**IMAGE CLASSIFICATION WITH NEURAL NETWORKS**

Pneumonia, a serious respiratory infection caused by viruses or bacteria, is a worldwide public health concern. It is the greatest infectious cause of death in children worldwide, killing around 740,180 children in 2019. Furthermore, vulnerable populations such as elders, smokers, and people with underlying health concerns are more vulnerable. While pneumonia has a terrible impact on communities and families, it can be prevented and treated with basic actions and low-cost drugs. The importance of early detection and proper diagnosis in saving lives and minimizing the burden cannot be overstated. Despite limited resources in low- and middle-income nations, focusing on prevention and access to reasonable therapies can lower death rates dramatically.

1. **Problem Statement**

Pan African Medical Training & Research Center has tasked the team with developing a neural network model to accurately detect pneumonia in patients by analyzing chest X-ray images in the Chest X-Ray Images (Pneumonia) dataset.

We focus on leveraging advanced machine learning techniques on this dataset containing numerous images. Thorough training on this dataset will ensure the model's reliability and performance across different scenarios. Once integrated, the model will provide automated analysis, assisting treatment decisions and reducing potential human errors. It will also serve as a valuable training tool, empowering medical professionals to enhance their diagnostic skills. Our goal is to improve patient outcomes, contribute to medical research, and advance the field of pneumonia diagnosis.

1. **Data understanding**

The Chest X-Ray dataset focuses on pneumonia detection in chest X-ray images. It consists of a collection of chest radiographs labeled as either "normal" or "pneumonia" cases. The dataset has been derived from multiple sources, including pediatric and adult patients, and encompasses various types of pneumonia, such as bacterial and viral.

The images are in JPEG format and typically grayscale, with varying resolutions and sizes. The dataset contains training and testing subsets, enabling the development and evaluation of machine-learning models for pneumonia detection.

[*https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia*](https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia)

1. **Objectives**

The main goal of this undertaking is to create a Neural Network model capable of reliably categorizing chest X-ray pictures as having or not having pneumonia.

Objectives:

* Train and refine a Neural Network model to obtain high accuracy in categorizing chest X-ray pictures as positive or negative for pneumonia.
* Validate the trained model using appropriate evaluation metrics and datasets to ensure its robustness and generalizability in real-world scenarios.
* Document and communicate the findings and insights from the project to contribute to the scientific community's knowledge and advance the field of pneumonia detection using a deep learning technique.

1. **Success Metrics**

The performance of the model is evaluated based on:

* Accuracy - achieving an accuracy of over 85%
* Recall - achieving a recall of over 65%
* Loss Function - minimize loss function to ensure the model is not overfitting

1. **Modeling**

For this study the following models were implemented and based on performance and the best was chosen based on its individual success metric. The models tested were:

* Baseline model: Densely Connected Neural Networks
* Convolutional Neural Networks Model (CNN)
* CNN with regularization & dropout
* CNN with architecture modifications
* CNN with transfer learning

1. **Evaluation**

Theappropriate model for this study was chosen based on its performance against the success metrics stated above. Ie. Accuracy, recall & loss function

Here is a breakdown of the evaluation results of the top two performing models:

**Model 2 Simple CNN:**

* Test Loss: ***0.3740***
* Test Accuracy: ***0.8317***

**Model 6 (Resized Images):**

* Test Loss: **0.5205**
* Test Accuracy: ***0.7628***

These results indicate the performance of each of the models on the test dataset. Lower values for test loss indicate better performance, while higher values for test accuracy indicate higher accuracy. Based on the results, Model 2 performs the best with a test loss of ***0.3740,*** a test accuracy of ***0.8317*** and a recall score of ***93*** percent for pneumonia and ***67*** percent for Normal.

A recall score of ***0.93*** for the "PNEUMONIA" class means that the model correctly identified ***93%*** of the actual pneumonia cases in the dataset. This indicates a high sensitivity of the model in detecting pneumonia, as it has a low likelihood of missing positive cases.

1. **Conclusion**

In conclusion, we have developed an image classification model for detecting pneumonia from X-ray images. After training the model on a dataset of X-ray images, we achieved an accuracy of ***83%*** on the training dataset and loss of ***0.37***.

A high recall score is desirable in medical diagnosis tasks like pneumonia detection because it is crucial to minimize false negatives (missing positive cases). In the context of pneumonia, a false negative would mean failing to identify a patient with pneumonia, leading to potential health risks and delayed treatment.

Therefore, a recall score of ***0.93*** suggests that the model has a strong ability to accurately detect pneumonia cases in the X-ray images, which is a positive outcome for the classification task.

Despite our efforts, it is important to acknowledge that the model's performance fell short of the targeted accuracy of ***85%***. While achieving high accuracy is desirable, it is not always guaranteed, especially in complex and challenging tasks like pneumonia detection from X-ray images.

Several factors might have contributed to the inability to reach the desired accuracy: limited data, complexity of the task, chosen model architecture and its capacity to capture relevant features, and sensitivity of the model to hyperparameters. Pneumonia detection from X-ray images can be affected by various factors, including imaging artifacts, inter-patient variability, and the presence of other lung conditions. These challenges can make accurate classification more difficult, even with a well-designed model.

1. **Recommendations**

* Pan African Medical Training and Research should use our CNN model for X-ray image classification in the field and also when training researchers in class.
* Deploy the developed image classification model for pneumonia detection from X-ray images in the real world. It will be useful in assisting healthcare professionals in the diagnosis of pneumonia., increase efficiency and speed in diagnosis and act as a second opinion.