X-RAY PNEUMONIA IMAGE CLASSIFICATION WITH DEEP LEARNING

Bao Tram Duong

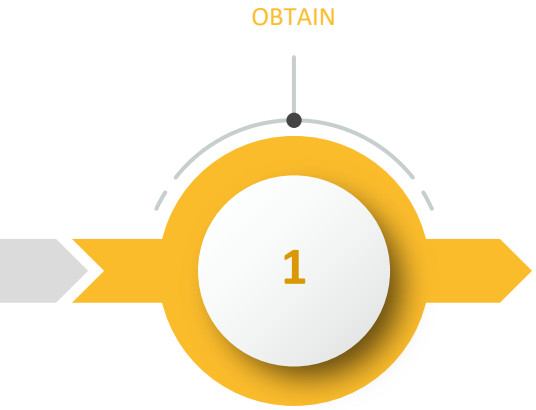
The Problem

- Globally, a child dies of pneumonia every 39 seconds.
- Pneumonia is a leading cause of morbidity and mortality in children younger than the age of 5 years, killing more children than HIV/AIDS, malaria, and measles combined.
- 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.

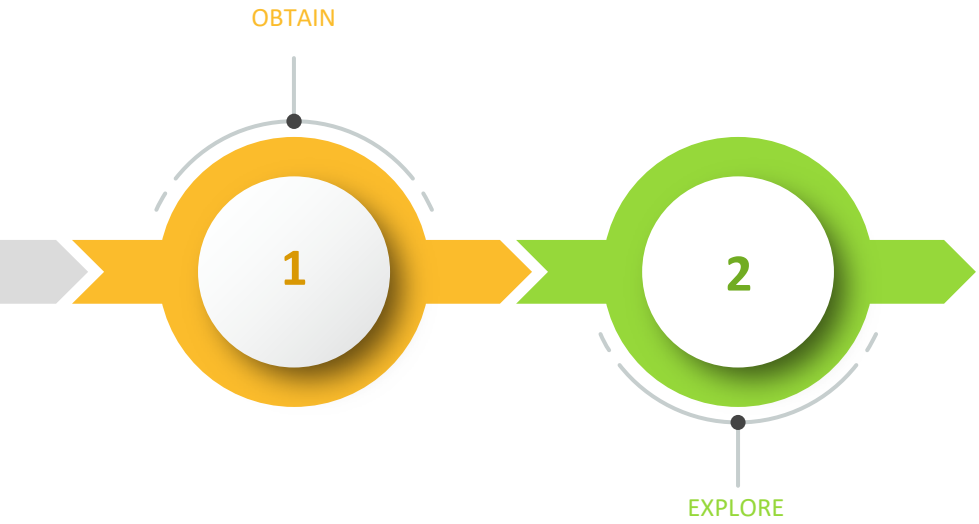
The Solution

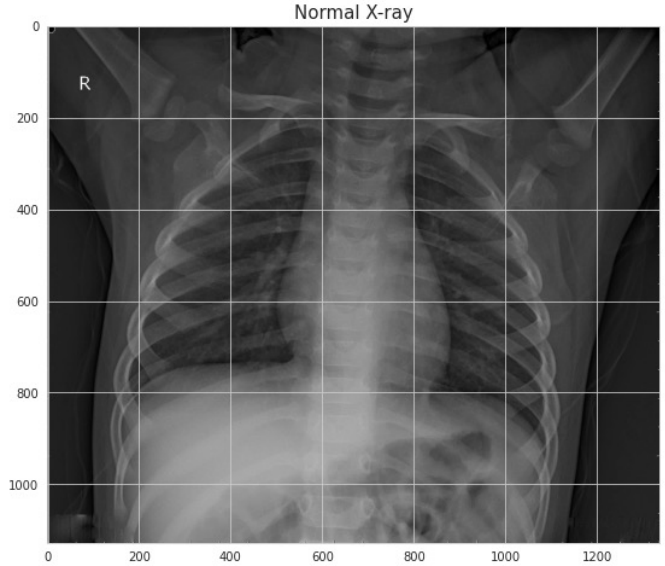
To solve this, deep learning (DL), a branch of machine learning (ML), inspired by the make-up of the human brain, are developed to detect hidden features in images which are not apparent or cannot be detected even by medical experts.

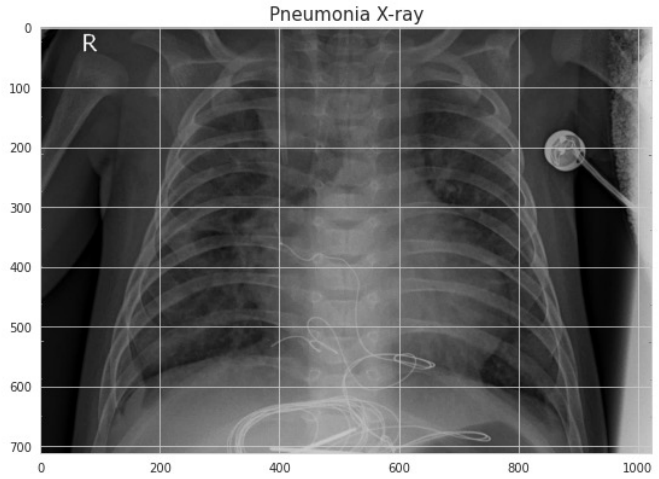
With AI system aiding medical experts in expediting the diagnosis, earlier treatment can be prescribed, resulting in improved clinical outcomes.



