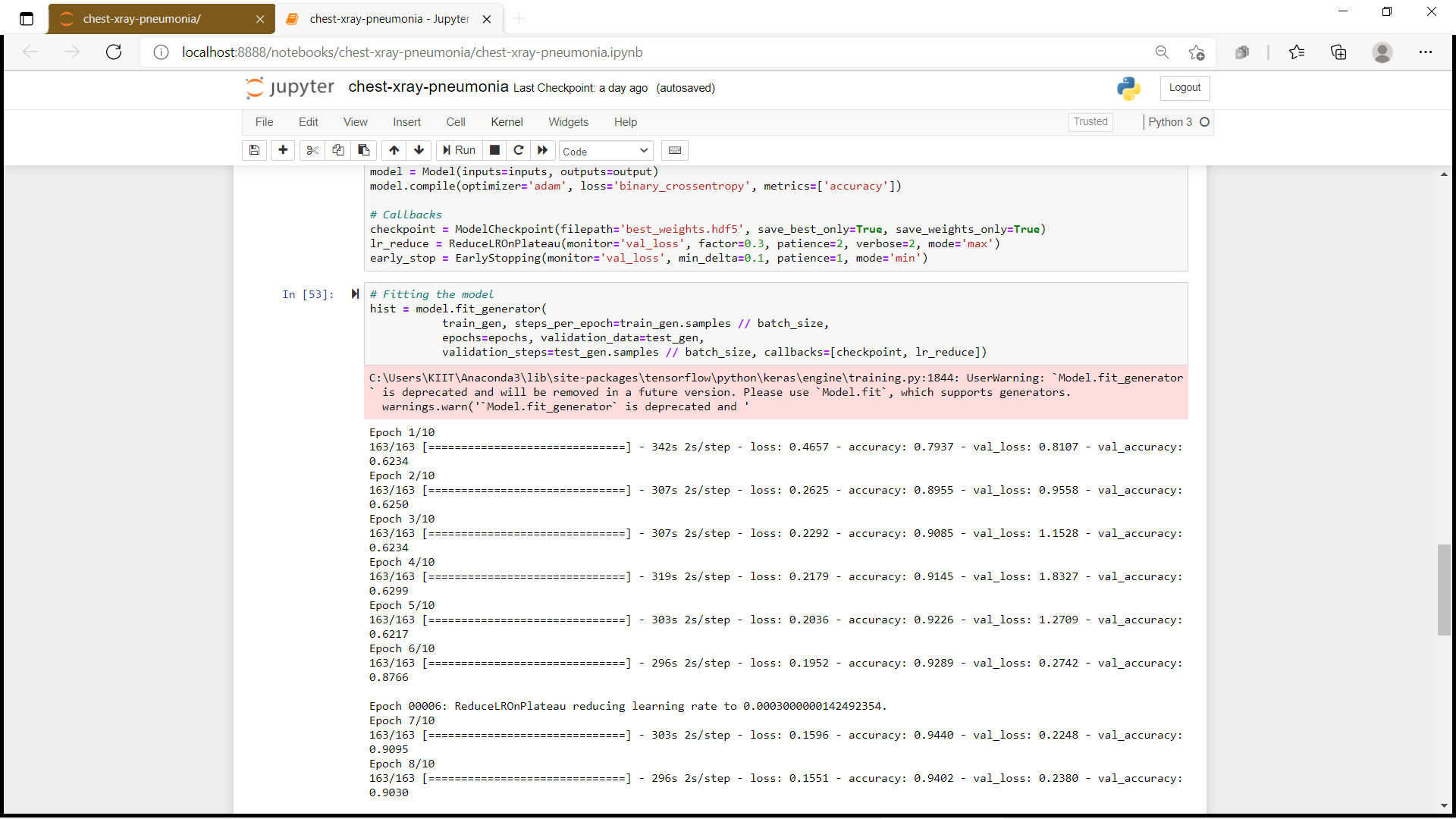
* ****ModelCheckpoint****: when training requires a lot of time to achieve a good result, often many iterations are required. In this case, it is better to save a copy of the best performing model only when an epoch that improves the metrics ends.
* ****EarlyStopping****: sometimes, during training we can notice that the generalization gap (i.e. the difference between training and validation error) starts to increase, instead of decreasing. This is a symptom of overfitting that can be solved in many ways (*reducing model capacity*, *increasing training data*, *data augumentation*, *regularization*, *dropout*, etc). Often a practical and efficient solution is to stop training when the generalization gap is getting worse.



