Diagnosing Pneumonia

Reducing Costs in Healthcare

Team Members:

Valentina Buritica
Eli Holden
Fazle Hameem
Matthew Musgrave
Emmanuel Korlewala

Motivation/Hypothesis:

Develop a machine learning model that is trained to accurately identify pneumonia in x-ray images. This technology can be used to increase diagnostic efficiency reducing costs for patients. This can also be used to help underserved nations with cheaper alternatives to healthcare which in some cases may save lives.
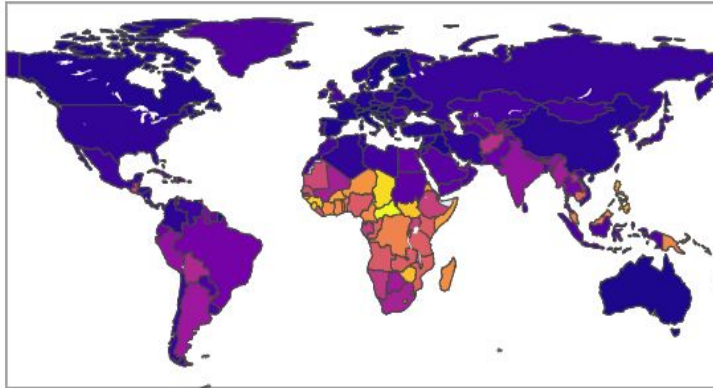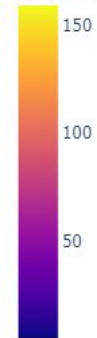
November 12

World Pneumonia Day
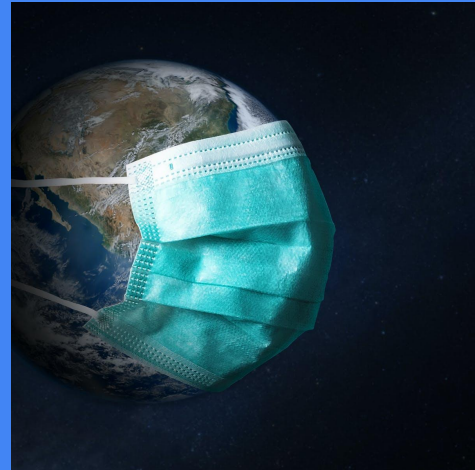
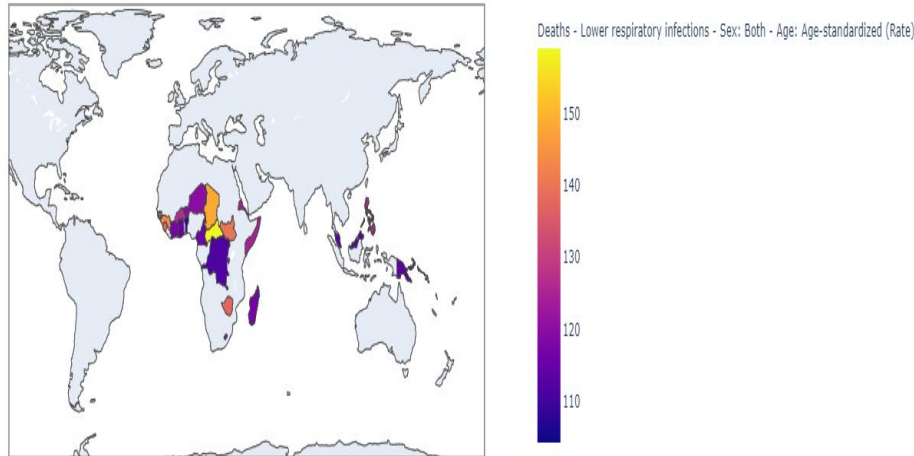# Pneumonia Statistics - Death Rates

Death Rate from Pneumonia 2017 (per 100,000)

Deaths - Lower respiratory infections - Sex: Both - Age: Age-standardized (Rate)

Countries with death rate over 100 per 100,000 cases

Deaths - Lower respiratory infections - Sex: Both - Age: Age-standardized (Rate)

What regions have the highest death rate?

