Detecting Pneumonia

A Computer Vision Task

Project and Presentation by: Samuel Middleton

Business Case

- Use technology and emerging fields of study to lower healthcare costs.
- Specifically, using a form of AI to detect pneumonia in chest x-rays.
- Work with clients to provide a cloud option for deployment.

WHAT WE ARE WORKING ON

Data

Chest X-Rays that show both normal lungs and pneumonia infected lungs.

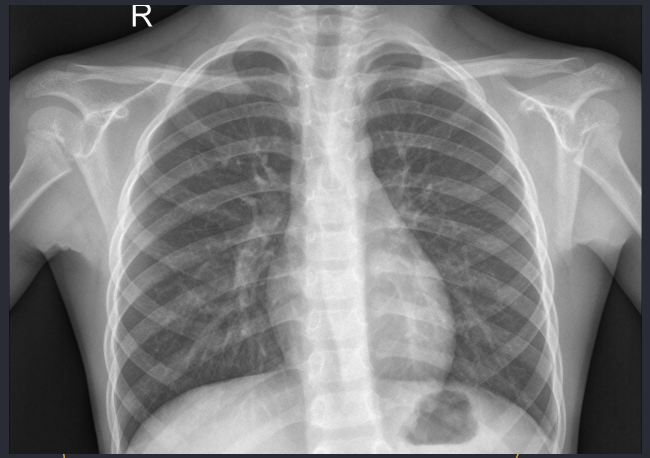
Models

Machine Learning
Using mathematics to derive relationships.

Communication

Our data driven recommendations.

ABOUT THE DATA



- Kaggle sourced.
- There are 5,863 X-Ray images and 2 categories, Pneumonia/Normal.
- Chest X-ray images from pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou.

Modeling

Neural Networks

Having a computer attempt to replicate how a brain learns.

Framework

PyTorch is a neural network framework developed by Facebook's AI team.

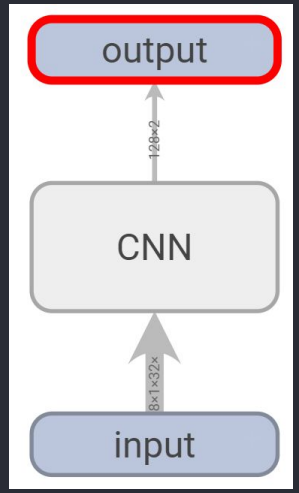
Insights

Our model has currently been trained and can predict, with a certain degree of accuracy pneumonia in patients.

PyTorch Model

- Deep Learning - A Neural Network that learns by feeding information through more mathematical transformations.

  Metric:
  Precision - A measure that will allow us to see at what rate we are correct when we predict a case.
  Optimizing this metric means we collect the most True Positive cases at the risk of collecting a few False Positive cases.



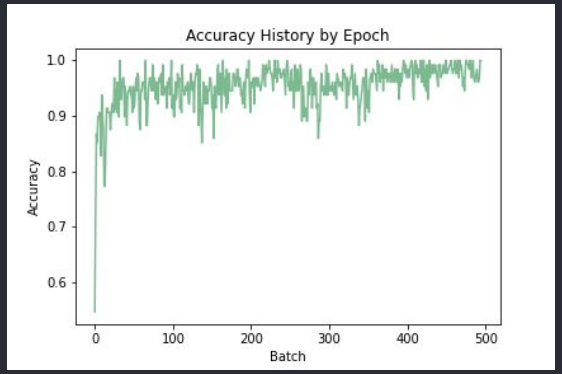
Graphical representation of model

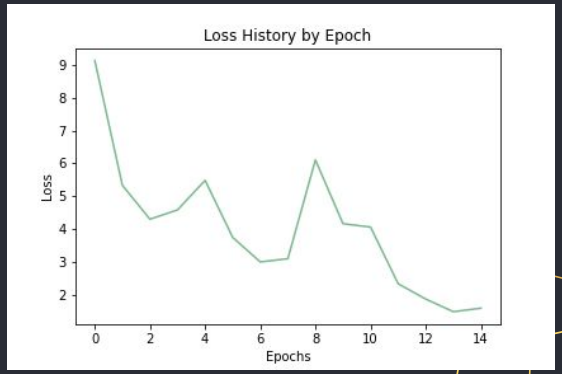
PyTorch Model

Training and Testing

- Loss - An amount that measures how wrong a model is.
- Accuracy - How correct a model is in percentage.

