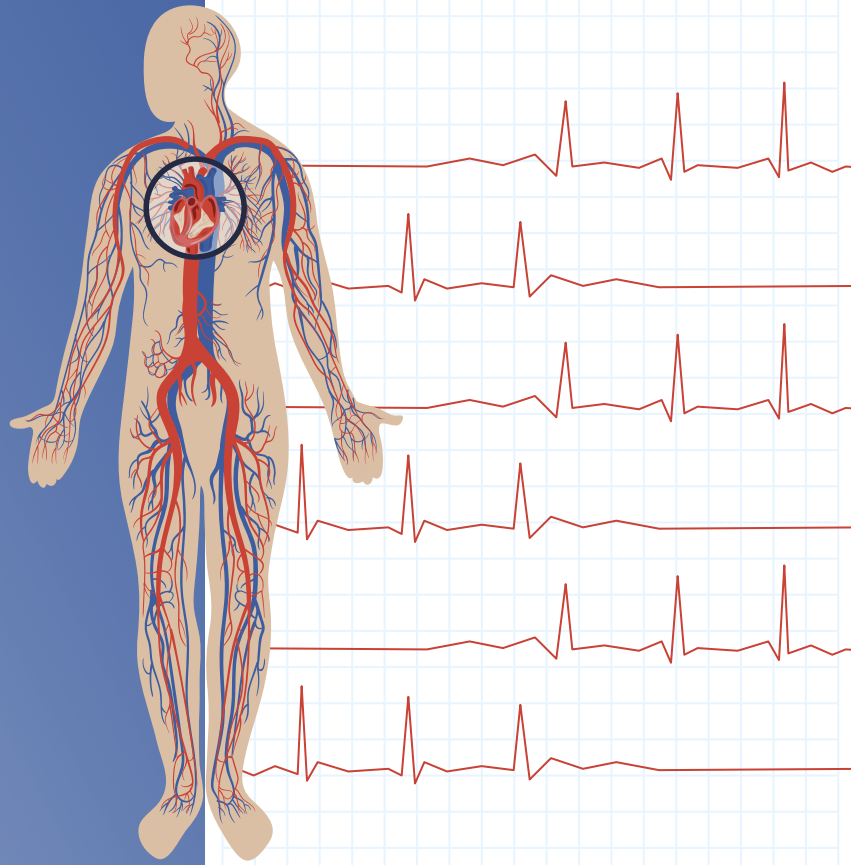
12 Lead EKG Image Classifier

By: Jose Gonzalez, Shannon Williams, Nancy Ulloa, and Arle Alcid

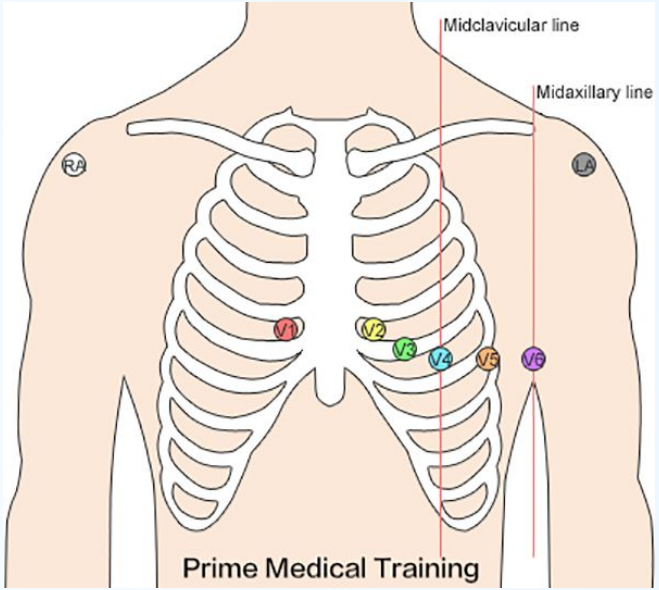


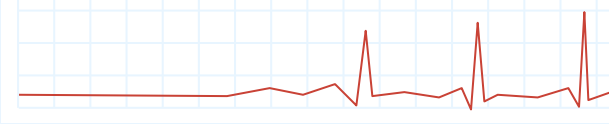
Purpose

The purpose of our project was to create a machine learning model that could classify EKGs as a healthy or arrhythmic heart rhythm

