Pneumonia Detection

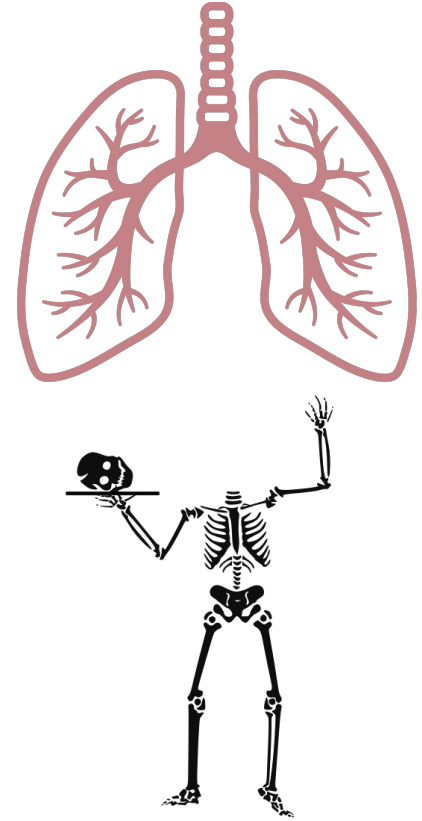By Armaan, Aarushi, Ella, Laeina, Aryan

What is Pneumonia?

- It is in your lungs!

- Small air sacs fill with fluid

- Can cause cough, difficulty breathing, etc.

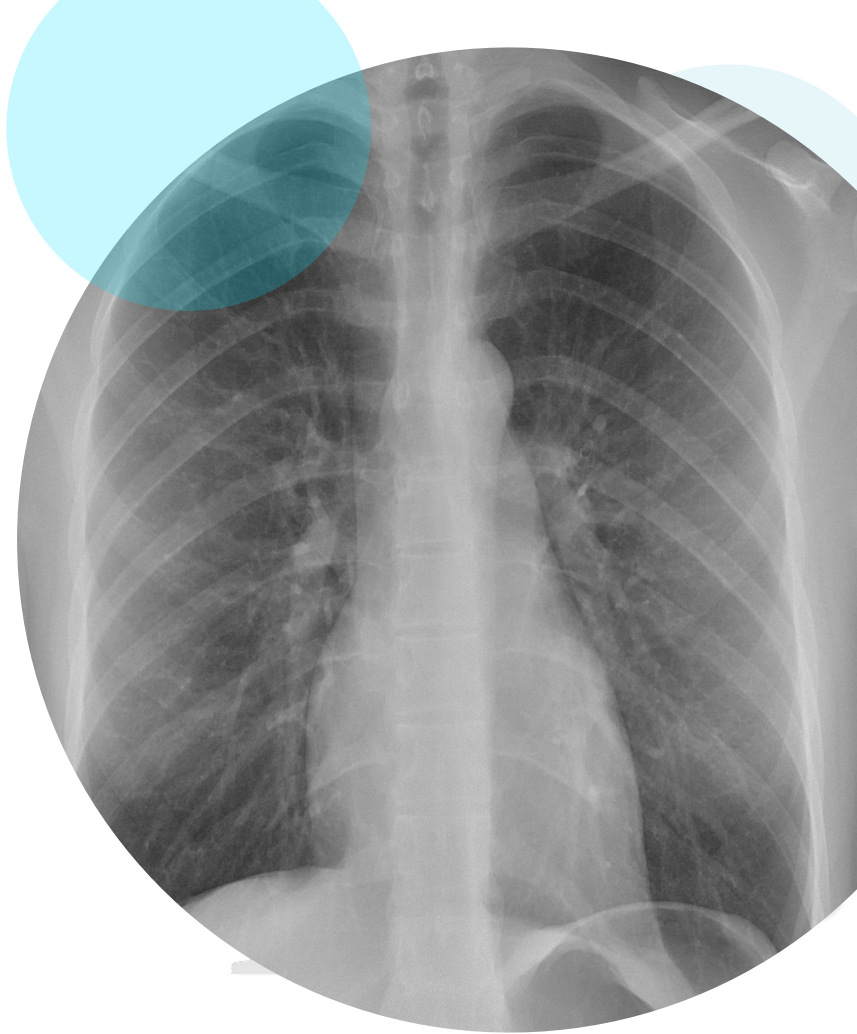- More commonly found in younger and older generations

But how can we detect this?   X-RAYS!

