Project Proposal

1. Project Title: Deep Learning for Pneumonia Diagnosis: A CNN-Based Approach
2. Team Information: Justin Brent
3. Project Overview: The problem that I aim to address in this project is detecting pneumonia from chest X-ray images using a CNN based approach. I will be using a dataset of over 5000 chest X-ray images that is publicly available on Kaggle.com. In addition to this, I will be using image augmentation techniques to increase the overall size of the dataset. Several diseases such as cancer, tumors, pneumonia, and more are currently using computer-aided diagnosis, and this is a rapidly growing field.
4. Objectives: The specific objectives and goals of this project for me, at first, is learning enough information about computer vision in order to accomplish this task. This will include watching online videos about this topic. In addition to this, I have located a few research articles that have done similar projects, and I wish to replicate the results (such as accuracy) of these projects by using similar techniques, such as transfer learning and data augmentation.
5. Background and Literature Review: Since this dataset is available on Kaggle, I believe that there are several people that have tried various solutions to this problem. There are also existing peer-reviewed articles about this topic. From my research, the main approach that most people have taken in this problem is utilizing transfer learning from larger, pre-trained models such as ResNet-50, and using data augmentation to produce higher accuracy. One of the major challenges of this project is the limited size of the dataset, so this will have to be mitigated through data augmentation to produce better results.
6. Methodology: The main approach that I will be taking to achieve my project objectives is learning enough information about computer vision to begin this project. This will mainly consist of watching online videos and reading related articles about this topic. Once I am comfortable enough in being able to program a very simple solution, without using any advanced techniques, I will start off with that and record my results. I will then begin to try using transfer-learning from various CNN’s and different data augmentation techniques to produce better results.
7. Scope: This project’s features and functionalities will include dataset preprocessing, CNN-Based Model development, model training and optimization, evaluation and performance metrics, and possibly a User Interface and model deployment.
8. Technical Requirements: Google Colab, PyTorch or TensorFlow, Python, NumPy, Pandas, Matplotlib
9. Project Timeline: I do not have a specific timeline for this as much of what I have to do is learning; however, I expect to be able to start programming this project by the beginning of March. During this timeframe of programming, the key steps are 1) Developing a simple CNN with no data-augmentation and recording the results 2) Developing a CNN using transfer-learning and data-augmentation and recording the results 3) Try other CNNs for transfer-learning and more data-augmentation techniques to produce greater accuracy.
10. Expected Outcomes: I expect to achieve an accuracy of at least 80% or higher for this project. Mainly, I expect to have gained a good understanding of computer vision that will help me in the future.
11. Conclusion: For this project I will be working on computer-aided diagnosis of chest X-rays for the detection of pneumonia. I will be utilizing a publicly available dataset of chest X-ray images, transfer-learning of CNNs, and data augmentation to build my own model to accomplish this task. Pneumonia is a leading cause of death worldwide, and the field of computer vision in healthcare is a current growing field.
12. References:
    1. Dataset: <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
    2. Peer-Reviewed Article: Automated detection of pneumonia cases using deep transfer learning with paediatric chest X-ray images
       1. Link to Peer Reviewed Article: <https://pubmed.ncbi.nlm.nih.gov/33861150/>