Introduction

- 12 Lead Electrocardiogram (EKG)





Data Retrieval

- PTB-XL
  - A large publicly available dataset
  - Over 21,700 test subjects
    - (https://physionet.org/content/ptb-xl/1.0.3/#files-panel)

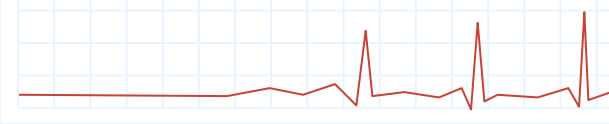
Folder Navigation: <base>

Name
records100
records500
LICENSE.txt
RECORDS
SHA256SUMS.txt
example_physionet.py
ptbxl_database.csv
ptbxl_v102_changelog.txt
ptbxl_v103_changelog.txt
scp_statements.csv

SCP-ECG Statement Description
normal ECG
inferior myocardial infarction
anteroseptal myocardial infarction
left ventricular hypertrophy

Folder Navigation: <base> / records100 / 01000

Name
Parent Directory
01000_lr.dat
01000_lr.hea
01001_lr.dat
01001_lr.hea
01002_lr.dat
01002_lr.hea



Data Cleaning

1st

Imported data to MongoDB

2nd

Exported database to our notebooks

3rd

Cleaned DataFrame from NANs/empty columns

4th

Separated “SCP_Codes” Column and kept the “NORM” or “ABNORM”

5th

For loop matching .hea to file names in csv

6th

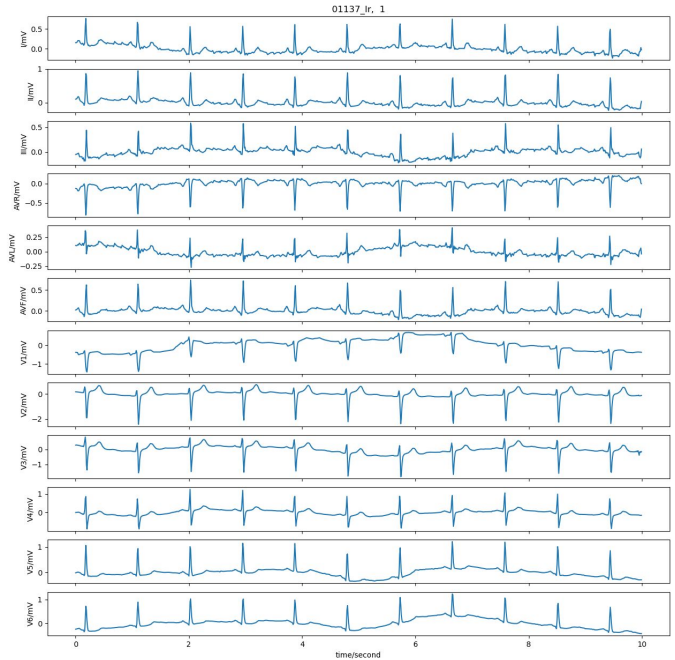
Waveform Database (wfdb) package to make EKGs