How do X-Rays Help?

By using a series of X-Rays, doctors can see our lungs, blood vessels, and heart in order to determine if u have pneumonia or not.
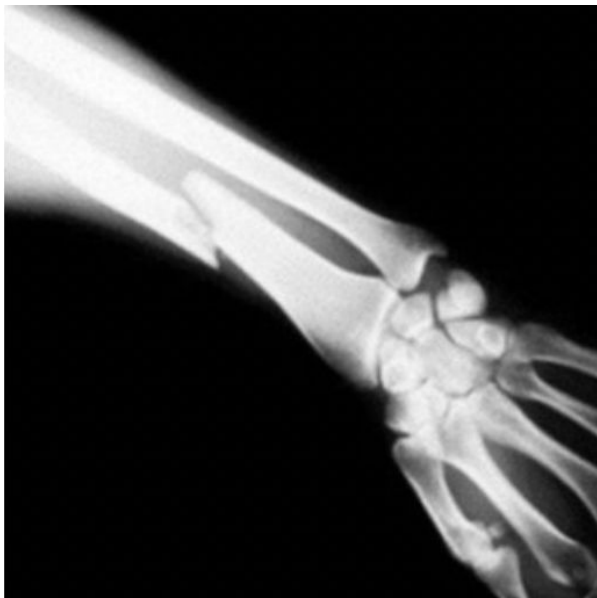
We can use AI!

How can we use AI?

By classifying our data, grouping it in different categories, we can give that data to a certain program

	class	split	index
0	0.0	train	0
1	0.0	train	1
2	1.0	train	2
3	0.0	train	3
4	1.0	train	4
...	...	...	...
2395	1.0	test	2395
2396	0.0	test	2396
2397	0.0	test	2397
2398	1.0	test	2398
2399	0.0	test	2399

Models/Implementations

Training vs. Testing Data

Training

- Initial Dataset
- Recognize patterns
- Perform the criteria

Testing

- Validation data
- Evaluate the model's accuracy

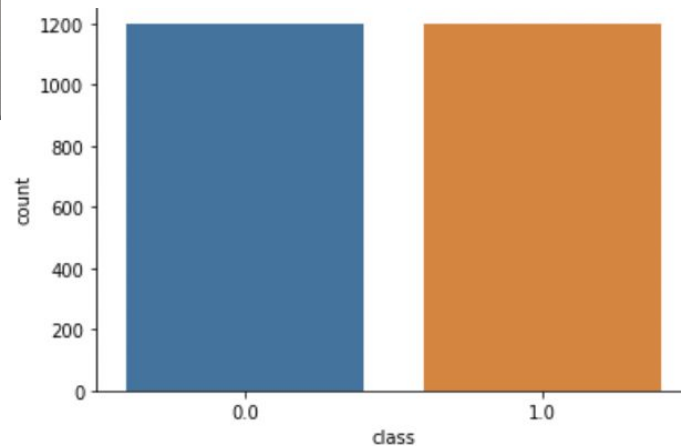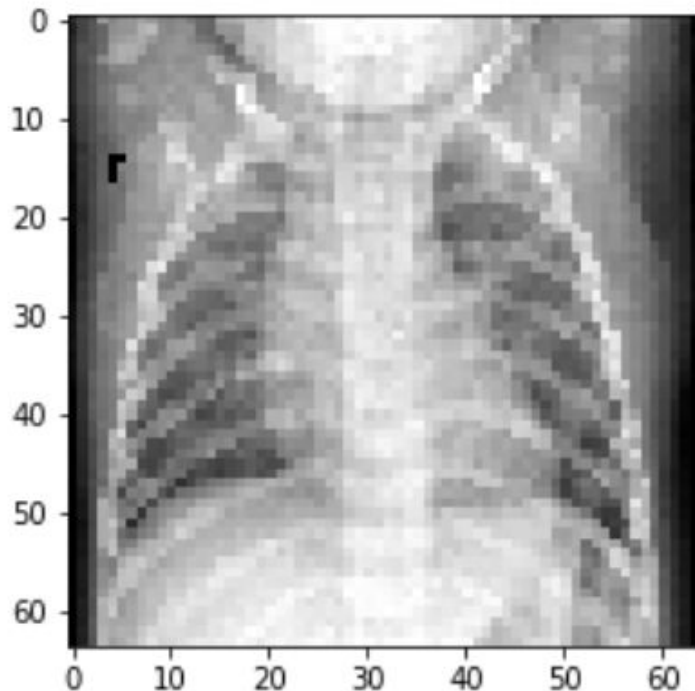|	class	split	index
0	0.0	train	0
1	0.0	train	1
2	1.0	train	2
3	0.0	train	3
4	1.0	train	4
...	...	...	...
2395	1.0	test	2395
2396	0.0	test	2396
2397	0.0	test	2397
2398	1.0	test	2398
2399	0.0	test	2399

