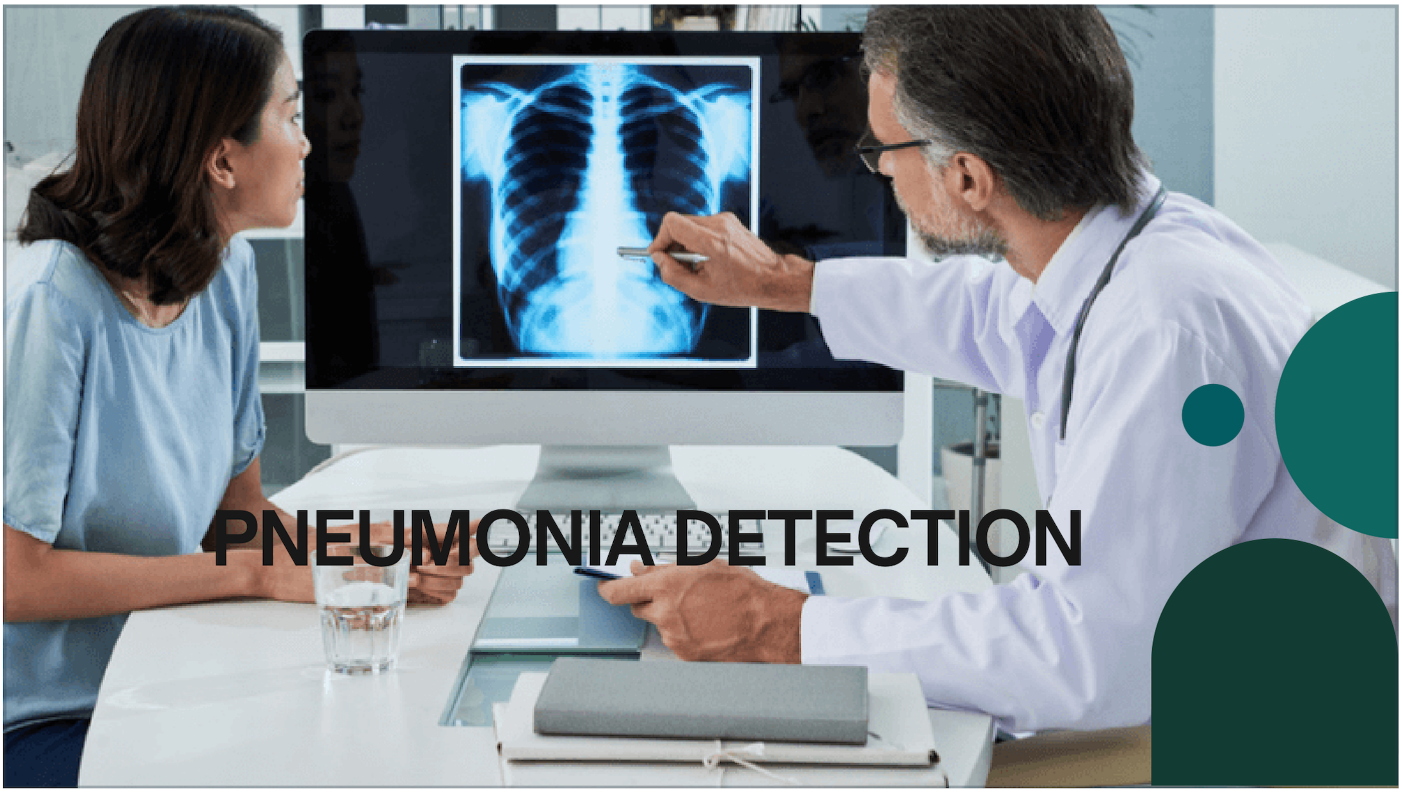


# PNEUMONIA DETECTION

# Project Overview

- The project focuses on leveraging a dataset of chest X-ray images to develop machine learning models capable of accurately diagnosing pneumonia. The dataset, sourced from Kaggle, comprises images classified into two categories: pneumonia and normal. This project aims to build and optimize convolutional neural network (CNN) models to achieve high diagnostic accuracy.