What these charts show is the relationship between the model learning more (lower loss) and how its accuracy improves as it learns.





PyTorch Model

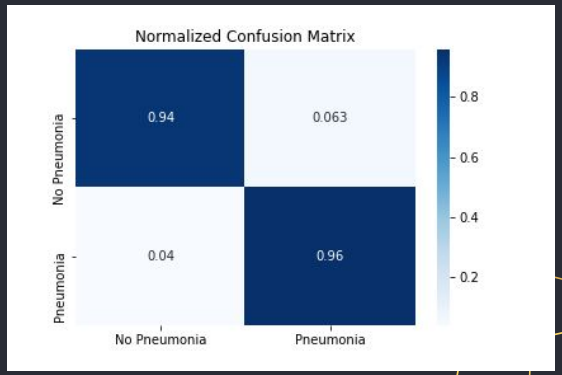
Results

- Confusion matrices are a common method of showing how well a model performed by show True Positives, False Positives, True Negatives, and False Negatives.
- Top matrix is actual counts.
- Bottom matrix displays normalizations for the top chart.

These can be used to determine other metrics, such as our chosen metric of precision.

Using the formula below our precision is **95.6%**, meaning almost 96% of our predictions of Pneumonia cases are correct.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$



Recommendations

Modeling

During the modeling phases of the project we were able to duplicate our original model, closely, in another popular deep learning framework.

This gave us deeper insights into the problem at hand and allowed us to compare the two frameworks.

Due to ease of development and deployability, as well as extensibility and flexibility, we recommend the use of the PyTorch framework.

Recommendations

Cloud Deployment

- Deep Learning Networks can be setup localcally but the start up costs, maintenance, and labor required to run them are prohibitive.
- These operations are significantly more likely to be cloud based deployments also known as PaaS (Platform as a Service).
- Generally, Cloud platforms are a pay-by-use system and will only incur costs during use.

Therefore, we recommend a cloud deployed modeling solution.

Recommendations

Storage Requirements

Storage can be a concern for cloud deployment, therefore our model has been engineered with these requirements in mind.

We accomplished this by using transformations to the images that allow for smaller image size and therefore less data used.

How we were able to achieve this spatial savings:

- Images taken in by the model only need be 32x32px, thereby significantly decreasing the drive space required for them.
- Reducing any images that are saved as RGB to single channel Grayscale.

Recommendations

In the end, we suggest that the way of the future in the healthcare industry is paved with machine learning. Our model could easily be deployed to a cloud solution. A specialized and user friendly front-end can shoulder the burden by allowing the layman to feed in the X-Rays. Finally, we have arrived at the conclusion that one framework was more desirable than another, but in the end that is more or less a matter of preference and willingness to invest resources.

Future Works

- Further refinement of our model for higher accuracy
- Cloud based deployment
- User friendly dashboard deployed for cloud solution

THANKS!

DO YOU HAVE ANY QUESTION?

samuelmiddleton93@gmail.com

https://www.linkedin.com/in/samuel-middleton-a8533491/

RESOURCES

PyTorch - https://pytorch.org/

Deep Learning - https://www.investopedia.com/terms/d/deep-learning.asp#:~:text=Deep%20learning%20is%20an%20AI,is%20both%20unstructured%20and%20unlabeled.

Amazon Web Services - https://aws.amazon.com/