

***Table of Contents:***

**1.1-Introduction 03**

**1.2 Objectives of Research 05**

**1.3 Problem Statement 06**

**2. Review of literature 07**

**3. Data Collection 09**

**4. Methodology 10**

**4.1 Exploratory Data Analysis**

**4.1.1 Figures and Tables**

**4.2 Data Modelling 13**

**5. Findings and Suggestions 16**

**6. Conclusion 17**

***Introduction***

***Python:***

Python is a general-purpose programming language which can be used for a wide variety of applications. A great language for beginners because of its readability and other structural elements designed to make it easy to understand, Python is not limited to basic usage. In fact, it powers some of the world's most complex applications and website.

Python is an interpreted language, meaning that programs written in Python don't need to be compiled in advance in order to run, making it easy to test small snippets of code and making code written in Python easier to move between platforms.

Python has an enormous user community. This means that no matter what problem you're trying to solve, chances are there is already strong documentation, tutorials, guides, and examples to help you along your way. There are numerous integrated development environments and other development tools to choose from, and thousands of open source packages available to extend Python to do just about anything you can think of.

Python's ease of use and compatibility across a variety of operating systems makes it an ideal language for a number of uses. Many complex websites either currently or historically have used Python to power their back ends, from YouTube to Instagram to Reddit, and thousands of other well-known examples. But Python isn't only a web language.

Python is the primary language used for the massive cloud computing project OpenStack, powering private and public clouds in data centers all over the world. It's also used to write desktop software, like Calibre, OpenShot, and the original client for BitTorrent. Many application written in other languages, such as Blender, allow for scripting by users in Python. It's also a popular language for machine learning, scientific, statistical, mathematical, and other types of specialized computing.

Python is-

**1) Readable and Maintainable Code**

**2) Multiple Programming Paradigms**

**3) Compatible with Major Platforms and Systems**

**4) Robust Standard Library**

**5) Many Open Source Frameworks and Tools**

**6) Simplify Complex Software Development**

**7) Adopt Test Driven Development**

**Artificial Intelligence:**

Artificial Intelligence is a way of making a computer, a computer-controlled robot, or a software think intelligently, in the similar manner the intelligent humans think.

AI is accomplished by studying how human brain thinks, and how humans learn, decide, and work while trying to solve a problem, and then using the outcomes of this study as a basis of developing intelligent software and systems.

The goals of AI are to create expert systems and to implement human intelligence in machines, in which the machine exhibits intelligent behavior, learn, understand, think and advice users and also behave like humans.

AI has been playing a dominant role in various fields such as-

1. Gaming
2. Nature Language Processing
3. Expert Systems
4. Vision Systems
5. Speech Recognition
6. Handwriting Recognition
7. Intelligent Robots

AI algorithms are capable of learning from data; they can enhance themselves by learning new heuristics (strategies, or "rules of thumb", that have worked well in the past), or can themselves write other algorithms.

***Objectives of Research***

The main objective of this research is to achieve solution to prevent the deaths caused due to improper generation of reports of patients suffering from pneumonia.

This study aimed to develop a predictive model for pneumonia hospitalization to accurately identify high-risk individuals to facilitate the efficient prediction of pneumonia.

***PROBLEM STATEMENT:***

**To predict the patient’s condition of pneumonia using chest x-ray images.**

**What actually is pneumonia?**

Pneumonia is an infection of the lungs with a range of possible causes. It can be a serious and life-threatening disease.

Pneumonia can occur in young and healthy people, but it is most dangerous for older adults, infants, people with other diseases, and those with impaired immune systems.

Sometimes pneumonia is difficult to diagnose because its symptoms are commonly seen in people suffering with colds or the flu. We may not diagnose until it lasts longer than these other conditions.