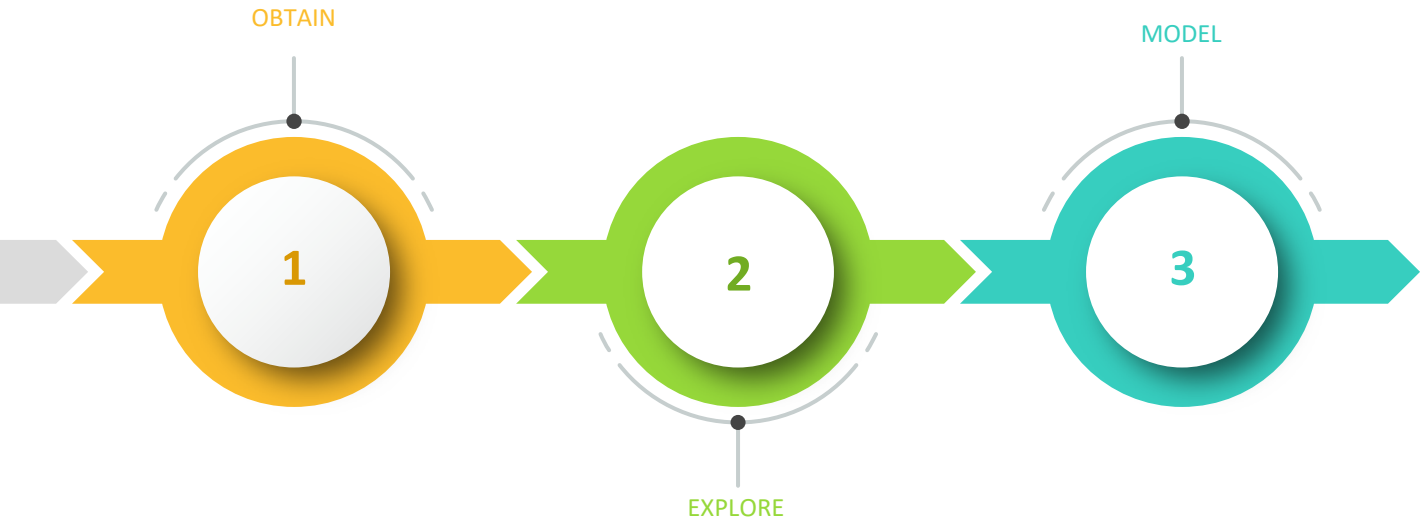
- Chest X-ray images (anterior-posterior) were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou. All chest X-ray imaging was performed as part of patients' routine clinical care.
- There are 5,863 X-Ray images (JPEG)
- 2 categories: Normal & Pneumonia

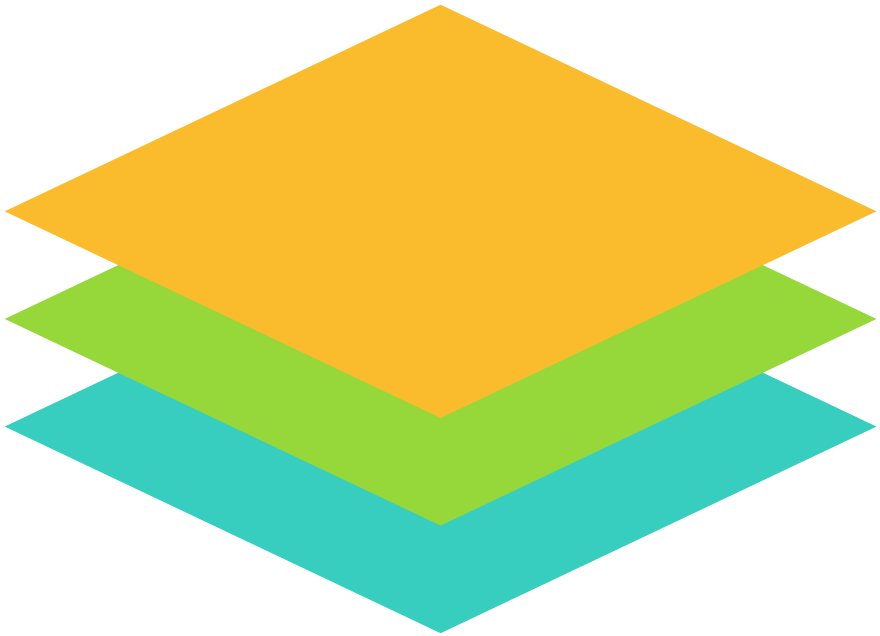

OBTAIN
1
2
EXPLORE


Normal X-ray
R
0
200
400
600
800
1000
0
200
400
600
800
1000
1200


Pneumonia X-ray
R
0
100
200
300
400
500
600
700
0
200
400
600
800
1000


X-ray Condition in Training Set
Count
4000
3500
3000
2500
2000
1500
1000
500
0
NORMAL
PNEUMONIA
Condition