- Sub-Saharan Africa
- South East Asia

What is common between these countries?
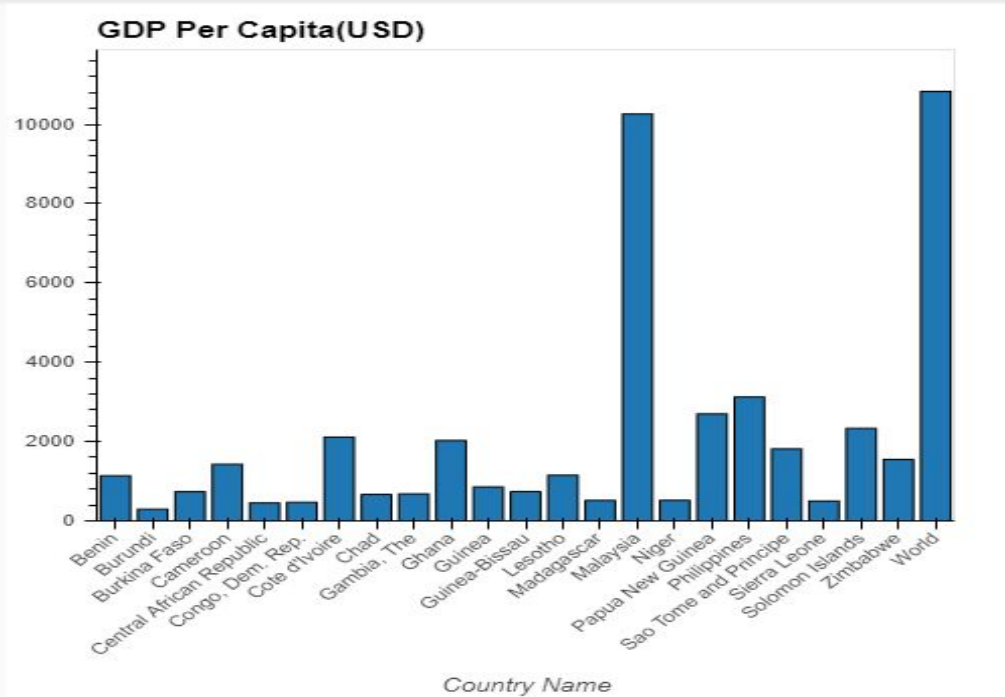
GDP Per Capita(USD) lower than world average

Pneumonia & Healthcare Challenges

Pneumonia is a huge burden on our healthcare systems.

- Pneumonia has been in the top 20 most expensive conditions seen during inpatient hospitalization
- Most cases of pneumonia can be treated with antibiotics costing less than $0.40 for a full course.
- However, poor accessibility and expensive diagnosis poses a challenge

Radiologist In Sub Saharan Africa

- The average radiologist in South Africa makes about - $55,252.89 USD
- The main issue is the lack of radiologist not only in South Africa but in many of the pneumonia stricken countries
- There are more radiologists working in the four teaching hospitals on Longwood Avenue in Boston, Massachusetts, than there are in West Africa.

The Model

The Model - Convolutional Neural Networks (CNN)

- CNN's are a class of deep neural networks commonly applied to analyzing visual imagery.