7th

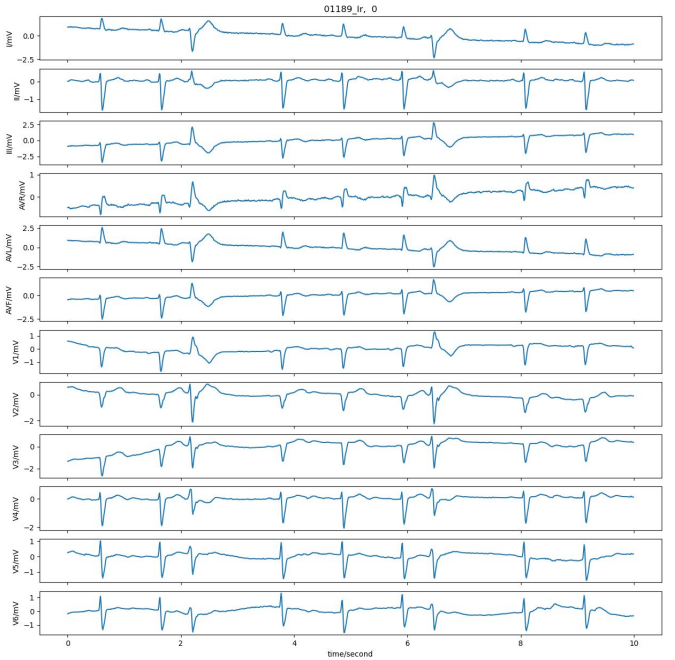
Saved 12 lead EKGs to corresponding folders

EKG Images

Normal



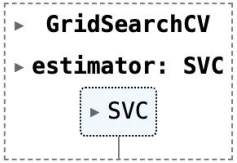
Abnormal



Scikit-Learn

- Limit size of model
  - Training size = 10%
  - Testing size = 10%
  - Shuffle
  - Stratify
- Training Classifier
  - Gamma (0.01, 0.001, 0.0001)
  - C (1, 10, 100)
- Test Performance
  - 52.27% of the samples were classified correctly

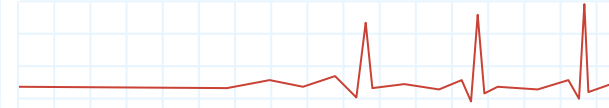
```
# Train Classifier
classifier = SVC()

parameters = [{'gamma': [0.01, 0.001, 0.0001],
               'C': [1, 10, 100, 1000]}]

grid_search = GridSearchCV(classifier, parameters)

grid_search.fit(X_train, y_train)

```




```
print(best_estimator)
```

```
SVC(C=1, gamma=0.01)
```

```
52.27272727272727% of samples were correctly classifed
```



TensorFlow

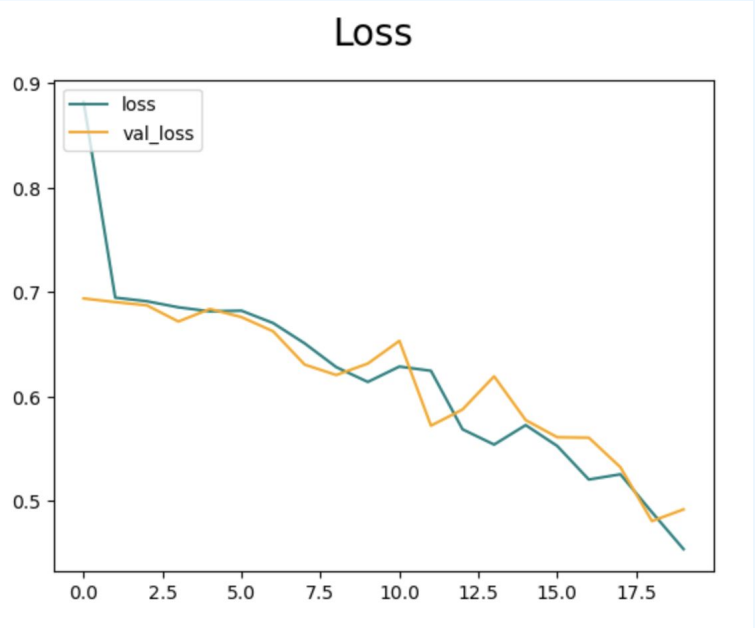
On our second attempt we tried to create a model using tensorflow.The goal with this model was to get something that could actually read the images and classify them. This model proved to be inaccurate but it did allow us to get a better understanding of how the image processing model works for our optimizing face