OBTAIN
MODEL
1
2
3
EXPLORE

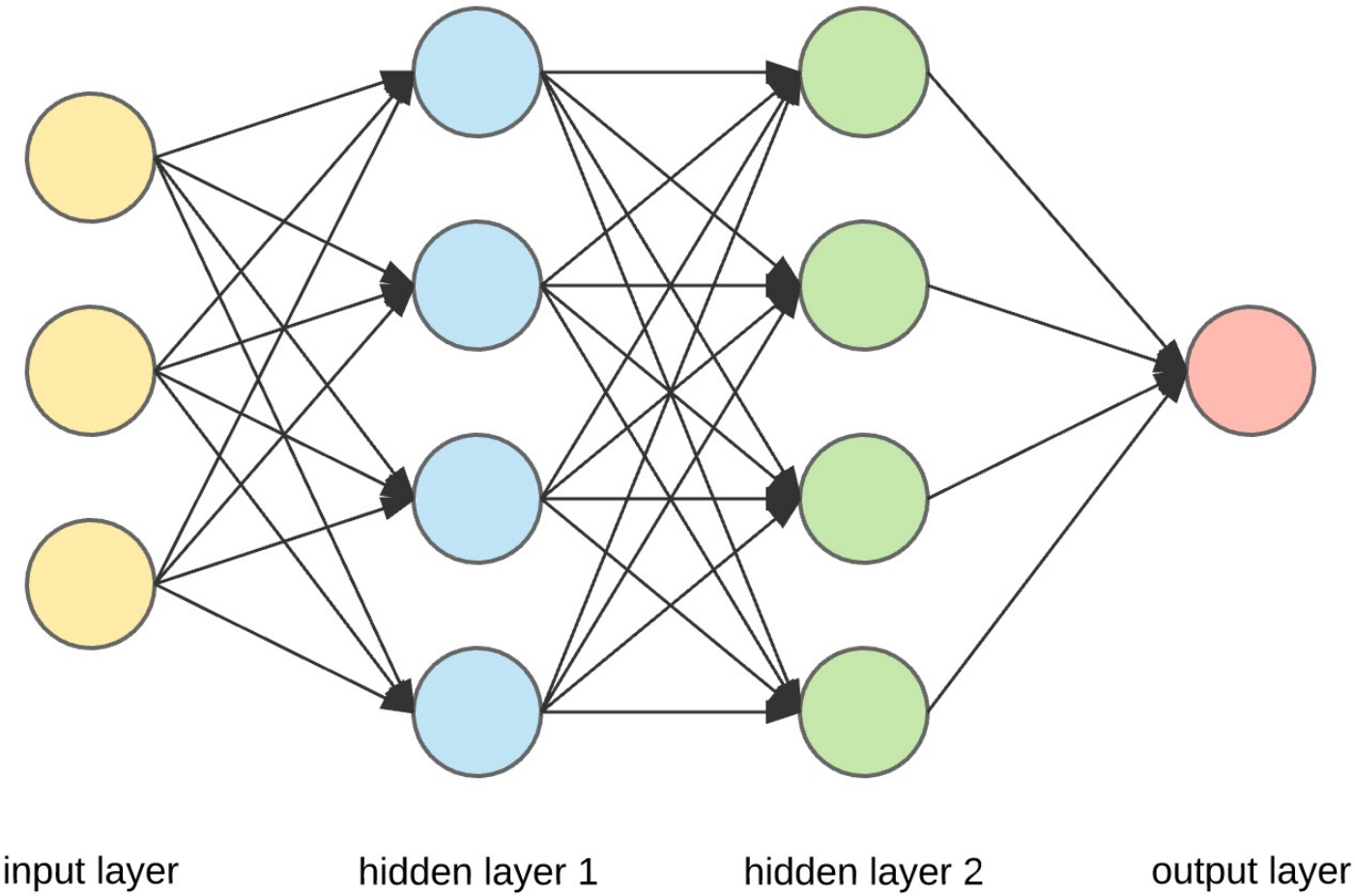


MULTILAYER PERCEPTRON MODEL

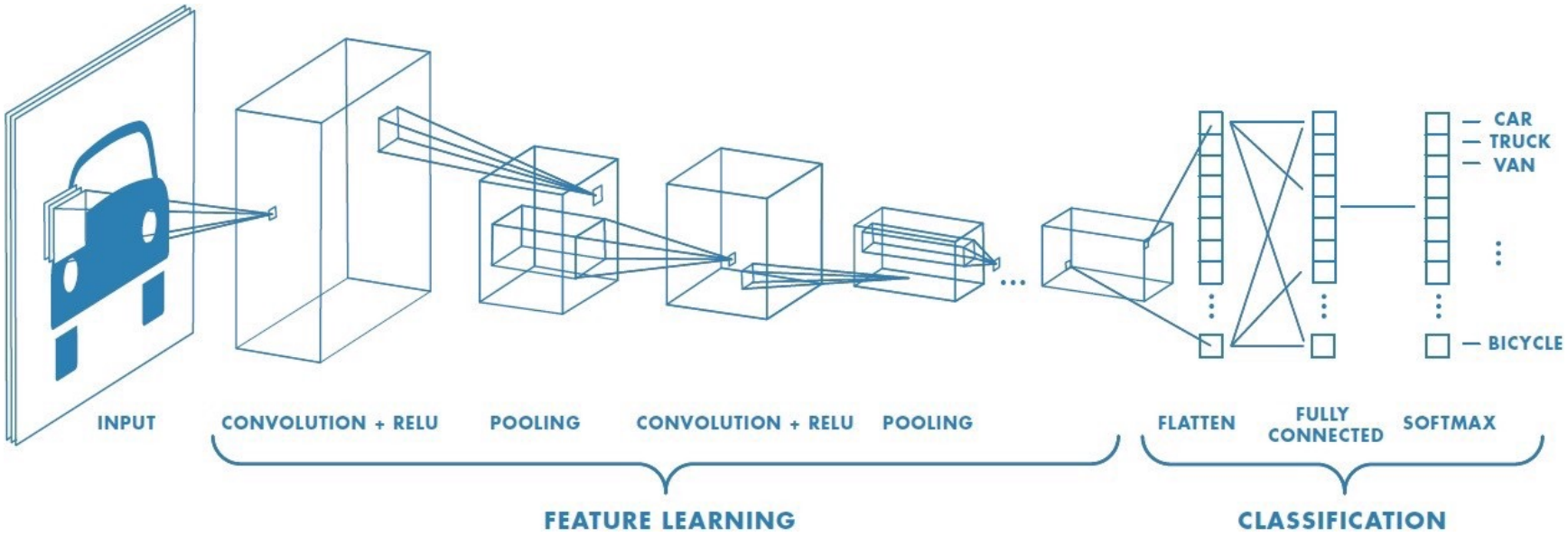
CONVOLUTIONAL NEURAL NETWORK MODEL

TRANSFER LEARNING VGG16 CNN MODEL

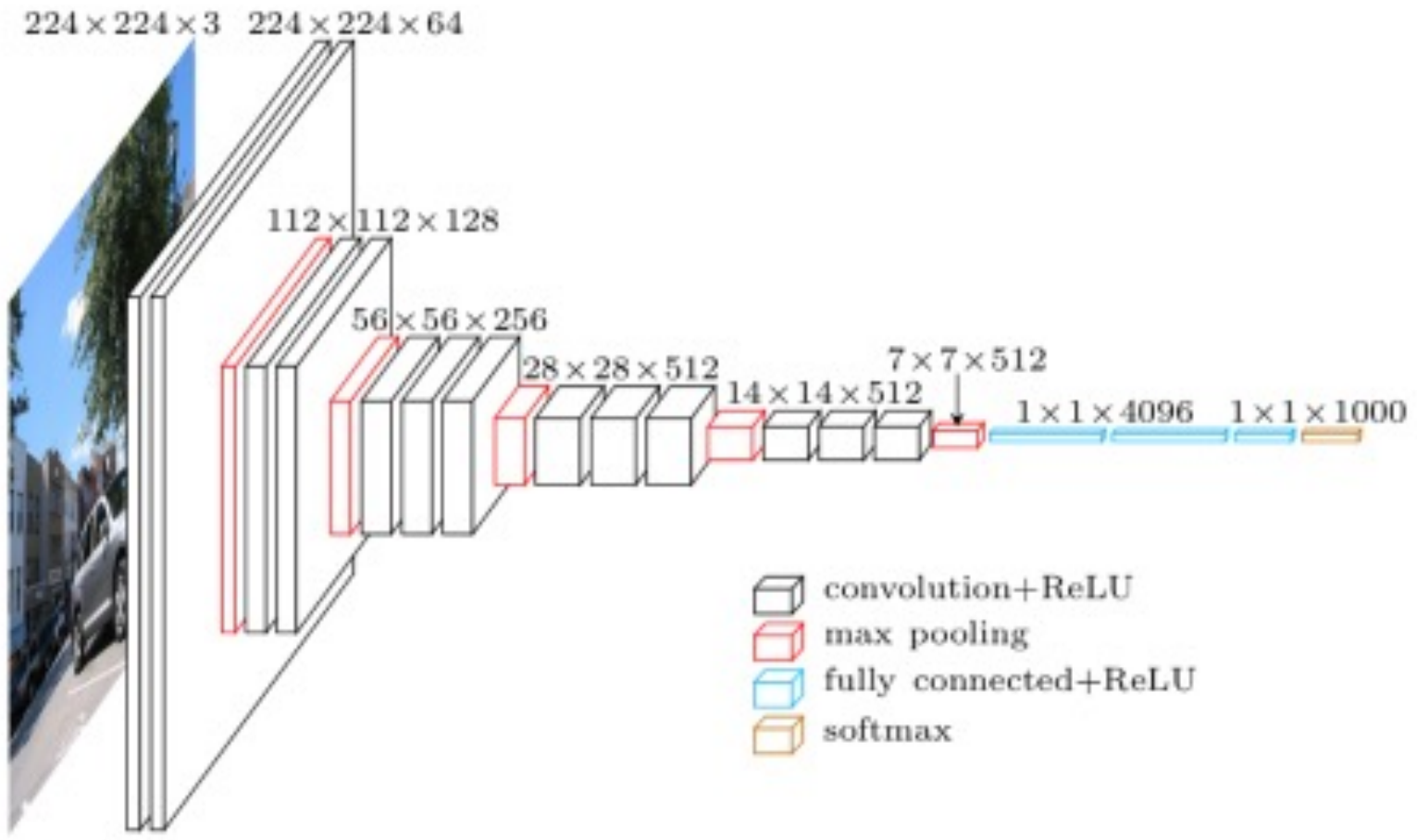
Multilayer Perceptron



Convolutional Neural Network

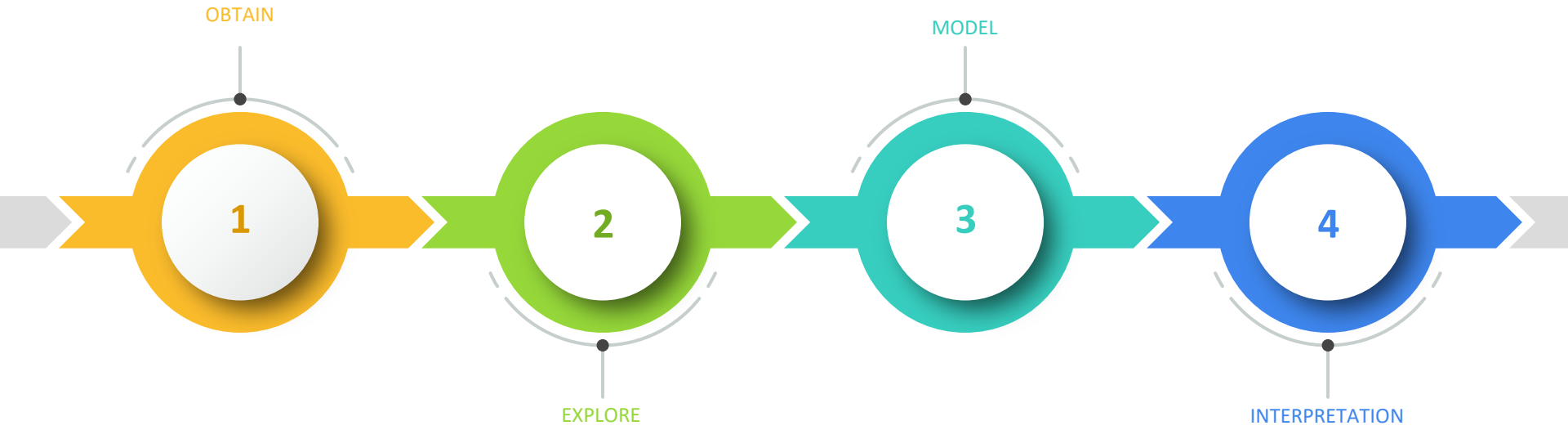