# TABLE OF CONTENT



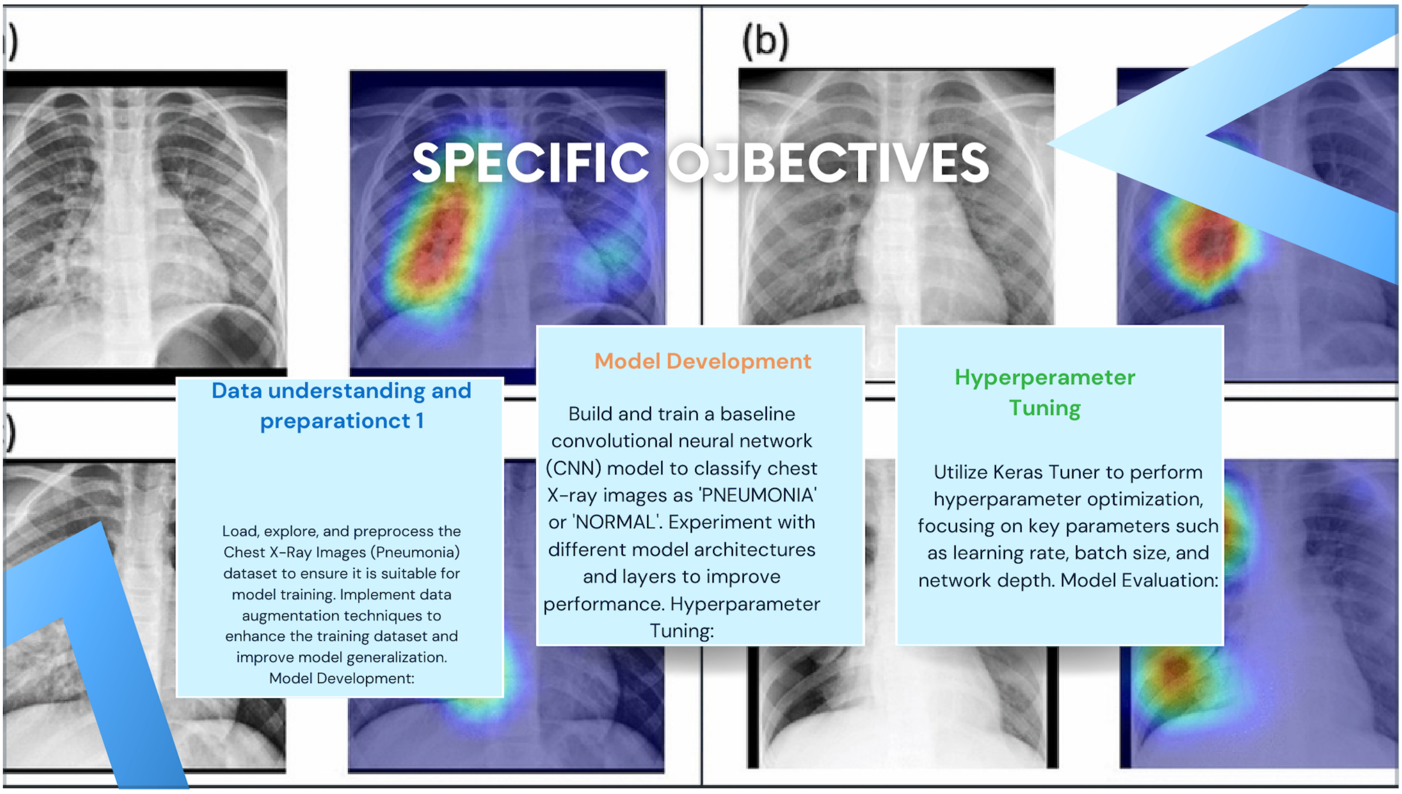
## GENERAL OBJECTIVE

The objective of this project is to leverage deep learning techniques to develop an automated system capable of accurately diagnosing pneumonia from chest X-ray images. This system aims to: Provide reliable diagnostic assistance to healthcare professionals. Reduce the diagnostic workload in hospitals and clinics. Improve diagnostic accuracy and speed, especially in areas with limited access to radiologists.

## PROBLEM STATEMENT

Pneumonia is a life-threatening infection of the lungs that results in millions of hospitalizations and deaths worldwide each year. Early and accurate diagnosis is critical for effective treatment and better patient outcomes. Chest X-ray imaging is the most commonly used diagnostic tool for pneumonia, but interpreting these images can be challenging and requires significant expertise. In many parts of the world, especially in resource-limited settings, there is a shortage of trained radiologists, which can delay diagnosis and treatment.


