

ML Project Proposal: Pneumonia Classification Using Chest X-Ray Images

Prabhat Soni 2018252

Lavanya Verma 2018155

Aditya Saini 2018125

1 Motivation and problem statement

Pneumonia is a deadly lung infection which impacts millions of people around the world. According to a report by the American Thoracic Society in 2015, it was the world's leading cause of death among children under 5 years of age. This project deals with classification of pneumonia based on chest X-ray images using Machine Learning based techniques. With this project, we hope to make the diagnostic process easier by providing quick and accurate results.

2 Data Acquisition effort

We'll be using a Kaggle dataset hosted publicly under the Creative Commons License at <https://data.mendeley.com/datasets/rscbjbr9sj/2>.

3 Preprocessing techniques to be explored

- Image Normalization
- Image Padding

4 Learning techniques

Feature Extraction techniques

- Histogram of Oriented Gradients(HoG)
- Local Binary Patterns(LBP)

Proposed Baseline: SVM+HoG/LBP
Advanced Models:

1. Artificial Neural Network + HoG/LBP
2. CNN

3. Experimental CNN with transfer learning - UNet, ResNet

5 Strategy for model selection and tuning hyper-parameters

K-fold cross validation, Mini Batch Gradient descent, ADAM Optimization

6 Training approach(es) to be explored

- Mini-batch Gradient descent
- Stochastic Gradient Descent Momentum(SGDM)

7 Ensemble approaches(None)

8 Evaluation metrics and Error Analysis approaches.

F1 Score, Accuracy, ROC-AUC curve, Confusion matrix

9 Deliverable of individual team members, described as clearly as possible.

- Lavanya: Evaluation Metric for CNN model; Feature extraction for baseline
- Aditya : Evaluation Metric for Transfer learning in CNN from UNet/ ResNet; Pre-processing of data
- Prabhat: Evaluation Metric for ANN model; Baseline implementation