Transfer Learning with VGG16



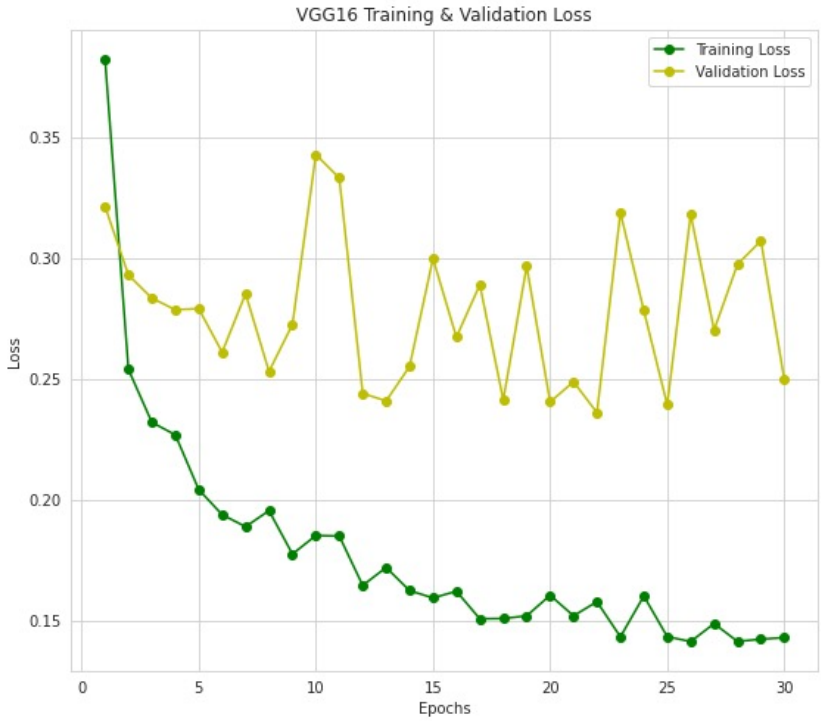
Summary of Key Findings

	Model	Accuracy	Precision	Recall	F1 Score	AUC
0	Multilayer Perceptron Model	0.74	0.83	0.66	0.66	0.66
1	Convolutional Neural Network Model	0.88	0.89	0.86	0.87	0.86
2	Transfer Learning: VGG16 CNN Model	0.92	0.92	0.90	0.91	0.90

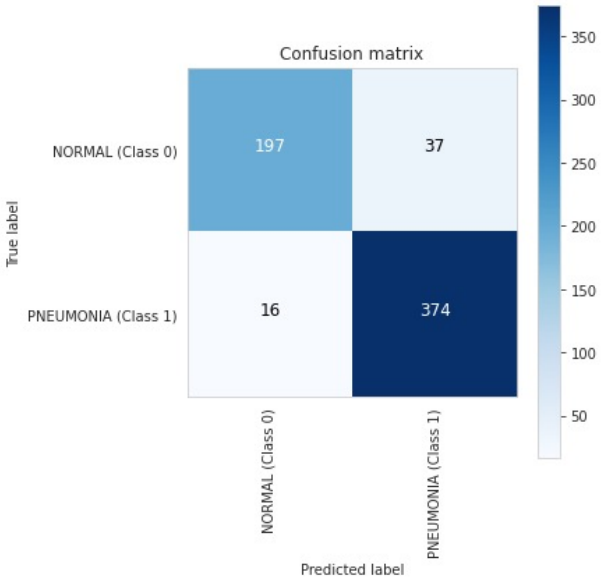