WORK DISTRIBUTION
The work was distributed into
Business Understanding
Data Exploration, Preview
Preprocessing
Model Archtecture
Model compiling and evaluation
General results evaluation




# SPECIFIC OJBECTIVES

### Data understanding and preparationct 1

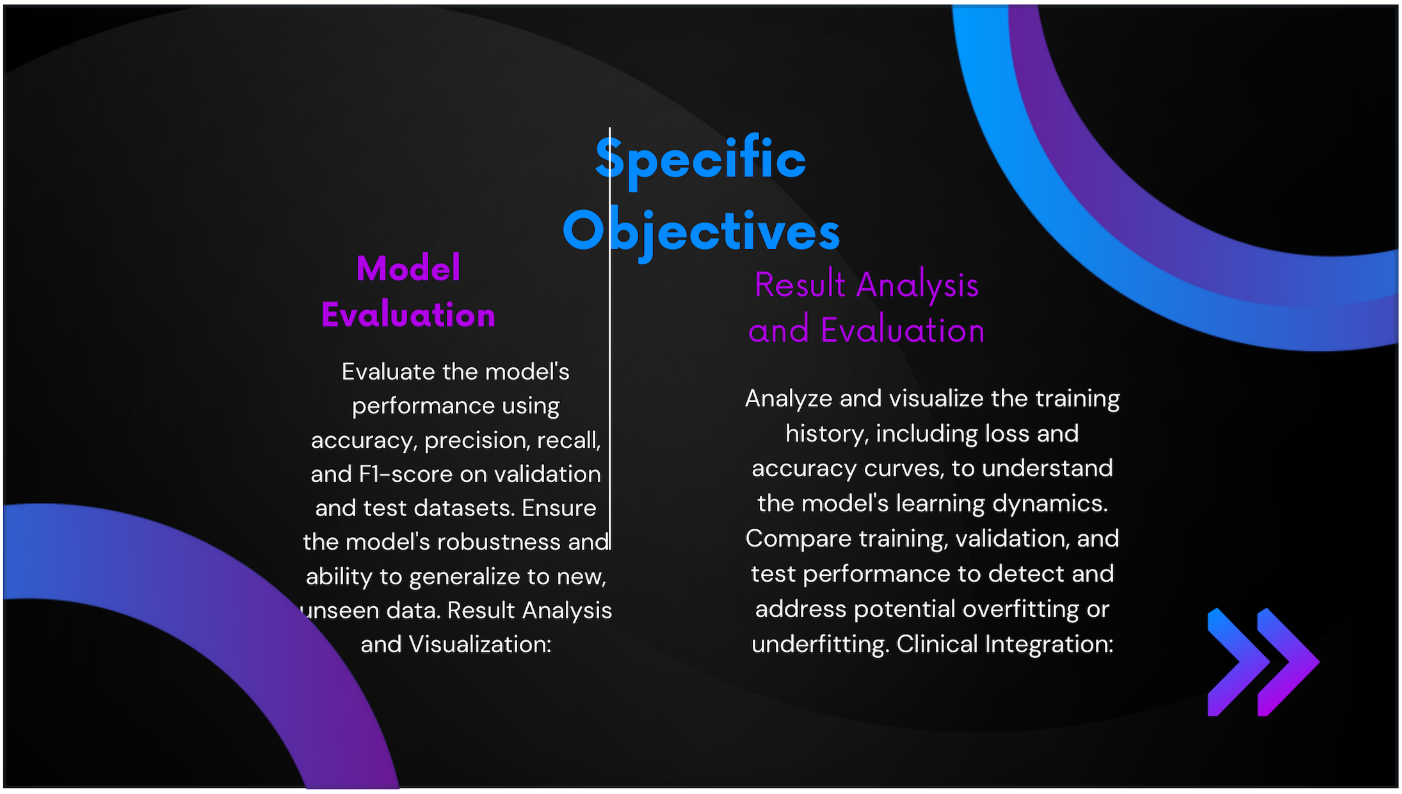
Load, explore, and preprocess the Chest X-Ray Images (Pneumonia) dataset to ensure it is suitable for model training. Implement data augmentation techniques to enhance the training dataset and improve model generalization. Model Development:

### Model Development

Build and train a baseline convolutional neural network (CNN) model to classify chest X-ray images as 'PNEUMONIA' or 'NORMAL'. Experiment with different model architectures and layers to improve performance. Hyperparameter Tuning:

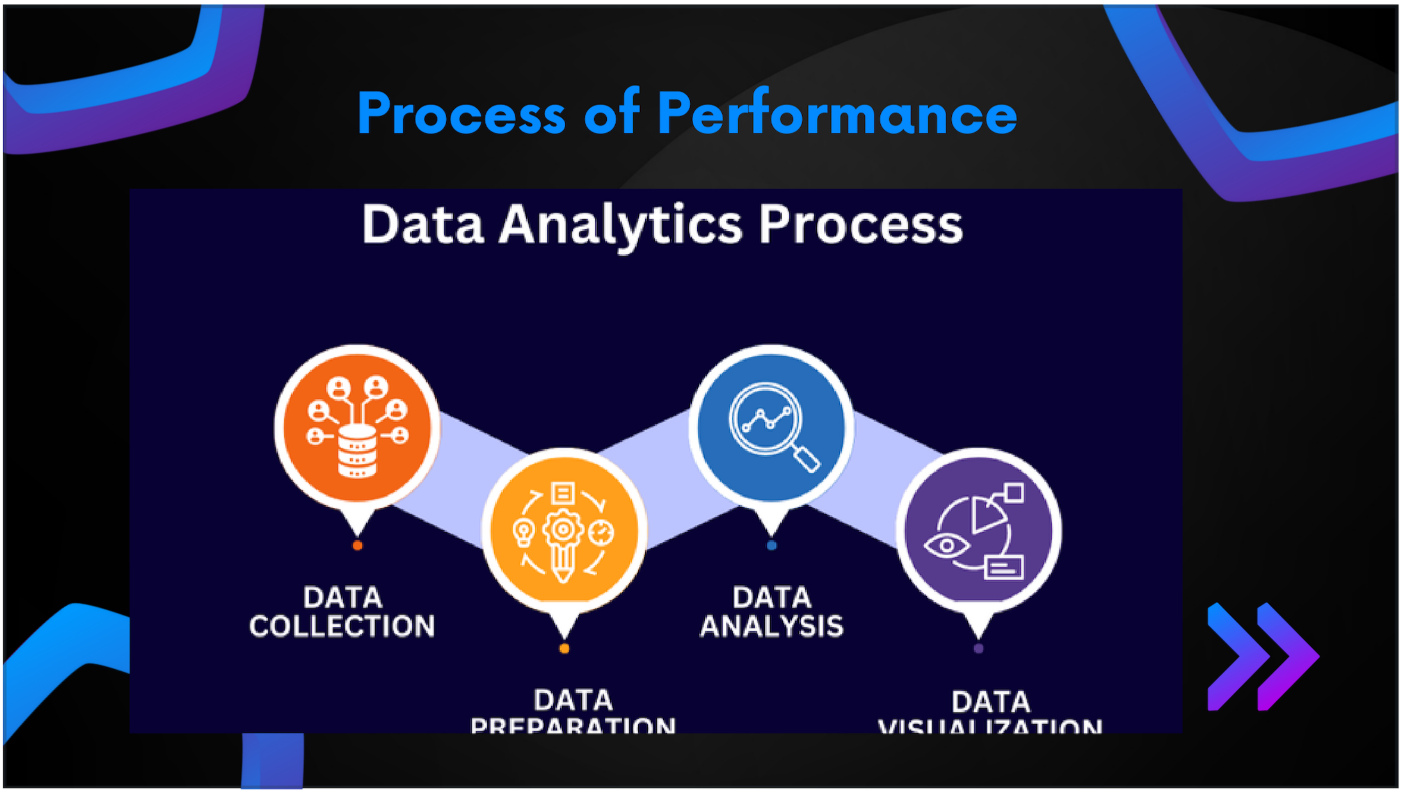
### Hyperperameter Tuning

Utilize Keras Tuner to perform hyperparameter optimization, focusing on key parameters such as learning rate, batch size, and network depth. Model Evaluation:


Specific Objectives
Model Evaluation
Evaluate the model's performance using accuracy, precision, recall, and F1-score on validation and test datasets. Ensure the model's robustness and ability to generalize to new, unseen data. Result Analysis and Visualization:
Result Analysis and Evaluation
Analyze and visualize the training history, including loss and accuracy curves, to understand the model's learning dynamics. Compare training, validation, and test performance to detect and address potential overfitting or underfitting. Clinical Integration:

# Clinical Integration

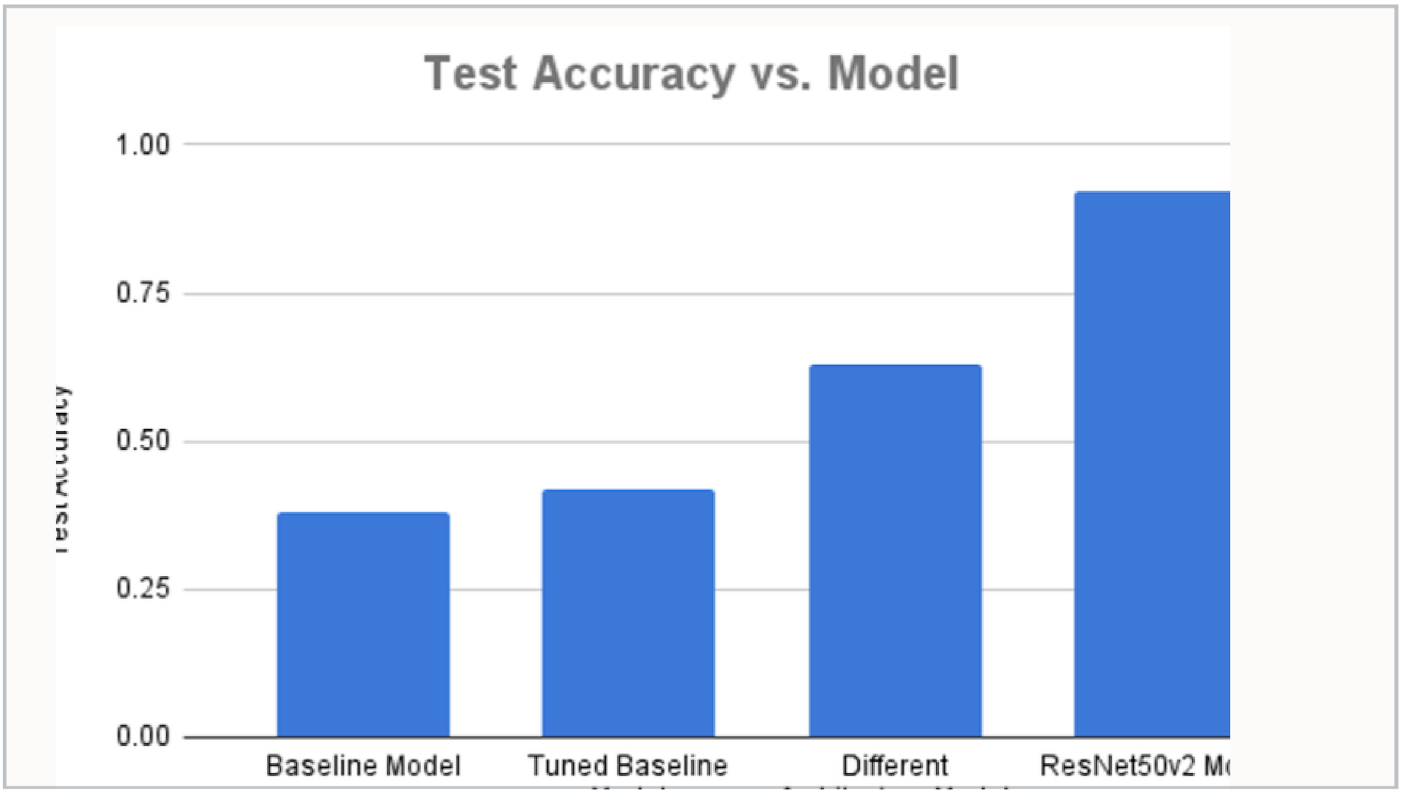
*Propose a strategy for integrating the developed model into clinical workflows to assist in the diagnosis of pneumonia.*