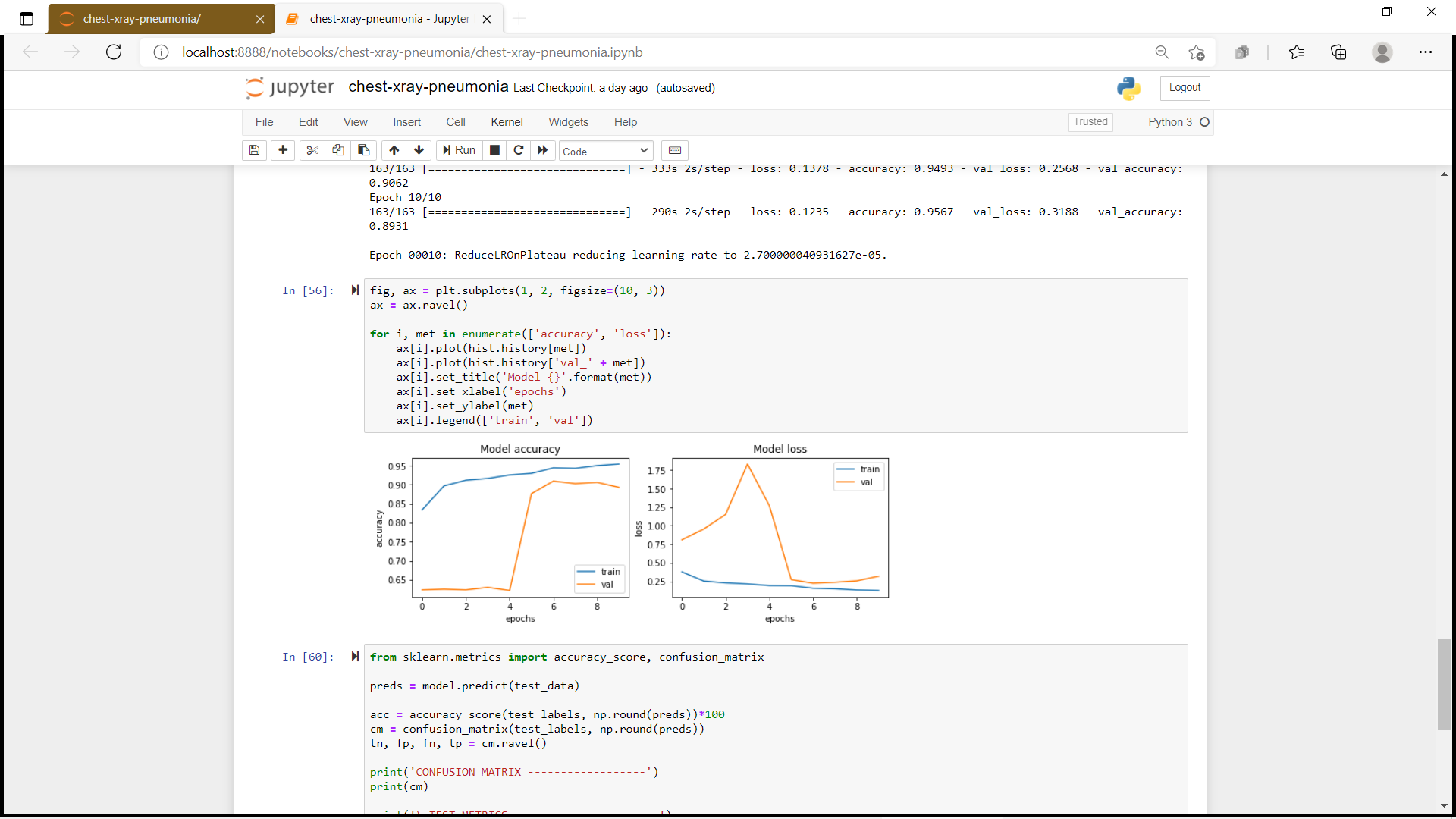




Next I trained the model for 10 epochs with a batch size of 32. Please note that usually a higher batch size gives better results but at the expense of higher computational burden. Some research also claim that there is an optimal batch size for best results which could be found by investing some time on hyper-parameter tuning.