*A chest x-ray will confirm whether the patient is suffering from pneumonia or not. A doctor can only evaluate x-rays up to certain limit and some evaluations may not be accurate.*

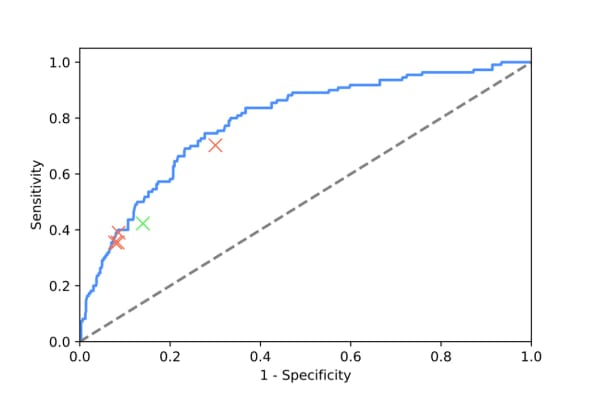
**Solution to the problem:**

Developing an application which uses the chest x-ray of the patient for evaluating the person’s condition of pneumonia.

***Review of Literature***

Artificial intelligence (AI) aims to mimic human cognitive functions. It is bringing a paradigm shift to healthcare, powered by increasing availability of healthcare data and rapid progress of analytics techniques. We survey the current status of AI applications in healthcare and discuss its future. AI can be applied to various types of healthcare data (structured and unstructured). Popular AI techniques include machine learning methods for structured data, such as the classical support vector machine and neural network, and the modern deep learning, as well as natural language processing for unstructured data. Major disease areas that use AI tools include cancer, neurology and cardiology. We then review in more details the AI applications in stroke, in the three major areas of early detection and diagnosis, treatment, as well as outcome prediction and prognosis evaluation.

The research team led by Stanford adjunct professor Andrew Ng, started by training the neural network with 112,120 chest X-ray images that were previously manually labeled with up to 14 different diseases. One of them was pneumonia. After training it for a month, the software beat previous computer-based methods to detect this type of infection. The Stanford Machine Learning Group team pitted its software against four Stanford radiologists, giving each of them 420 X-ray images



Many deep learning models based on CNN algorithm have been developed to detect pneumonia in our lungs. The accuracies of these models typically vary from 75% to 90%.

There have been various attempts to improve upon these accuracy figures in early 2000s but they couldn't be achieved because of hardware limitations

But today with accelerated GPUs we can achieve accuracies higher than 95%.

But to clock an accuracy higher than 95% we need more epoch for which high end workstations or supercomputers are being used.

Nevertheless, AI and deep learning in medical sector has been pretty disruptive and this trend of using Artificial Intelligence in this field will continue for the welfare of mankind.

Importantly, it has a simple user experience: Doctors input an X-ray of lungs, and automatically get the numeric probability of those lungs being infected with pneumonia or not. They also get a color map that highlights the level of infection throughout the tissue. Using these insights, the doctors can then make decisions about how to approach treatment.

The team hopes to see its research applied not only in hospitals but all around the world

***DATA COLLECTION:***

To predict whether the patient is suffering from pneumonia or not, the model is to be trained with a set of x-ray images which contains chest x-rays of normal person and chest x-rays of person suffering from Pneumonia.

The model is to be trained because, these x-rays are used as reference and the uploaded x-ray of the patient is compared to the existing sets, there by predicting the result of the patient.

The source link through which the data is collected is mentioned below:

[https://medium.com/datadriveninvestor/detecting-pneumonia-in-chestx-rays-with-custom-vision-and-pytorch-e270e071e982](https://medium.com/datadriveninvestor/detecting-pneumonia-in-chest-x-rays-with-custom-vision-and-pytorch-e270e071e982)

In the link we can find the set of data images of chest x-rays of normal person and chest x-rays of person suffering from Pneumonia.

The data set can be downloaded directly and can be used to train the model using those set of x-rays.

The process of training the data and working of the model will be explained in the further sections.

