MEMORANDUM

Date: Feb 22, 2022
To: Bill and Melinda Gates Foundation
From: Shu Zhang
Subject: Deep Learning Model on Children Pneumonia Diagnose

Executive Summary

Based on my research, the application of CNN(Convolutional Neural Network) model can diagnose pneumonia very accurately from chest x-ray screening results. I hope the Gates Foundation can provide $2,000,000 USD in funding to scale up my deep learning model to help improve pneumonia detection accuracy in some developing countries where pneumonia diagnose accuracy is low.

Background

More than 60% of pneumonia deaths occur in sub-Saharan Africa, and two-thirds of these are in children under the age of five. In some Asian countries with low levels of medical detection, the accuracy of radiologists in detecting pneumonia in children is less than 85%. Based on this situation, I believe it is necessary to pilot the use of my deep learning model in developing countries where specialized radiologists are scarce and/or have the insufficient diagnostic capacity to help avoid incorrect diagnoses. Also, the application of the model can help save the loss caused by misdiagnosis.

Description of Analysis

A dataset containing x-ray images of 2900 pediatric patients will be used to train and validate models that will attempt to predict pneumonia in 200 test cases.

Model Building and Testing

The CNN model is a type of neural network commonly used in computer vision and natural language processing tasks. In this case, I applied a CNN model for pneumonia diagnosis based on chest X-ray examination. The model architecture is shown in Table 1. Through testing, I determined that this set of parameters enables the model to achieve a high level of accuracy, which is 95.36%.

Model Evaluation

Three ways are applied to evaluate the accuracy of the model.

Evaluation	Result
Confusion Matrix	Based on Chart 1, The true health(TN) rate is 0.3248, the true diagnose(TP) rate is 0.6387, the false diagnose(FP) rate is 0.0092, and the false health(FN) rate is 0.0371, showing a very high prediction accuracy in this case.
ROC	The ROC result in Chart 2 shows that the AUC is close to 1, indicating the model performs well.
Top N% Predictions	From Chart 3, we can see that the prediction accuracy indicators of Healthy and Pneumonia are both close to 1.

Based on the above three assessment methods, it can be confirmed that the model has a high accuracy and can effectively detect whether the person has pneumonia.

Recommendations based on Cost - Benefit Analysis

The Cost-Benefit Analysis method was utilized to assess the potential benefit of the model extension due to its high accuracy.

Take South Africa areas as example, the percentage of false negatives(false health) and the percentage of false positives(false diagnose) of chest X-ray tests for pneumonia in children in the South African region is relatively high. According to a study published in the International Journal of Public Health[1], pneumonia diagnoses are performed with chest X-ray in 20 different types of health care facilities in Limpopo Province, South Africa. The results found that the majority of these institutions had between 20% and 30% false negatives and between 10% and 20% false positives. Thus we can assume the false health rate is 25% and the false diagnose rate is 15%.

In terms of financial cost, the cost of private health care in South Africa can vary widely depending on the type of treatment required, the medical facility used, and the individual insurance coverage. A 2018 report by the South African Private Practitioners Forum estimated the average cost of a hospital stay in a private hospital to be around ZAR 7,500 ($500 USD) per day[2]. Since the expense ratio of public hospital and private hospital is around 1:5, we can assume the expense in public hospital is $100. According to a 2020 study published in the South African Medical Journal, the average cost of ICU care in South Africa is approximately ZAR 25,000 ($1,700 USD) per day[3].

Besides, according to a study published in the South African Medical Journal in 2016, the average length of stay in the ICU for patients with community-acquired pneumonia (CAP) was 5.6 days[4]. In addition, the South African national treatment guidelines for community-acquired pneumonia recommend a treatment duration of 5 to 7 days for uncomplicated pneumonia, and 7 to 10 days for severe or complicated cases. Thus, we can assume the average days of treatment in South Africa is 7 days.

Rate			Unit Cost	Cost
FN	Original	0.25	$100 * 7d	$175
	Improved	0.0371		$25.97
FP	Original	0.15	$1,700 * 5.6d	$1428
	Improved	0.0092		$87.584

Based on above information, the benefit of applying this model for each child can be (175+1428)-(25.97+87.584)=$1489.446. Since the number of children aged 0-5 in South Africa was about 13.06 million in 2019[5], it is reasonable to estimate the total benefit as $1489.446*13.06M=19.452B. This means that every dollar in that $2 million spent in scaling this model is worth nearly 10,000 times more than it is. This is merely in South Africa.