OBTAIN
1
EXPLORE
2
MODEL
3
INTERPRETATION
4

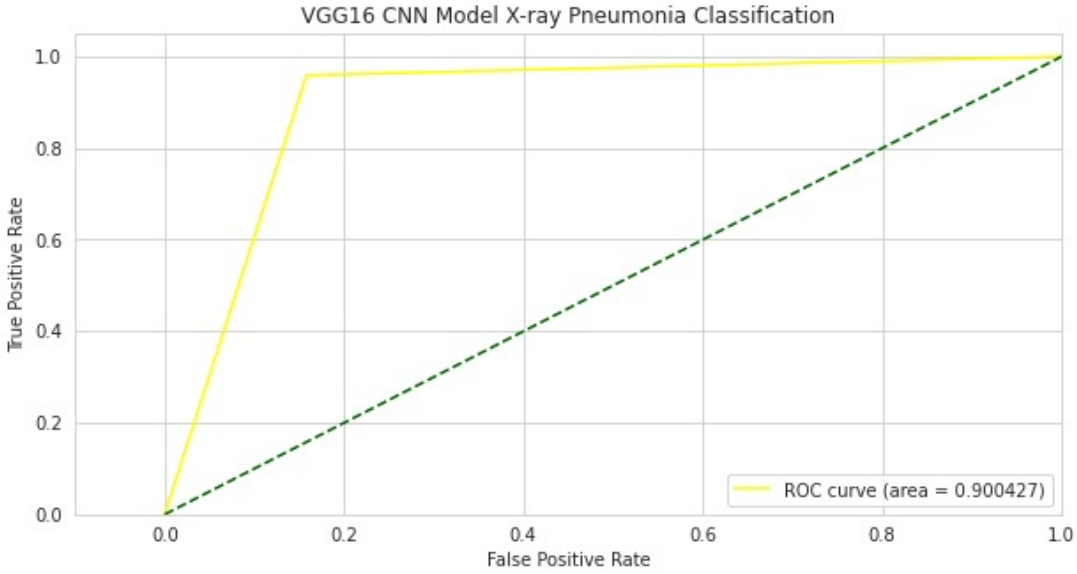
Best Model

Transfer Learning with VGG16



Train accuracy = 95%
Validation accuracy = 92%





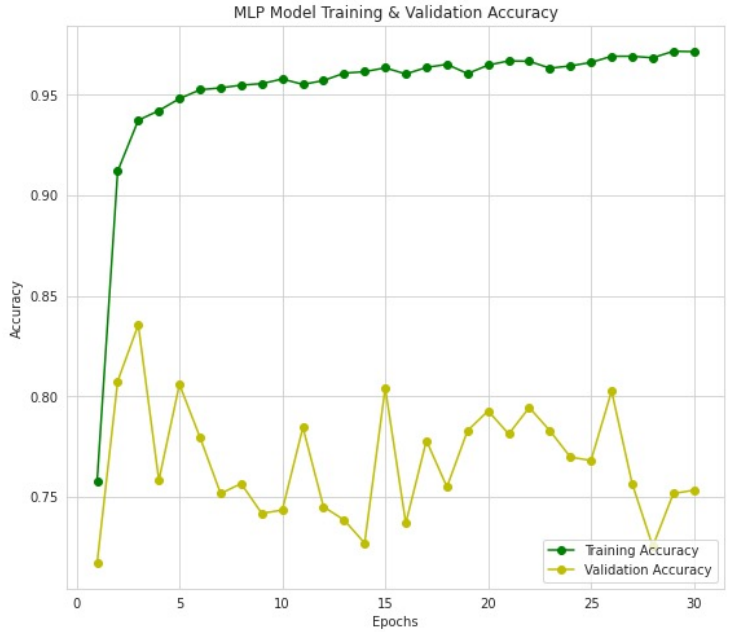
Accuracy of 92%
Recall/sensitivity of 90%
Precision/specificity of 92%
The area under the ROC curve of 90%