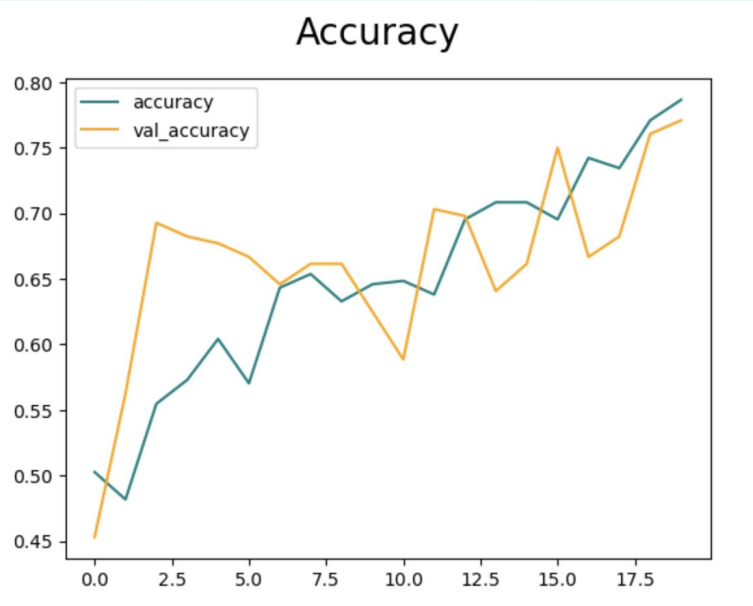
```
train_ds = tf.keras.utils.image_dataset_from_directory(
  data_dir,
  validation_split=0.2,
  subset="training",
  seed=123,
  image_size=(img_height, img_width),
  batch_size=batch_size)
```

```
Found 434 files belonging to 2 classes.
Using 348 files for training.
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 rescaling_1 (Rescaling)     (None, 300, 300, 3)       0

 conv2d (Conv2D)             (None, 298, 298, 32)      896

 max_pooling2d (MaxPooling2  (None, 149, 149, 32)      0
 D)

 conv2d_1 (Conv2D)           (None, 147, 147, 32)      9248

 max_pooling2d_1 (MaxPoolin  (None, 73, 73, 32)        0
 g2D)

 conv2d_2 (Conv2D)           (None, 71, 71, 32)        9248

 max_pooling2d_2 (MaxPoolin  (None, 35, 35, 32)        0
 g2D)

 flatten (Flatten)           (None, 39200)             0

 dense (Dense)               (None, 128)               5017728

...
Total params: 5037378 (19.22 MB)
Trainable params: 5037378 (19.22 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Optimization Model






```
print(pre.result(), re.result(), acc.result())
```

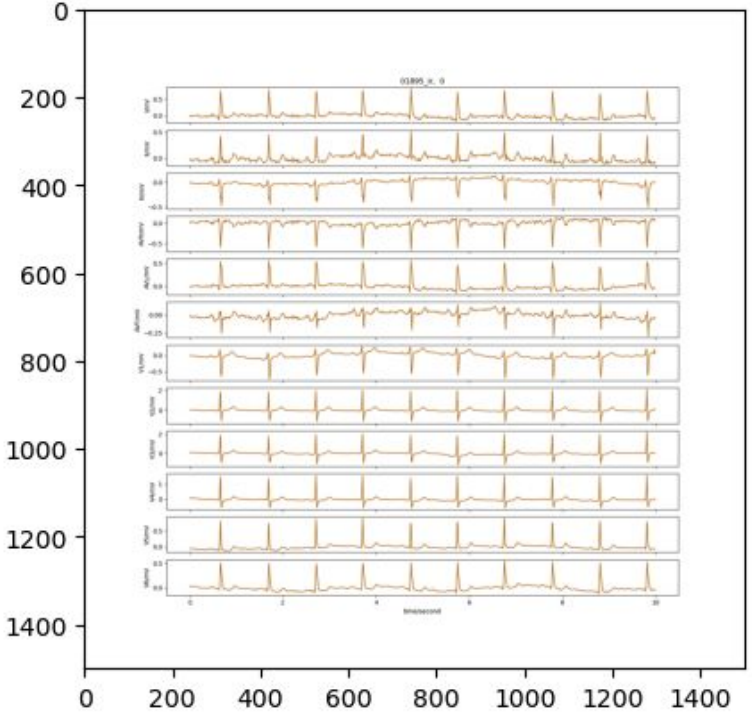
```
tf.Tensor(0.72727275, shape=(), dtype=float32) tf.Tensor(0.6956522, shape=(), dtype=float32) tf.Tensor(0.7291667, shape=(),
dtype=float32)
```

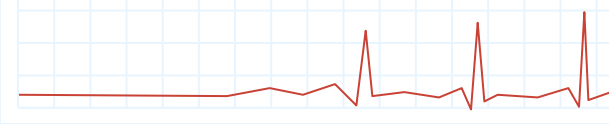
Optimization Model Cont.

