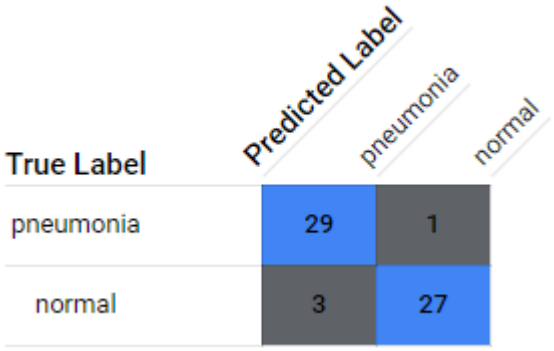
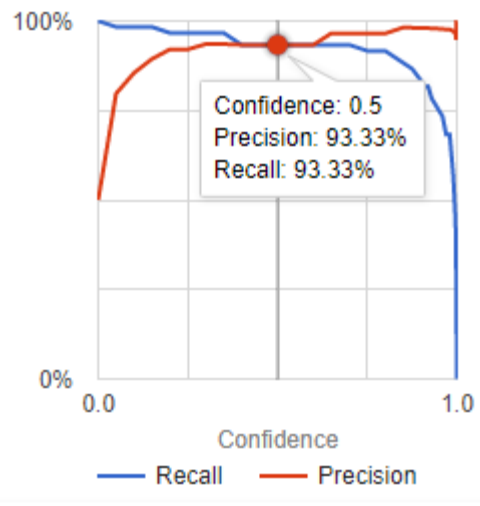
AutoML Modeling Report

Vinícius da Silva Vale

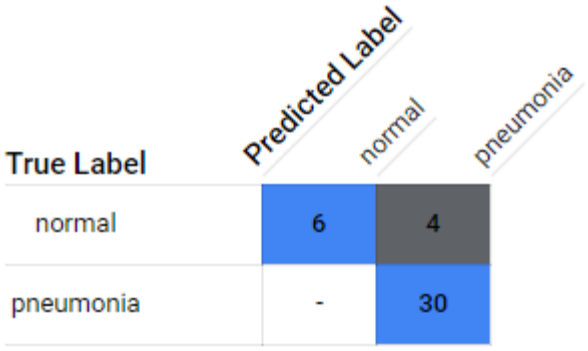
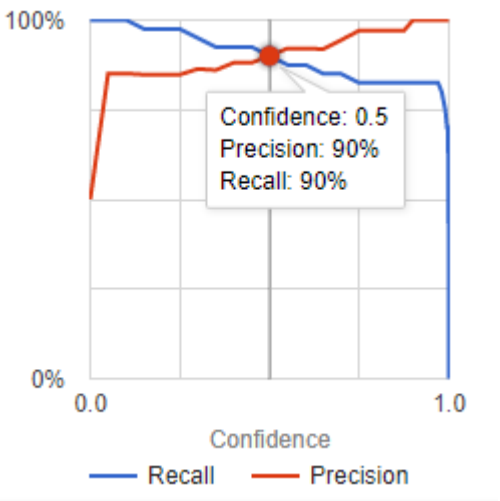
Binary Classifier with Clean/Balanced Data

Train/Test Split How much data was used for training? How much data was used for testing?	240 Normal images and 240 Pneumonia images for train. For test 30 normal and 30 Pneumonia
Confusion Matrix What do each of the cells in the confusion matrix describe? What values did you observe (include a screenshot)? What is the true positive rate for the “pneumonia” class? What is the false positive rate for the “normal” class?	This table shows how often the model classified each label correctly (in blue), and which labels were most often confused for that label (in gray).  97% is the true positive rate for the “pneumonia”, 3% of false positive, 10% of false negative and 90% of true negative. In the perspective of the normal class 10% is the false positive rate.
Precision and Recall What does precision measure? What does recall measure? What precision and recall did the model achieve (report the values for a score threshold of 0.5)?	Precision measure the frequency of predictions that were correct (positive). The higher the precision, the fewer false positives predicted. The higher the recall, the fewer false negatives, or the fewer predictions missed. The model score a Precision of 93.33% and a Recall of 93.33%.

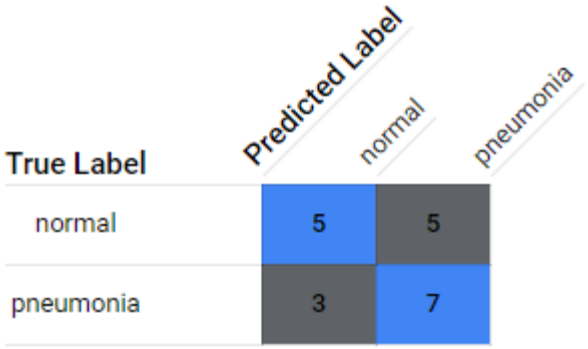
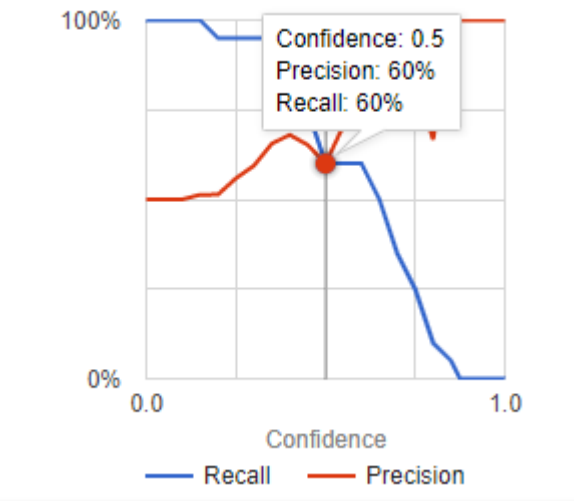
	