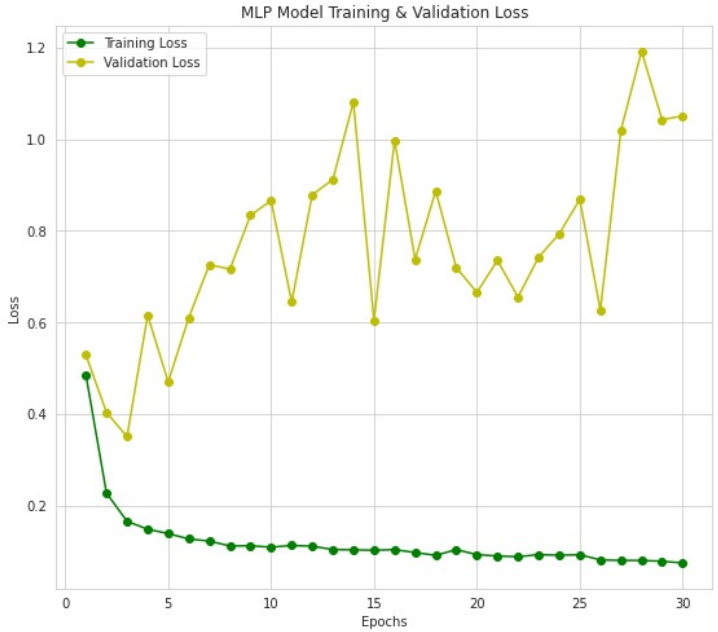
Future Work

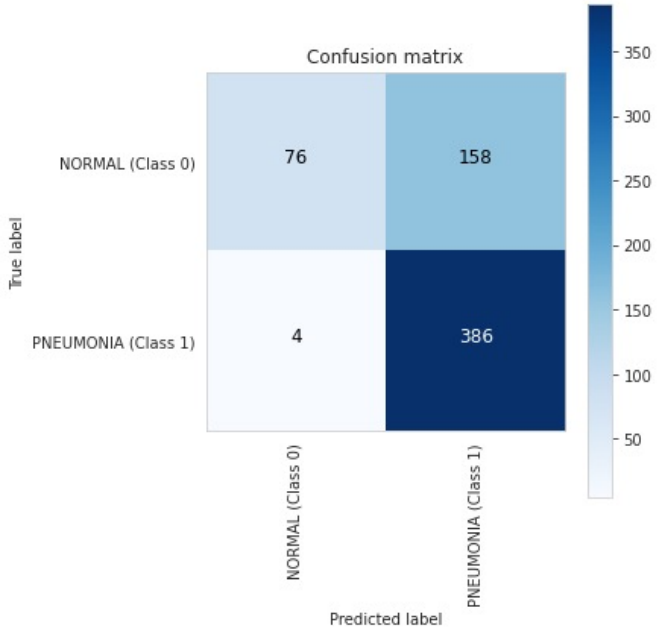
1. Build a multi-class classification deep learning model to distinguish between Normal, Viral Pneumonia, and Bacterial Pneumonia
2. Combine CNN models with other classifiers such as Support Vector Machine (SVM)
3. Tune parameters such as learning rate, batch size, try another optimizer, number of layers, different types of layer, number of neurons per layer, and the type of activation functions for each layer. GridSearchCV or RandomizedSearchSV can be used to achieve this.