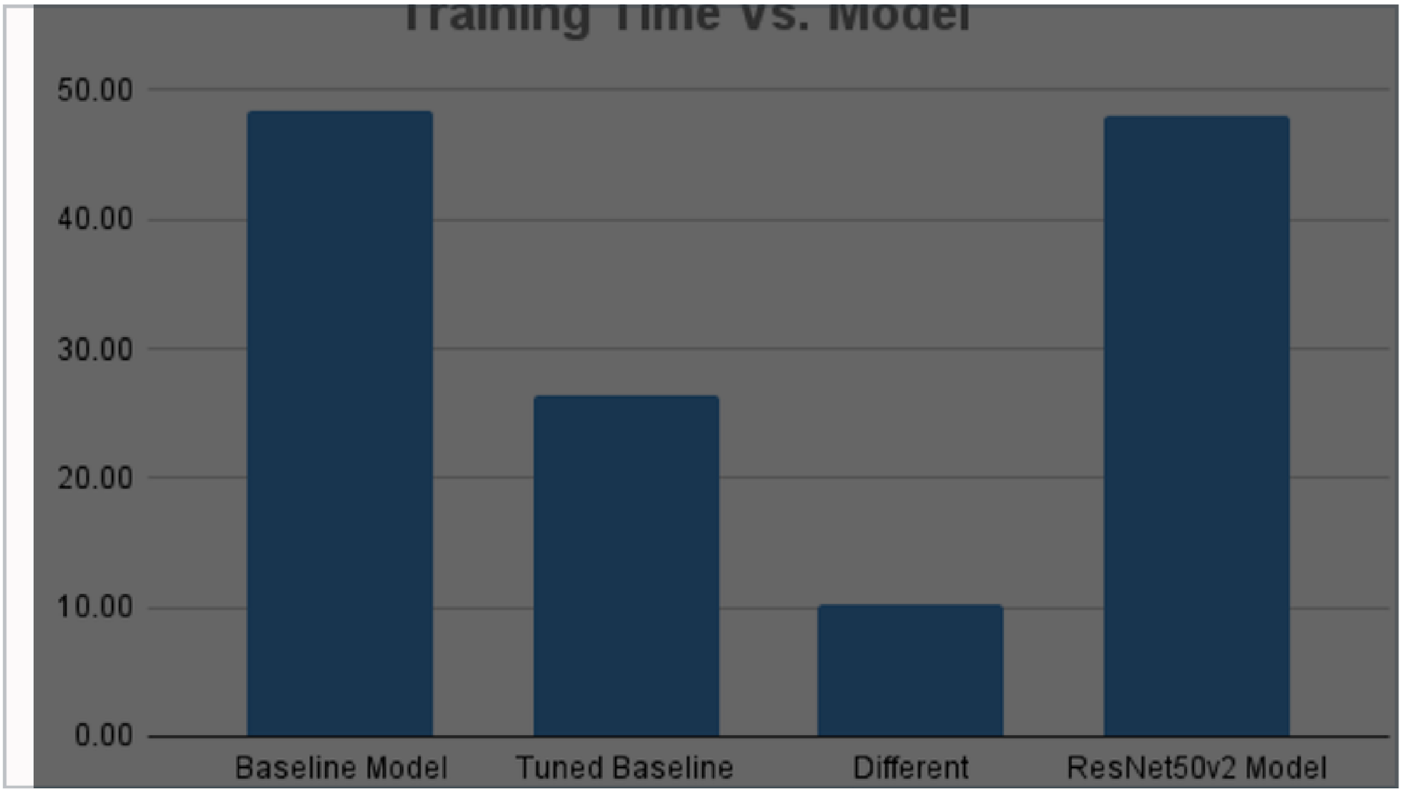

Process of Performance
Data Analytics Process
DATA COLLECTION
DATA PREPARATION
DATA ANALYSIS
DATA VISUALIZATION

## Modeling Architecture

We focused on below modeling for accuracy

Baseline CNN
Tuned baseline CNN
Complex Model
ResNET50V Model


Test Accuracy vs. Model
1.00
0.75
0.50
0.25
0.00
Baseline Model
Tuned Baseline
Different
ResNet50v2 M