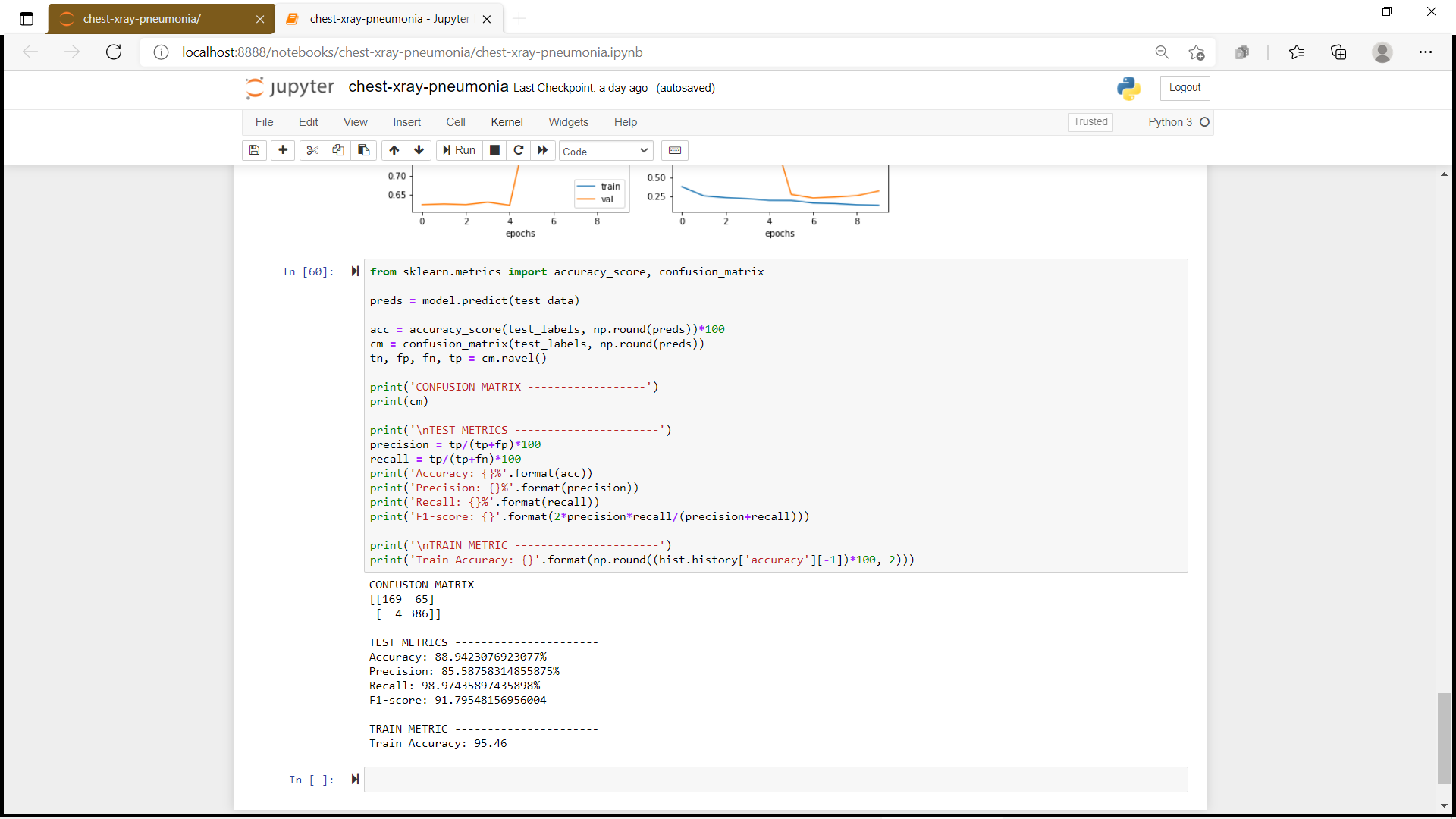


Let’s visualize the loss and accuracy plots.



So far so good. The model is converging which can be observed from the decrease in loss and validation loss with epochs. Also it is able to reach 90% validation accuracy in just 10 epochs.

Let’s plot the confusion matrix and get some of the other results also like precision, recall, F1 score and accuracy.



