Title: Medical Vision Classification Using Deep Learning: A Case Study on Pneumonia Detection from Chest X-Ray Images

Abstract
This project focuses on the development of a medical vision classification system using Convolutional Neural Networks (CNN) to detect pneumonia from chest X-ray images. Initially set to work on a hyper-tuning summarization chatbot based on GPT-NEO, the project shifted focus due to the other group planing for the same. The key challenge was to develop a Convolutional Neural Network (CNN) model that could accurately distinguish between normal and pneumonia-affected chest X-rays. Despite achieving high accuracy on the training set, the model struggled with overfitting, resulting in lower test set accuracy. Various strategies, including learning rate adjustments and dropout regularization, were employed to improve performance.

Introduction
The project aimed to develop a deep learning-based approach for medical image classification, focusing on detecting pneumonia from chest X-ray images sourced from Kaggle. Pneumonia detection is a critical task in medical diagnostics, where rapid and accurate identification can significantly impact patient treatment outcomes. The motivation behind this project was to leverage the capabilities of CNNs in image recognition to assist in medical diagnoses.

Problem and Motivation
The primary challenge addressed was the accurate classification of chest X-rays into 'Normal' (labeled as 0) and 'Pneumonia' (labeled as 1) categories. The motivation for choosing this problem lies in its potential to aid healthcare professionals in diagnosing pneumonia more efficiently and accurately, thereby improving patient care.

Approach
The approach involved using a CNN, renowned for its effectiveness in image classification tasks. The initial model was trained for 10 epochs without dropout. Subsequent iterations experimented with learning rate decay and dropout regularization to mitigate overfitting and improve generalization. The final model configuration included 100 epochs with a learning rate decay at the 10th and 30th epochs and dropout layers.

Rationale
The use of CNNs was inspired by their proven success in similar image classification tasks. The learning rate decay and dropout were implemented to address overfitting, a common issue in deep learning models, particularly when dealing with high-dimensional data like images.

Key Components and Limitations

The model's key components included convolutional layers for feature extraction and dropout layers for regularization. One limitation observed was the model's tendency to overfit, as evidenced by the disparity between training and test accuracies.

Setup

- Dataset: The dataset comprised chest X-ray images, categorized as 'Normal' and 'Pneumonia'. https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia
- Experimental Setup: The experiments were conducted using various CNN architectures, with parameters like learning rate and dropout rate being iteratively adjusted.
- Computing Environment: The model was trained and tested on a GPU-enabled environment, ensuring efficient handling of the computationally intensive tasks.

Results

- Main Results: The best model achieved about 80% accuracy on the test set, with near-perfect training accuracy, indicating overfitting.
- Supplementary Results: Learning rate adjustments and the introduction of dropout were critical in addressing overfitting, though they did not completely eliminate it.

Discussion

While the model demonstrated reasonable performance, its tendency to overfit suggests a need for further refinement, possibly through more sophisticated regularization techniques or more extensive data augmentation. Comparatively, existing approaches in literature might offer insights into alternative architectures or training strategies that could yield better generalization.

Conclusion

This project highlights the challenges and potential of using CNNs for medical image classification. Despite achieving high accuracy on training data, the model's performance on unseen test data underscores the complexity of generalizing deep learning models in medical diagnostics. Future work could explore more advanced techniques to improve the model's robustness and accuracy.