If we could apply this model to more countries, this would yield even higher value that can change the life of so many children. Currently, I plan to set up a start-up to put my model into practice. I sincerely hope that the Gates Foundation will provide $2,000,000 for my project, helping more children live their healthy life.

Appendix

[1] N. N. Ntuli, J. E. Mokoena, and K. J. Sithole: Assessment of the Impact of a Community-based Health Programme on Adult Mortality, International Journal of Public Health, PMID: 28321264, 2017.
[2] Annual Report 2018, South African Private Practitioners Forum(SAPPF), 2018.
[3] The Cost of Intensive Care in South Africa, South African Medical Journal, 2020.
[4] South African Medical Journal, 2016.
[5] Article index: SP.POP.0004.TO.ZS(South Africa), The World Bank, 2019.

Table 1: The Architecture of CNN Model

numLayers	3
numFilters	32
kernelSize	5
dropoutValue	0.4
maxPooling	3
numClasses	2
batchSize	32
learningRate	0.0001
epochs	60

Chart 1: Confusion Matrix

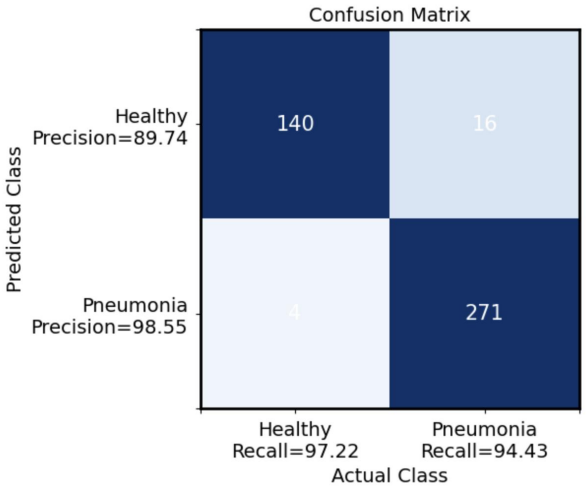


Chart 2: ROC Result

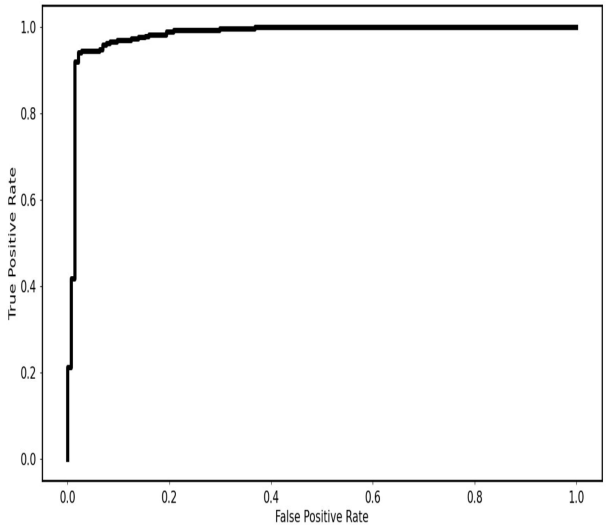
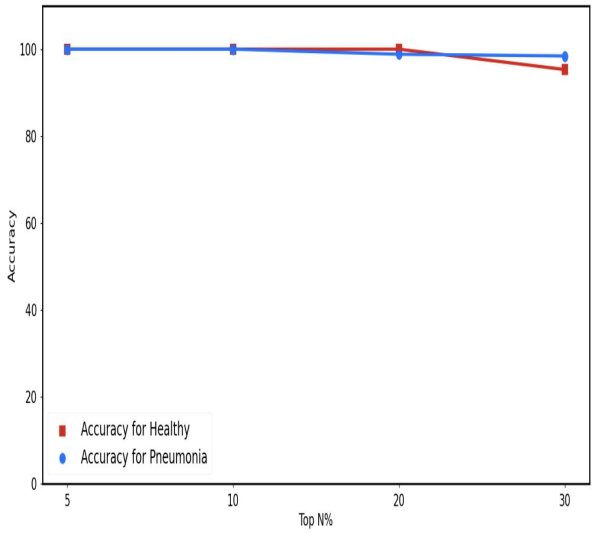


Chart 3: Top N% Predictions



Deepnote Link:
https://deepnote.com/workspace/shu-a130-1179f4e9-b851-4875-b4d7-38aaa3c9e408/project/Deep-Learning-for-Image-Classification-Duplicate-3d48f559-bc57-45eb-b918-cc56126236f9/notebook/DeepLearningCNNPlaybook-d966260334694dab8bc64f9aa278ab8a#/sidebar/project