Score Threshold When you increase the threshold what happens to precision? What happens to recall? Why?	Precision increase and Recall decrease. Confidence threshold refers to the level of confidence the model needs to have to assign a category to a test item, higher means lower risk of classifying images incorrectly, increasing precision, but increase false negatives, since the model can classify fewer images, decreasing recall.

Binary Classifier with Clean/Unbalanced Data

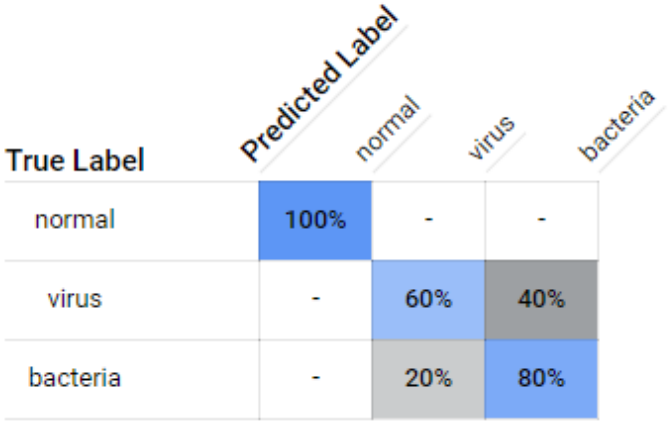
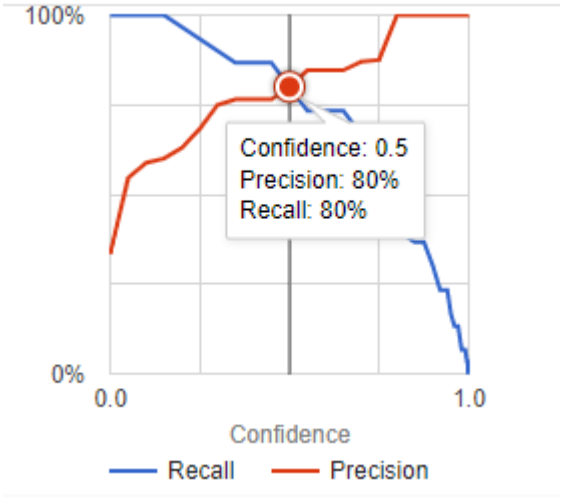
Train/Test Split How much data was used for training? How much data was used for testing?	 In normal class 80 for train, 10 forvalidation and 10 for test. In pneumonia class 240 for train, 30 forvalidation and 30 used for testing.

Confusion Matrix How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix.	 100% is the true positive rate for the “pneumonia”, 0% of false positive, 40% of false negative and 60% of true negative. The capacity for detecting cases of pneumonia increased, as did the amount of false negative
Precision and Recall How have the model’s precision and recall been affected by the unbalanced data (report the values for a score threshold of 0.5)?	 Both decreased in relation to the balanced model
Unbalanced Classes From what you have observed, how do unbalanced classed affect a machine learning model?	Favors the class with more cases

Binary Classifier with Dirty/Balanced Data

Confusion Matrix How has the confusion matrix been affected by the dirty data? Include a screenshot of the new confusion matrix.	 70% is the true positive rate for the “pneumonia”, 30% of false positive, 50% of false negative and 50 % of true negative. The ability to detect a normal lung has decreased and it is not much different from flipping a coin and the number of false positives has increased considerably
Precision and Recall How have the model’s precision and recall been affected by the dirty data (report the values for a score threshold of 0.5)? Of the binary classifiers, which has the highest precision? Which has the highest recall?	 Errors in test cases decreased the model's precision and recall. Using Clean / Balanced Data is the best option, followed by Clean / Unbalanced Data, Dirty / Balanced Data is not recommended
Dirty Data From what you have observed, how does dirty data affect a machine learning model?	Garbage in garbage out

3-Class Model

Confusion Matrix Summarize the 3-class confusion matrix. Which classes is the model most likely to confuse? Which class(es) is the model most likely to get right? Why might you do to try to remedy the model's "confusion"? Include a screenshot of the new confusion matrix.	
Precision and Recall What are the model's precision and recall? How are these values calculated (report the values for a score threshold of 0.5)?	 Precision 0.8 and Recall 0.8. It is an average between the precision values of each class, the same applies for recall
F1 Score What is this model's F1 score?	0.8