2400 rows × 3 columns

Looking at the results/ Using KNN
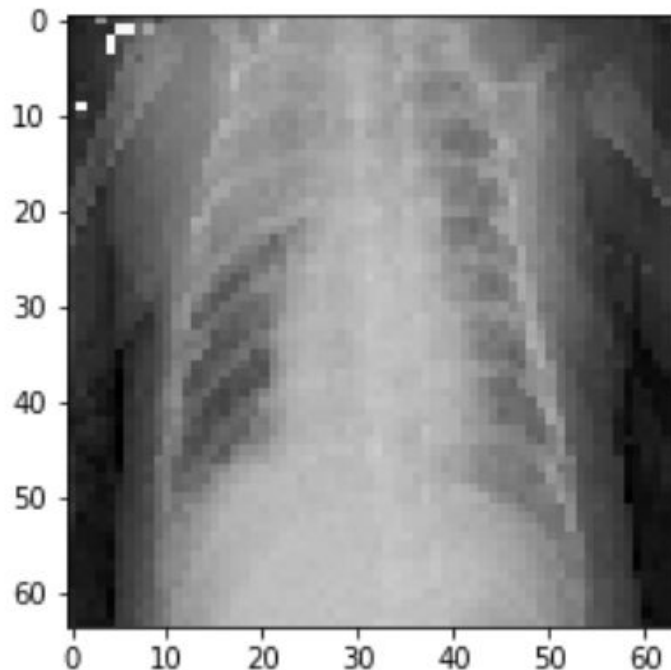
What is KNN?
K nearest neighbors