```
image_path = 'test_images/test_abnorm/01895_lr.png'
image_loaded = tf.keras.utils.load_img(image_path)
# Printing the obtained image.
image_loaded

img = cv2.imread(image_path)
plt.imshow(img)
plt.show()

```

```
if yhat > 0.5:
    print(f'Predicted class is norm')
else:
    print(f'Predicted class is abnorm')
```

```
Predicted class is abnorm
```





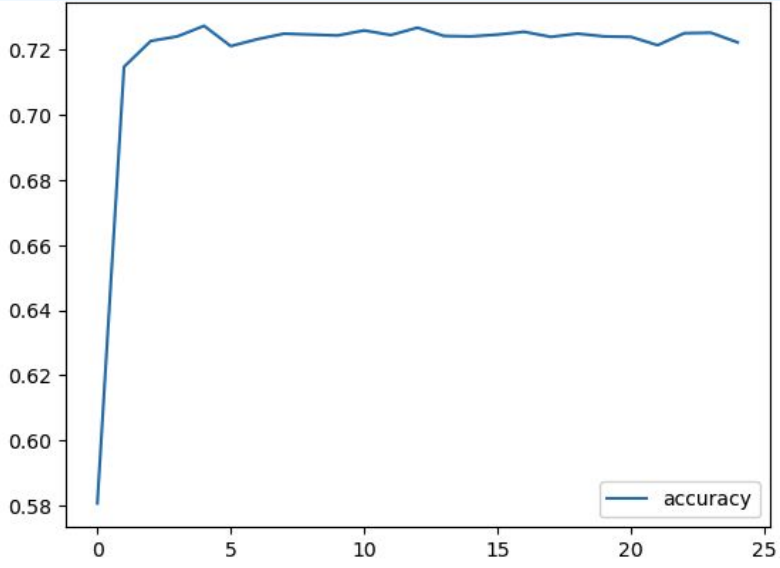
Neural Network Model

- A neural network algorithm was created in order to classify a normal or arrhythmic EKG, based off the patient's personal record
- Patient’s most relevant demographics were kept for the training model

	height	age	sex	weight	scp_codes_NORM
100	172.0	48.0	1	72.0	True
115	174.0	49.0	1	74.0	False
135	167.0	52.0	0	72.0	True

Continued...

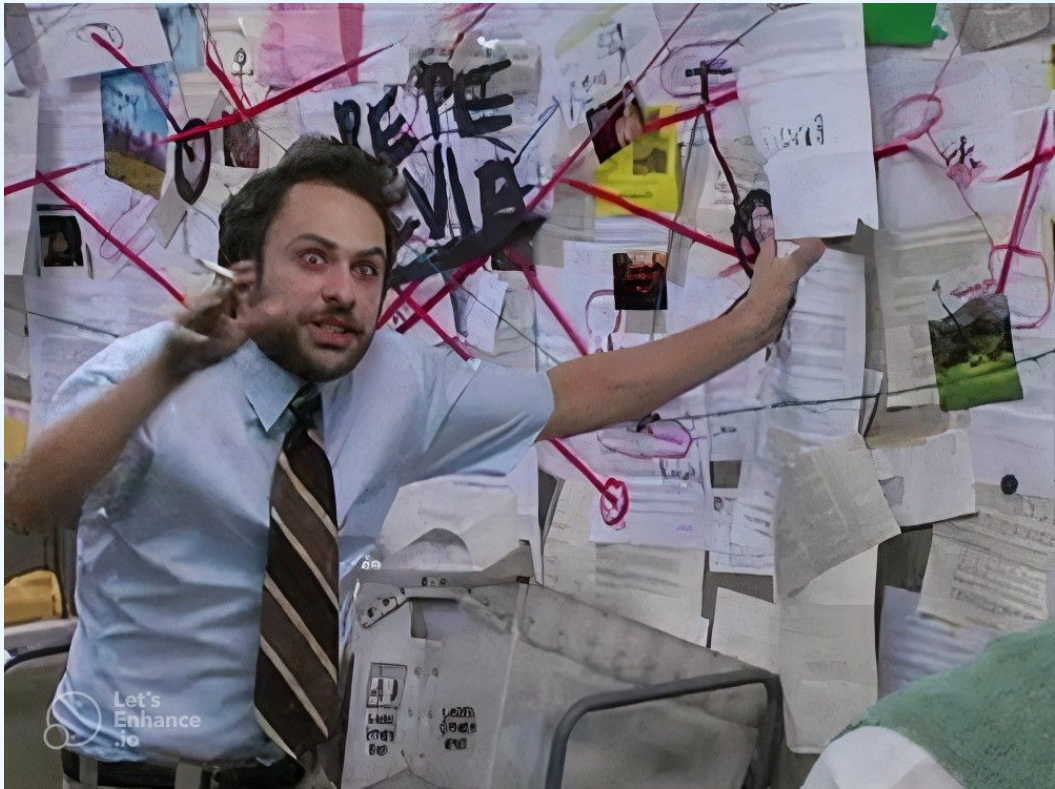
- 3 hidden layers
- 15, 10, 5 nodes per layer
- The model had a 89.7% accuracy rate

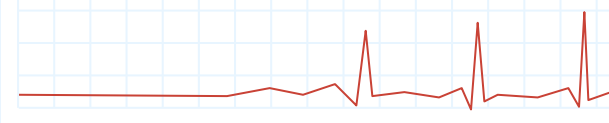


Limitations

- RAM – Our laptops melted
- To get the program to run we had to downscale the images and data might have been lost there.
- Lots of nuances throughout our abnormal classification

Questions?





Resources

- Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.
- Wagner, Patrick, et al. "PTB-XL, a large publicly available electrocardiography dataset" (version 1.0.3). *PhysioNet* (2022), https://doi.org/10.13026/kfzx-aw45.
- Wagner, P., Strodthoff, N., Bousseljot, R.-D., Kreiseler, D., Lunze, F.I., Samek, W., Schaeffter, T. (2020), PTB-XL: A Large Publicly Available ECG Dataset. Scientific Data. https://doi.org/10.1038/s41597-020-0495-6

Medical Advice Disclaimer

DISCLAIMER: THIS WEBSITE DOES NOT PROVIDE MEDICAL ADVICE
The information, including but not limited to, text, graphics, images and other material contained on this website are for informational purposes only. No material on this site is intended to be a substitute for professional medical advice, diagnosis or treatment. Always seek the advice of your physician or other qualified health care provider with any questions you may have regarding a medical condition or treatment and before undertaking a new health care regimen, and never disregard professional medical advice or delay in seeking it because of something you have read on this website.