

## Trabalho Final em Grupo – CNN para classificação de Raio X

Grupos: 4-5 alunos

Nota: 70%

### Descrição

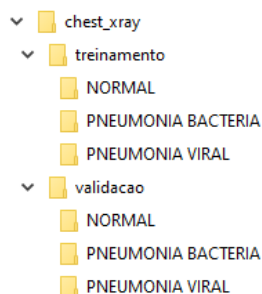
O objetivo do trabalho é criar uma rede neural convolucional que seja capaz de classificar imagens de Raio X em 'normal', 'pneumonia bacteriana' ou 'pneumonia viral', conforme exemplo abaixo:



Os dados são de crianças de 1 a 5 anos de Guangzhou (China) e podem ser acessados através de:

<https://drive.google.com/file/d/1c0QYUUwwgdbc1KNeMaQcojEQe7Rw25xJ/view?usp=sharing>

Os alunos devem treinar uma CNN usando as imagens contidas na pasta “treinamento” e aplicar o modelo nas imagens contidas em “validação”, para cálculo da performance do modelo:



Além da classificação, cada grupo deve escolher pelo menos 1 dos seguintes pontos abaixo para abordar no trabalho:

- 1) Use data augmentation para criar mais dados para o treino
- 2) Baseado no output ou neurônios das redes treinadas, como vocês criariam um índice para indicar o grau de gravidade da pneumonia?
- 3) Resolva este problema utilizando transfer learning
- 4) Utilize ferramentas de 'Explainable AI' para explicar quais as partes da figura foram mais relevantes para a classificação (ex. <https://github.com/slundberg/shap>)

O trabalho final deve conter:

- Notebook com o código utilizado
- Explicação das escolhas de arquitetura da rede e tratamentos dos dados utilizados
- Performance final do modelo multiclasse. Além da matriz de confusão final, para cada classe (normal, bactéria, viral), fornecer Precision, Recall e F1-score.

### **Observações**

- São critérios de avaliação:
  - Processo de resolução do problema
  - Arquitetura e tratamentos realizados, juntamente com as explicações
  - Organização do código
  - Performance final, avaliada no set de validação
- Este pode ser um trabalho cujo período de treinamento é demorado e intensivo de recurso computacional. Não deixe para a última hora!
- Trabalho é baseado em um caso público do Kaggle (plataforma de compartilhamento de desafios/ dados de machine learning). O aluno é incentivado a pesquisar soluções de problemas similares na internet, mas não deve copiar códigos prontos!
- Busque alternativas para facilitar o cálculo. Por exemplo, o Google Colab oferece alternativas de notebooks que utilizam GPUs como processamento, o que pode agilizar muito o treinamento da rede.