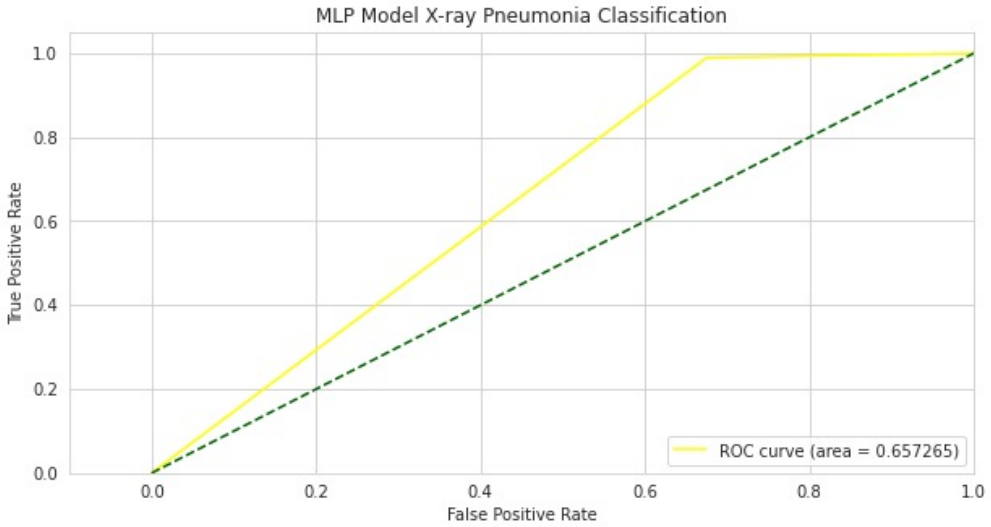
THANK YOU

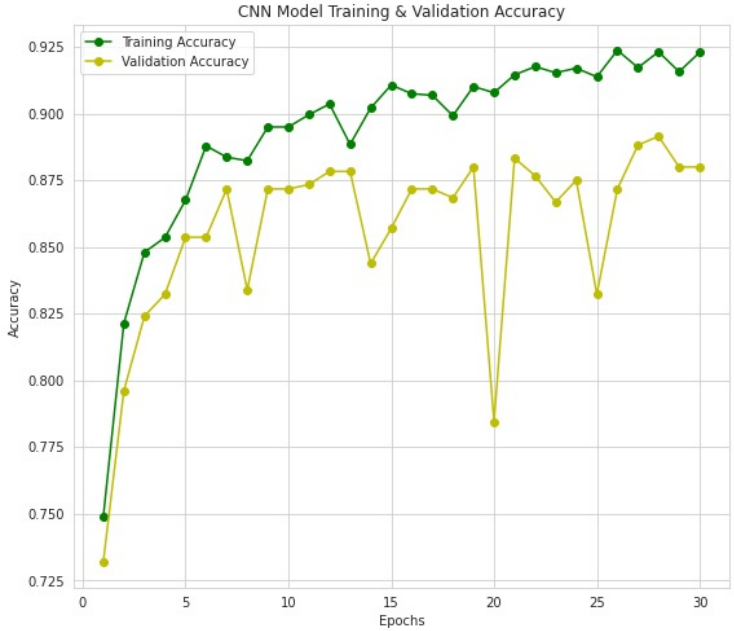
APPENDIX

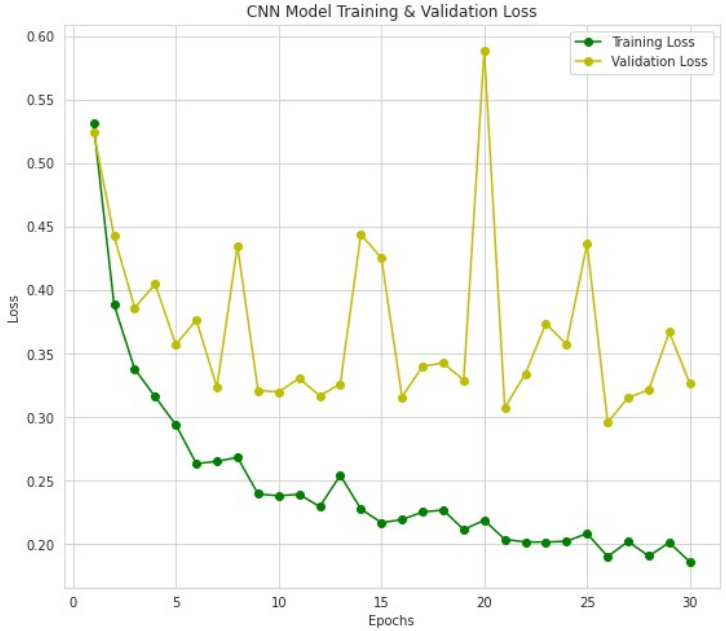

MLP Model Training & Validation Accuracy
Accuracy
Epochs
Training Accuracy
Validation Accuracy

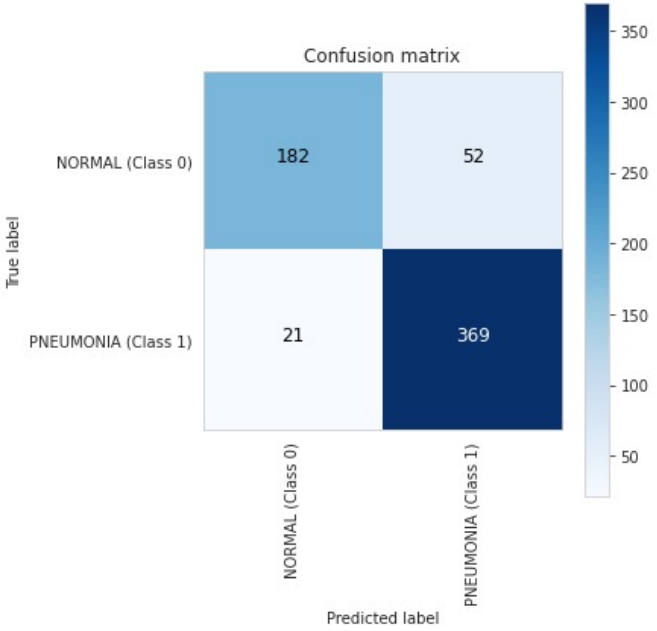

MLP Model Training & Validation Loss
Loss
Epochs
Training Loss
Validation Loss

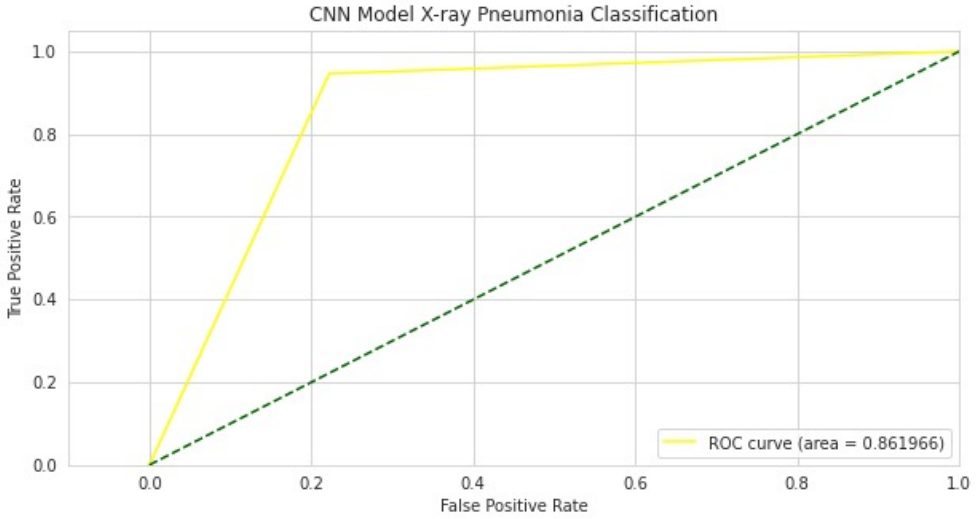

Confusion matrix
NORMAL (Class 0)
PNEUMONIA (Class 1)
True label
76
158
4
386
NORMAL (Class 0)
PNEUMONIA (Class 1)
Predicted label