Training Time vs. Model
50.00
40.00
30.00
20.00
10.00
0.00
Baseline Model
Tuned Baseline
Different
ResNet50v2 Model

# Results Evaluation

## Accuray

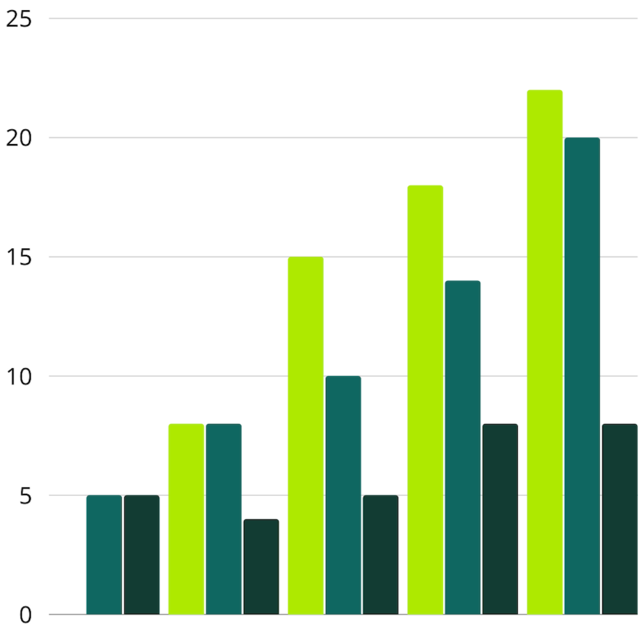
Baseline model accuracy was 0.875

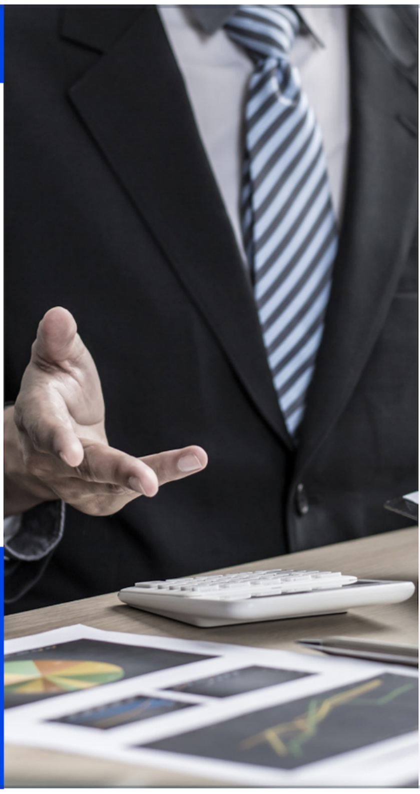
The tunned baseline model gave 0.875

Complex architecture model produced 1.833 approximate

The ResNET50V2 prduced

Tuned ResNET50V2





# DURATION

01 Baseline model took 2hrs 42min 18secs

02 Tunned baseline: 8hrs 44min 19secs for the best trial

03 Complex architecture:2HRS 20min 15 sec

04 ResNET50V2 model: 1hr 30min

# Model Comparison

| SN: | Model Name | Model Architecture | Training Time | Validation Accuracy | Test Accuracy | Test Loss |
|---|---|---|---|---|---|---|
| 1 | Baseline Model | CNN with original architecture | 5hr 11min | 0.875 | 0.74 | 3.89 |
| 2 | Tuned Baseline Model | CNN with original architecture | 2hrs 42min | 0.875 | 0.85 | 3.00 |
| 3 | Complex Architecture | CNN with modified architecture | 10.25min | 0.45 | 0.63 | 19.45 |
| 4 | ResNET50V2 | Pretrained ResNet50v2 | 10.25 mins | 0.88 | 0.92 | 0.23 |

# Recommendation

Further improvement may be consideredValidating the developed model through clinical trials is needed to ensure its reliability for diagnoses.

Focus on creating a system that functions effectively with limited computational resources.

Further optimization of hyperparameters such as dropout rate and batch size to enhance model performance.

Thynk Unlimited

# THANK YOU

MEMBERS

Jane Martha
Najma Abdi
Frida Oyucho
Eunice Ngunjiri
Brian Ochieng
Jubilant Mutuku