***Exploratory Data Analysis***

**Figures and Tables**

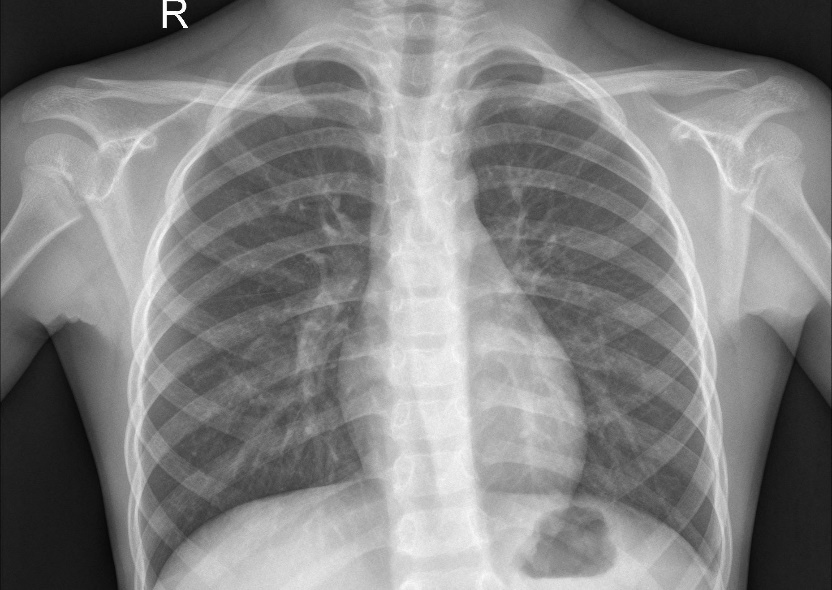
**CONFUSION MATRIX:**

A confusion matrix is a summary of prediction results on a classification problem.  
The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix.  
The confusion matrix shows the ways in which your classification model is confused when it makes predictions.  
It gives us insight not only into the errors being made by a classifier but more importantly the types of errors that are being made.



• Positive (P) : Observation is positive (for example: is an apple).  
• Negative (N) : Observation is not positive (for example: is not an apple).  
• True Positive (TP) : Observation is positive, and is predicted to be positive.  
• False Negative (FN) : Observation is positive, but is predicted negative.  
• True Negative (TN) : Observation is negative, and is predicted to be negative.  
• False Positive (FP) : Observation is negative, but is predicted positive.

# In this application we work with x-ray images to evaluate the patient’s condition on pneumonia.



The above x-ray image is the chest x-ray of normal person.

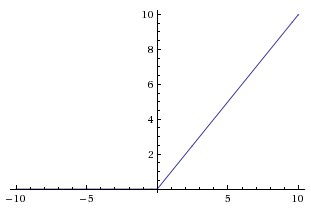


The above x-ray image is the chest x-ray of the person suffering from pneumonia.

The difference in the two images is the person’s x-ray who is suffering from pneumonia consist of white spots in the lung’s area of the x-ray image this will decide whether the person is suffering from pneumonia or not.

**RELU GRAPH:**

The Rectified Linear Unit is the most commonly used activation function in deep learning models



***Data Modelling***

The life-threatening disease Pneumonia can be predicted using this model. In this model we have used Convolution Neural Network algorithm (CNN), which is an AI algorithm.

What is a CNN?

A convolutional neural network (CNN) is a class of deep neural networks, most commonly applied to analysing visual imagery.

CNNs are regularized versions of multilayer perceptron’s. Multilayer perceptron’s usually refer to fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer.

A CNN works similar to the human eye.

It has a pixel range of 0 to 255.