- CNN's can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.

```
Model: "sequential_5"
_____
Layer (type)                 Output Shape              Param #
=================================================================
sequential_4 (Sequential)    (None, 180, 180, 3)       0
_____
rescaling_5 (Rescaling)      (None, 180, 180, 3)       0
_____
conv2d_9 (Conv2D)            (None, 180, 180, 16)      448
_____
max_pooling2d_9 (MaxPooling2 (None, 90, 90, 16)        0
_____
conv2d_10 (Conv2D)           (None, 90, 90, 32)        4640
_____
max_pooling2d_10 (MaxPooling (None, 45, 45, 32)        0
_____
conv2d_11 (Conv2D)           (None, 45, 45, 64)        18496
_____
max_pooling2d_11 (MaxPooling (None, 22, 22, 64)        0
_____
dropout_1 (Dropout)          (None, 22, 22, 64)        0
_____
flatten_3 (Flatten)          (None, 30976)             0
_____
dense_6 (Dense)              (None, 128)               3965056
_____
dense_7 (Dense)              (None, 2)                 258
=================================================================
Total params: 3,988,898
Trainable params: 3,988,898
Non-trainable params: 0
_____
```
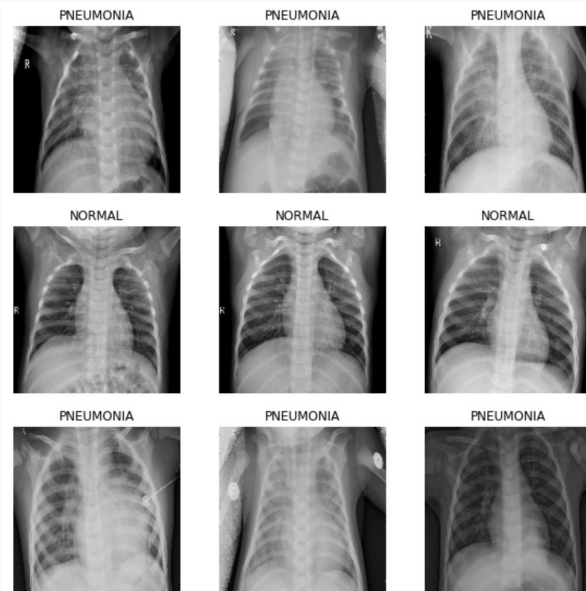
What is a Convolutional Neural Network?

Each convolutional layer has a number of filters that move through and image. That process is called convolving. The filter convolves the pixels of the image (changes the values) and passes it to the next layer in the CNN model.
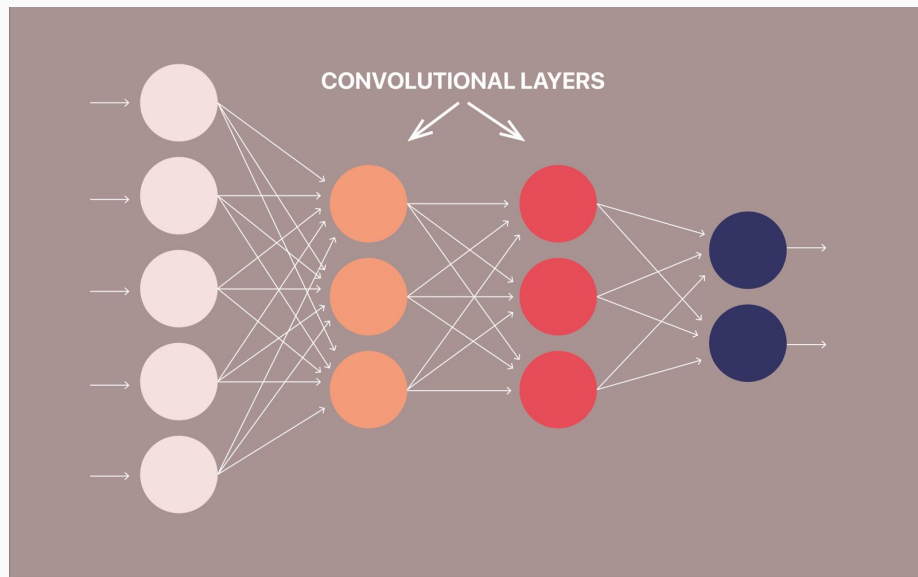
Read more about this here: https://lionbridge.ai/articles/difference-between-cnn-and-rnn/

How CNN's work:

- The CNN model focuses on key features of the image by passing the image through different layers and honing in those features.
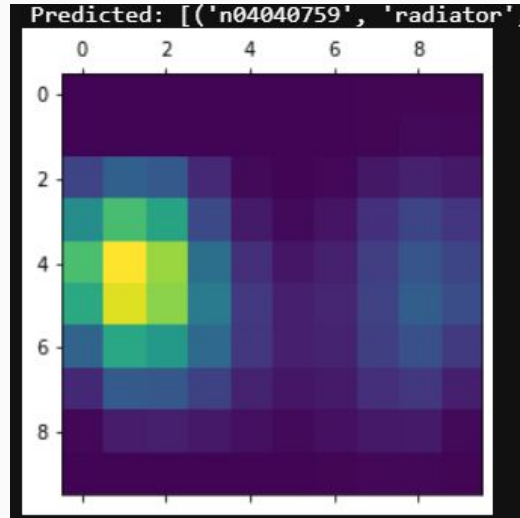
Data Preparation & Model Training

- Use of data augmentation to create "new" images by rotating and flipping
- Incorporated dropout layers to reduce the impacts of overfitting
- Assigned weights to balance classes of data
- Use re-scaling to normalize the images

```
Weight for class 0: 2.03
Weight for class 1: 0.66
```

Model Performance

Training and Validation Accuracy

Training and Validation Loss
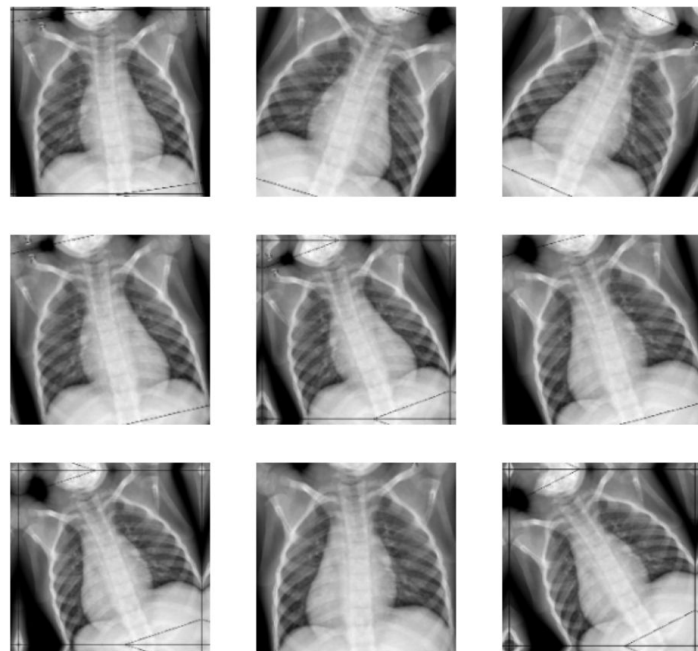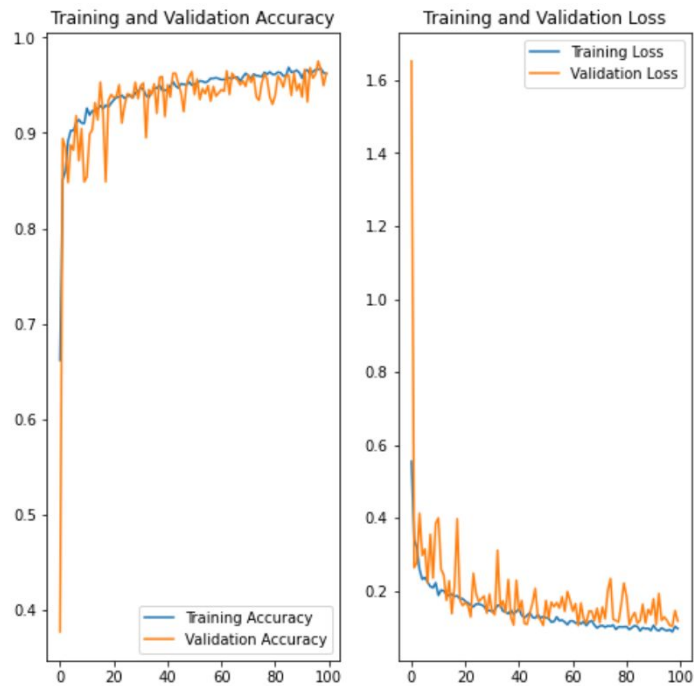
Confusion Matrix