The model is able to achieve an accuracy of 91.02% which is quite good considering the size of data that is used.

Although this project is far from complete but it is remarkable to see the success of deep learning in such varied real world problems. I have demonstrated how to classify positive and negative pneumonia data from a collection of X-ray images. The model was made from scratch, which separates it from other methods that rely heavily on transfer learning approach. In the future this work could be extended to detect and classify X-ray images consisting of lung cancer and pneumonia. Distinguishing X-ray images that contain lung cancer and pneumonia has been a big issue in recent times, and our next approach should be to tackle this problem.

**Future Work:**

Below are a variety of techniques and ideas one might pursue to expand on the work of this project.

1. ***Different Features:***

One could try running PCA on the images to decorrelate the data and use these results on a logistic regression model. (In other words, use a subset of principal components as features.) It may also be worth exploring different image feature extraction techniques to get a better baseline.

1. ***Improved Logistic Regression Baseline:***

Although we were able to run logistic regression on 32x32 and 128x128 down-sampled images, we were unable to both do a hyperparameter sweep on images larger than 32x32 and use the same image size as our deep learning model, 224x224. With more time, we would implement a Spark cluster on Google Cloud and attempt to fifind the best regularization value for 224x224 images. We predict that our baseline would improve fairly signifificantly as a result.

***C. More Error Analysis:***

One might explore the DenseNet’s activations using Class Activation Maps in order to better understand its behavior on this dataset. This might inform why it achieves such low AUC as compared to similar work.

***D. Bounding Boxes:***

A small subset of the dataset has bounding boxes around diseased areas. One could use these bounding boxes to train a CNN to not only classify images with pneumonia, but also identify where in the image the pneumonia is located.

**CONCLUSION**

**6.1 SUMMARY**

With high rates of patients suffering from pneumonia we need more accurate pneumonia detection system with minimal losses.So, in our project we developed a model to detect and pneumonia from chest X-ray images taken at high validation accuracy. The algorithm begins classification of images by the convolutional neural network framework, which extracts features from the images and classifies them. We developed two models one without batch normalization and dropout and one with batch normalization and dropout. .To validate the performance of the trained model on different chest X-ray images, we varied the sizes of the training still obtained relatively similar results. Because of the effectiveness of our trained CNN model for identifying pneumonia from chest X-ray images, the validation accuracy of our model was significantly higher with minimal losses .

Although this project is far from complete but it is remarkable to see the success of deep learning in such varied real world problems. I have demonstrated how to classify positive and negative pneumonia data from a collection of X-ray images. The model was made from scratch, which separates it from other methods that rely heavily on transfer learning approach. In the future this work could be extended to detect and classify X-ray images consisting of lung cancer and pneumonia. Distinguishing X-ray images that contain lung cancer and pneumonia has been a big issue in recent times, and our next approach should be to tackle this problem.

**6.2 COST ANALYSIS**

Since the project work is based on python programming , there is no actual procurement cost required.

**REFERENCES**

1. Rui P, Kang K. National Ambulatory Medical Care Survey: 2015 Emergency Department Summary Tables. Table 27. Available from: www.cdc.gov/nchs/data/nhamcs/webtables/2015edwebtables.pdf

2. Deaths: Final Data for 2015. Supplemental Tables. Tables I-21, I-22. Available from: www.cdc.gov/nchs/data/nvsr/nvsr66/nvsr6606tables.pdf

3. Franquet T. Imaging of community-acquired pneumonia. J Thorac Imaging 2018 (epub ahead of print). PMID 30036297

4. Kelly B. The Chest Radiograph. Ulster Med J 2012;81(3):143-148

5. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017, http://openaccess.thecvf.com/content\_cvpr\_2017/papers/Wang\_ChestX-ray8\_Hospital-Scale\_Chest\_CVPR\_2017\_paper.pdf

6. https://www.medicalnewstoday.com/articles/151632

7. https://ieeexplore.ieee.org/document/8869364

8. <https://www.sas.com/en_in/insights/analytics/machine> learning.html#:~:text=Machine%20learning%20is%20a%20method,decisions%20with%20minimal%20human%20intervention.

9. <https://www.geeksforgeeks.org/machine-learning/>

10. https://towardsdatascience.com/pneumonia-detection-using-deep-learning-5dc66468eb9