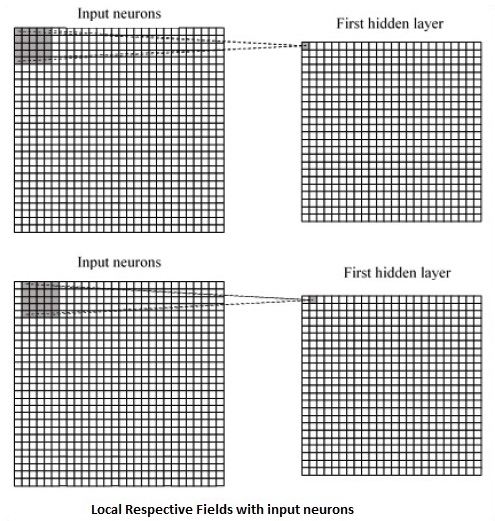
A CNN can be used in image classification.

Convolutional Neural networks are designed to process data through multiple layers of arrays. This type of neural networks is used in applications like image recognition or face recognition. The primary difference between CNN and any other ordinary neural network is that CNN takes input as a two-dimensional array and operates directly on the images rather than focusing on feature extraction which other neural networks focus on.

CNN utilizes spatial correlations that exist within the input data. Each concurrent layer of a neural network connects some input neurons. This specific region is called local receptive field. Local receptive field focusses on the hidden neurons. The hidden neurons process the input data inside the mentioned field not realizing the changes outside the specific boundary.

The mapping of connections from the input layer to the hidden feature map is defined as “shared weights” and bias included is called “shared bias”.

CNN or convolutional neural networks use pooling layers, which are the layers, positioned immediately after CNN declaration. It takes the input from the user as a feature map that comes out of convolutional networks and prepares a condensed feature map. Pooling layers helps in creating layers with neurons of previous layers.



Implementation of CNN using TensorFlow**:**

The following are steps that require the execution and proper dimension of the entire network:

**Step-1-** Include the necessary modules for TensorFlow and the data set modules, which are needed to compute the CNN model.

**Step-2-**  Declare a function, which includes various parameters and optimization variables with declaration of data placeholders. These optimization variables will declare the training pattern.

**Step-3-**  In this step, we will declare the training data placeholders with input parameters - for 64 x 64 pixels. This is the flattened image data.

We can reshape the tensor according to our requirements. The first value tells function to dynamically shape that dimension based on the amount of data passed to it. The two middle dimensions are set to the image size (i.e. 64 x 64).

**Step-4-**  Now it is important to create some convolutional layers

**Step-5-**  Let us flatten the output ready for the fully connected output stage - after two layers of stride 2 pooling with the dimensions. To create the fully connected with "dense" layer. We can set up some weights and bias values for this layer, then activate with ReLU.

**Step-6-**  Another layer with specific SoftMax activations with the required optimizer defines the accuracy assessment, which makes the setup of initialization operator.

**Step-7-**  We should set up recording variables. This adds up a summary to store the accuracy of data.

***Findings and Reviews***

We have gone through many resource findings on Pneumonia and its predictions using Deep Learning Algorithm.

Here are some of the source links,

1. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia#chest_xray.zip>
2. <https://medium.com/datadriveninvestor/detecting-pneumonia-in-chest-x-rays-with-custom-vision-and-pytorch-e270e071e982>
3. <https://www.youtube.com/watch?v=JrWHyqonGj8>

Our model only predicts the existence of Pneumonia, but it doesn’t give any information about the extent and severity of Pneumonia.

Further improvisations of this model can include the detailing of the state/level of the Pneumonia and also prescribed medications to be followed based on the patient’s previous health records.

Suggestions:

We had hypothesized that the inclusion of detailed risk profiles in addition to age and comorbidities would enable more accurate predictions of future pneumonia events.

***Conclusion***

This study developed and internally validated a model to predict pneumonia hospitalization based on x-ray images. We had hypothesized that the inclusion of detailed risk profiles in addition to age and comorbidities would enable more accurate predictions of future pneumonia events. Our findings support this hypothesis, and future improvements to the model may facilitate efficient prevention of pneumonia in an aging society.