Label: 0.0

Label: 1.0

Neural Network - CNN


Neural Network

CNN

- Computer Vision: trains computers to interpret and understand the visual world
- Neural Network: precise tool for computer vision
- Input = Information/Data you give the computer
- Hidden = training weights and bias
- Output = classification or regression
- CNN: convolutional neural network, subclass of NN
- input, **convolutional layer**, output
- computer vision problems
- key difference: convolutional layers
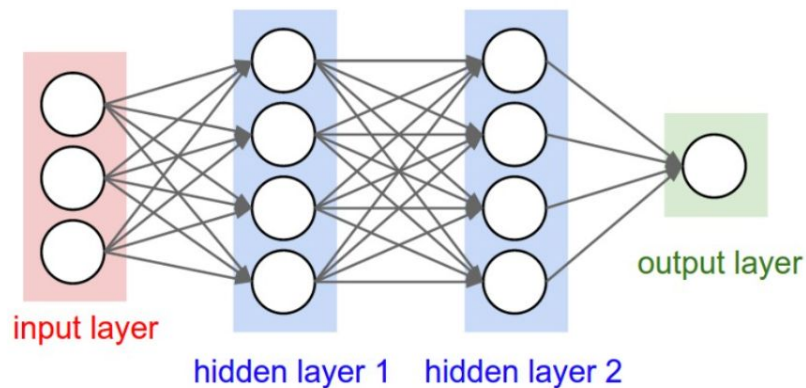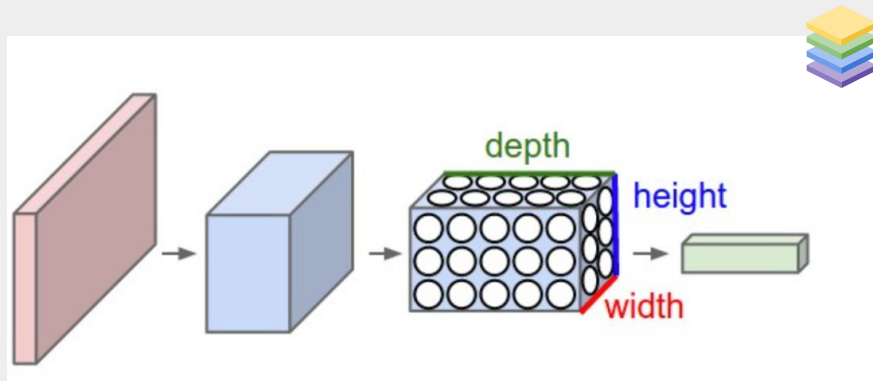
FIELD DATA

What is field data?

data outside our normal testing and training data! We'll also learn techniques for data augmentation, creating new "fake data" so that our model can generalize more effectively. (images taken from other x rays)
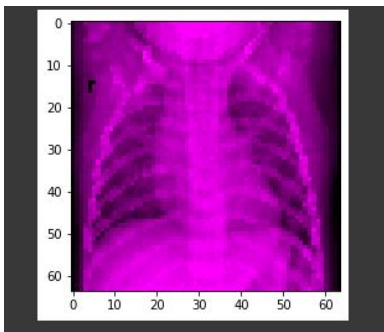
FIELD DATA

How is it different from our other data?

Field data is different because it is essentially just data that is not the same as testing and training. As an example it could be augmented or from a different machine.

←Testing/training data

Field data →

FIELD DATA

We implemented it by using the field data in our original model and finding the accuracy. We then augmented the data and adjusted our model to give us better results. In real life field data matters because there is always some variation in real data and not all of them are the same type.

How did we implement it? Why does it matter?

```
1 train_data, train_labels = get_train_data()
2 test_data, test_labels   = get_test_data()
3 field_data, field_labels  = get_field_data()
4
5 ### YOUR CODE HERE
6 cnn = CNNClassifier(num_hidden_layers = 1)
7 model_history = cnn.fit(all_data, all_labels, epochs = 100, validation_data = (test_data, test_labels), shuffle = True, callbacks =[monitor])
8
9
10 ### END CODE
```

FIELD DATA

By augmenting our data, we have gotten our models working better, even on messy field data! The accuracy of the model on field data came out to approximately 64% but varies between each model analyzed and how the model was altered

What were the results

```
[ ]   1 predictions = (cnn.predict(field_data) > 0.5).astype("int32")
      2 accuracy = accuracy_score(field_labels, predictions)
      3 print(accuracy)

      0.64
```

Ethics in AI + Our Model

- AI ethics consist of moral principles + techniques intended to help responsible use of artificial intelligence technology

- AI code of ethics, also called an AI value platform, is policy statement that formally defines role of artificial intelligence as it applies to continued development of human race

Why is Accuracy Important in AI + Our Model?

- Machine learning model accuracy is measurement used to determine which model is best at identifying relationships + patterns between variables in dataset based on input, or training, data

- Depending on what the model is for, the accuracy is going to hold different weights
  - Ie. Cancer Detection vs. Which Ad you will click on

- Any guesses on what our accuracy was?

THANKS!

Any questions?