Medical AI Wizards Milestone 2

Ethan Sandoval, Kevin Chen, Jonah Levine

Topic: Identifying and Classifying Skin Cancer

Project Recap

- To classify images of skin cancers as benign or malignant

Baseline Models

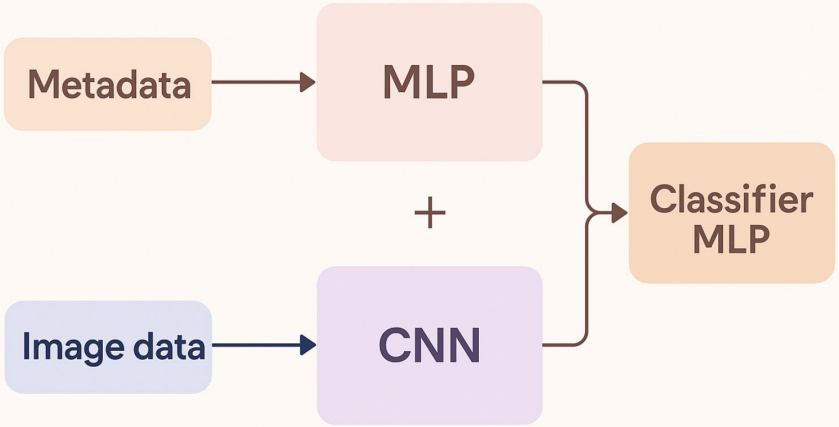
- We will implement 2 models
  - Majority classifier to compare accuracy
  - Linear regression model with just metadata to compare precision, recall, f1 score

Majority Classifier

- This would classify everything as benign
- There are 480700 benign images in the dataset and 18630 malignant images, so this simple model would already have 96% accuracy
- It would not have a precision or recall score since it would have no positives predicted, which is why we want another model

Linear Model

- We will train a linear regression model on just the metadata
- This should categorize some instances as malignant, so we will have a precision and recall



Deep Learning Model Architecture

Balancing Observations

- We have 96% benign observations, so it will be important to balance
- We will augment our malignant images to generate many more malignant observations
- If needed we can also remove some benign observations

Evaluation Methodology

Methods

- Accuracy

$$\frac{TP+TN}{TP+TN+FP+FN}$$

- Recall

$$\frac{TP}{TP+FN}$$

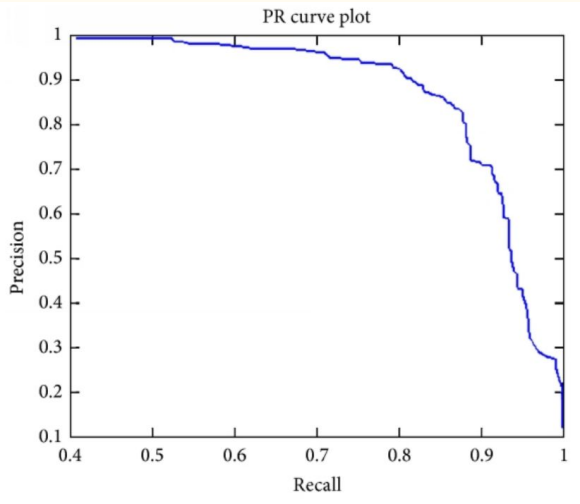
- Precision

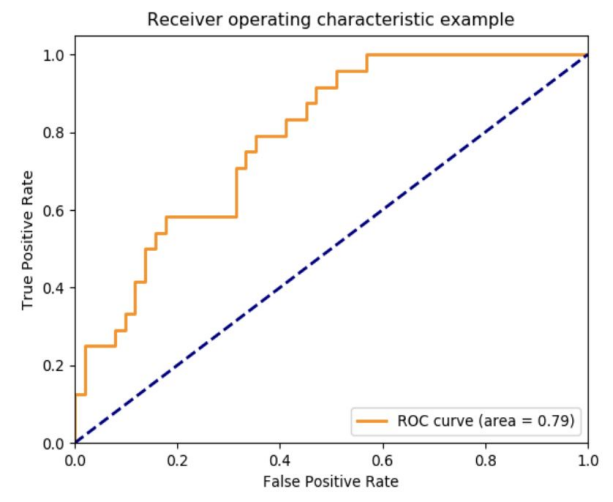
$$\frac{TP}{TP+FP}$$

- F1

$$\frac{2TP}{2TP+FP+FN}$$

Visualizations for Evaluation





Disease		Y_predicted	
	Condition	Positive	Negative
Y_target	Positive	15 True Positive (TP)	3+8=11 False Negative (FN)
	Negative	4+9=13 False Positive (FP)	23+19+1+2=45 True Negative (TN)