MLP Model X-ray Pneumonia Classification
True Positive Rate
False Positive Rate
ROC curve (area = 0.657265)

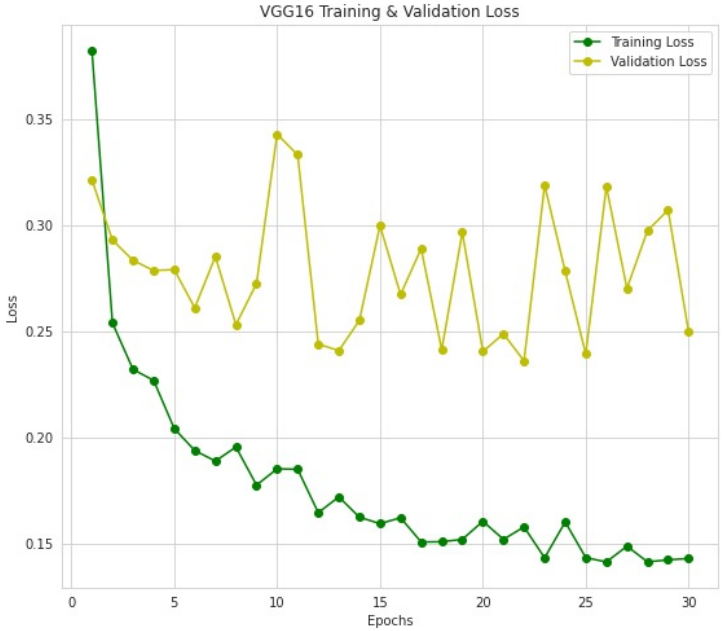

CNN Model Training & Validation Accuracy
Training Accuracy
Validation Accuracy
Accuracy
Epochs

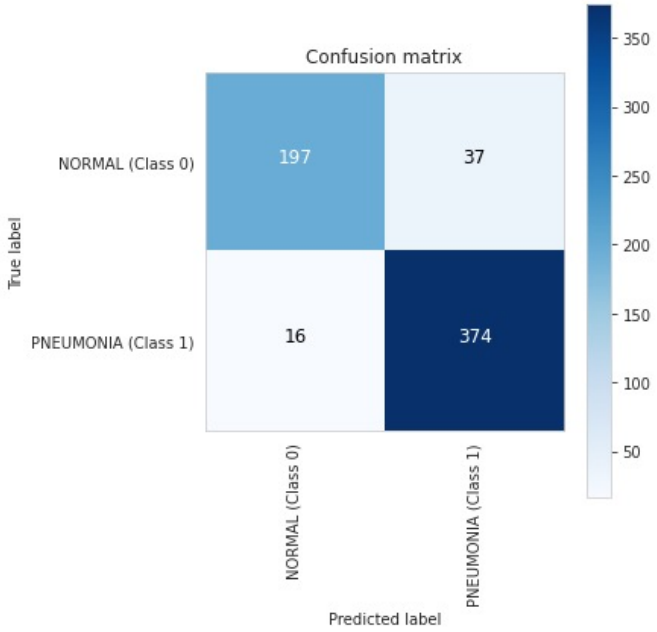

CNN Model Training & Validation Loss
Training Loss
Validation Loss
Loss
Epochs

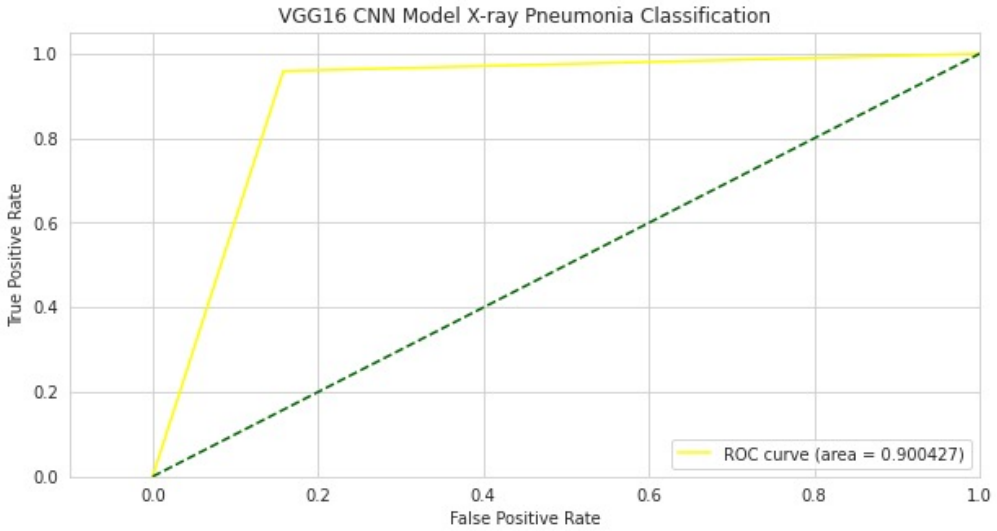

Confusion matrix
NORMAL (Class 0)
PNEUMONIA (Class 1)
182
52
21
369
True label
Predicted label


CNN Model X-ray Pneumonia Classification
True Positive Rate
False Positive Rate
ROC curve (area = 0.861966)


VGG16 Model Training & Validation Accuracy
Training Accuracy
Validation Accuracy
Accuracy
Epochs


VGG16 Training & Validation Loss
Training Loss
Validation Loss
Loss
Epochs


Confusion matrix
NORMAL (Class 0)
PNEUMONIA (Class 1)
197
37
16
374
True label
Predicted label


VGG16 CNN Model X-ray Pneumonia Classification
True Positive Rate
False Positive Rate
ROC curve (area = 0.900427)