	Precision	Recall	f1-score	Support
Normal	0.89	0.97	0.93	288
Pneumonia	0.99	0.96	0.97	883
Accuracy			0.96	1171
Macro avg	0.94	0.97	0.95	1171
Weighted avg	0.96	0.96	0.96	1171

Conclusions & Predictions

Medical Cost Savings in Developing Nations

Health Expenditure per Capita / GDP Per Capita (USD)
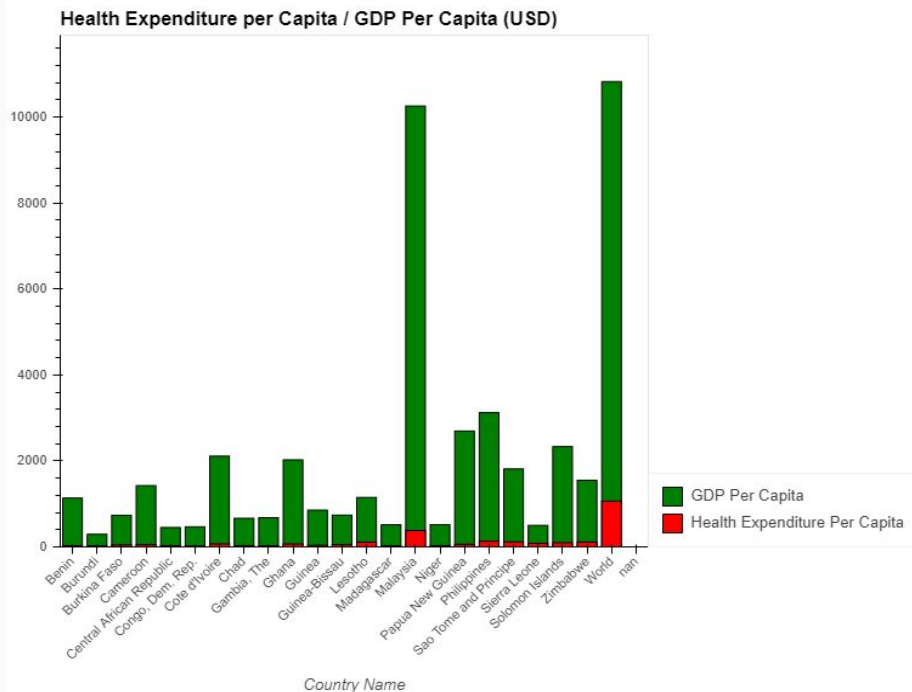
Mean (SD) Laboratory cost: $7.72

Our Machine Learning Model will not only give an accurate diagnostic results but will also save $7.72 per patient

For every 100 million cases, it will save $772,000,000

There are about 150 million cases each year

Diagnostic Robot

Live Demonstration

Challenges & Opportunities

- Adapting the deep neural network concepts learned in class to image recognition and analysis.
- Identifying useful applications of our deep neural network machine that would result in cost savings.
- We could refine our use case by digging deeper into the cost savings of deploying this machine to areas that would most benefit.
- Having more time for training would allow the complex model to produce better results (images slow the machine learning significantly)
- More labeled x-ray images in the data would likely add to accuracy of the model

Sources

- https://www.tensorflow.org/tutorials/images/classification
- https://keras.io/examples/vision/grad_cam/
- https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
- https://www.cdc.gov/dotw/pneumonia/index.html
- https://ourworldindata.org/pneumonia
- https://www.medscape.com/answers/967822-23537/what-is-the-global-prevalence-of-pneumonia#:~:text=The%20WHO%20Child%20Health%20Epidemiology,enough%20to%20require%20hospital%20admission.
- https://gh.bmj.com/content/2/1/e000166
- Current health expenditure (% of GDP) - Chad, Benin, Burundi, Burkina Faso, Cameroon, Congo, Dem. Rep., Cote d'Ivoire, Gambia, The, Guinea, Guinea-Bissau, Lesotho, Madagascar, Malaysia, Niger | Data (worldbank.org)