

# AutoML Modeling Report





Rosemary Nwosu-Ihueze

## Binary Classifier with Clean/Balanced Data

### Train/Test Split

How much data was used for training? How much data was used for testing?

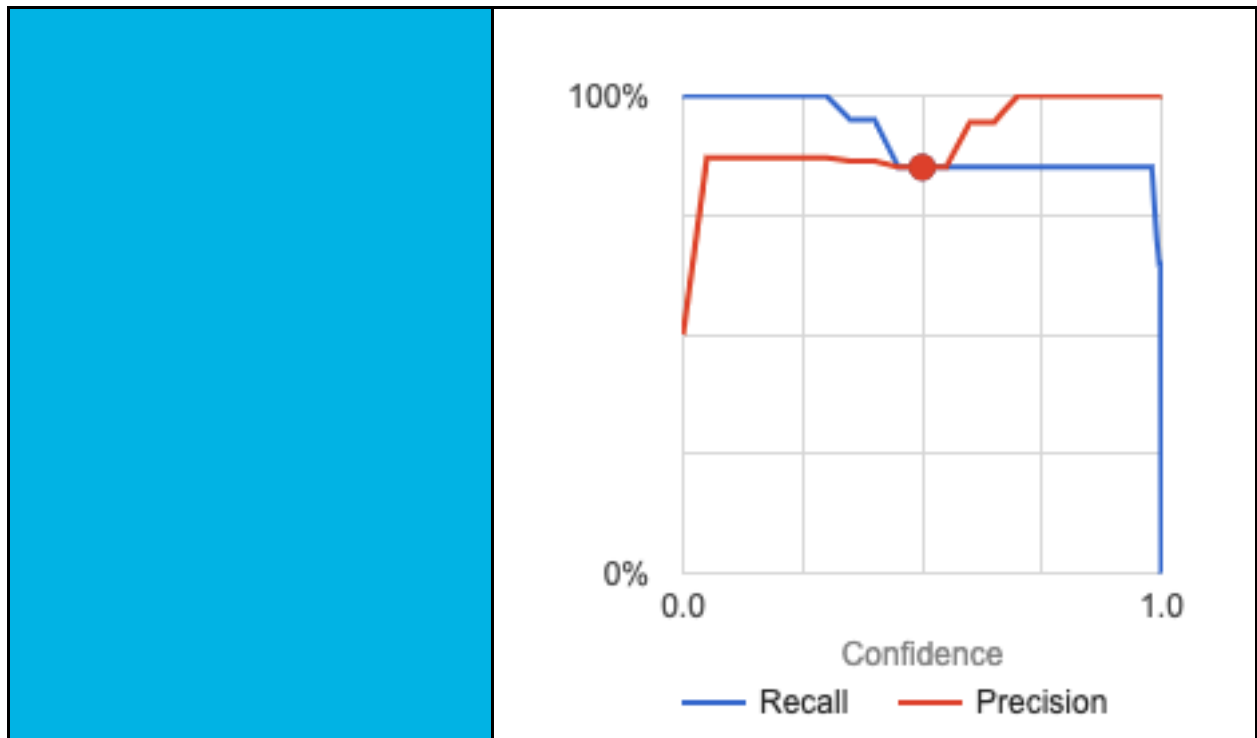
Labels	Images	
normal		100
pneumonia		100
Train	Validation	Test
80	10	10
80	10	10

### 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?

True positives are the correct classification of pneumonia while false positive is the incorrect classification of pneumonia. In the confusion matrix, it shows the model was better at classifying pneumonia diseased lungs than healthy lungs.

	<table><tr><th rowspan="2">True Label</th><th colspan="2">Predicted Label</th></tr><tr><th>normal</th><th>pneumonia</th></tr><tr><th>normal</th><td>70%</td><td>30%</td></tr><tr><th>pneumonia</th><td>-</td><td>100%</td></tr></table>	True Label	Predicted Label		normal	pneumonia	normal	70%	30%	pneumonia	-	100%
True Label	Predicted Label											
	normal	pneumonia										
normal	70%	30%										
pneumonia	-	100%										
<p><b>Precision and Recall</b></p> <p>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)?</p>	<p>Precision measures the ratio of true positives to the predicted values.</p> <p>Recall measures the ratio of true positives to actual positives.</p> <p>Binary_20210705115844</p> <div><div></div><div><p>Average precision ?</p><p>0.98</p><p>Precision* ? 85%</p><p>Recall* ? 85%</p><p>* Using a score threshold of 0.5</p></div></div>											
<p><b>Score Threshold</b></p> <p>When you increase the threshold what happens to precision? What happens to recall? Why?</p>	<p>Increase in threshold leads to increase in precision as is approaches 1 while the recall approaches 0.</p>											



## Binary Classifier with Clean/Unbalanced Data

### Train/Test Split

How much data was used for training? How much data was used for testing?

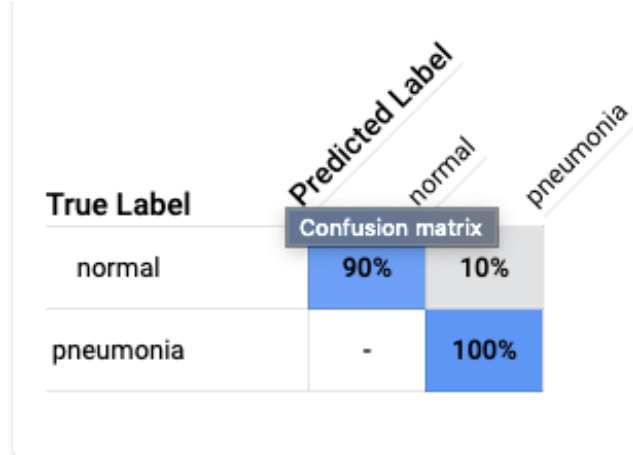
Labels	Images		
normal	100		
pneumonia	299		
Train	Validation	Test	
80	10	10	
239	30	30	

### Confusion Matrix

It improved the performance level of the model in

How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix.

classifying the healthy lungs in the confusion matrix with no effect on the pneumonia classification.



A confusion matrix for pneumonia classification. The y-axis is labeled 'True Label' with categories 'normal' and 'pneumonia'. The x-axis is labeled 'Predicted Label' with categories 'normal' and 'pneumonia'. The matrix shows that 90% of 'normal' cases were correctly classified, 10% were misclassified as 'pneumonia', and 100% of 'pneumonia' cases were correctly classified. A blue box labeled 'Confusion matrix' is overlaid on the table.

True Label	Predicted Label	
	normal	pneumonia
normal	90%	10%
pneumonia	-	100%

### 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)?

There was significant improvement on the precision and recall of the model.

Unbalanced\_20210706124043



Average precision ?

0.996

Precision\* ?

97.5%

Recall\* ?

97.5%

\* Using a score threshold of 0.5

### Unbalanced Classes

From what you have observed, how do unbalanced classes affect a machine learning model?

Unbalanced data introduces bias which impacts the accuracy of the model. With the increase in accuracy, the recall on both didn't change.



## 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.

The model performed poorly as seen in the diagram.

True Label	Predicted Label	
	pneumonia	normal
pneumonia	30%	70%
normal	40%	60%

### 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?

The precision and recall values went to a very low value of 47.7%. The highest precision and recall values were observed with clean and unbalanced data.

Dirty\_20210706011459



Average precision ?

0.477

Precision\* ?

45%

Recall\* ?

45%

\* Using a score threshold of 0.5

### Dirty Data

From what you have observed, how does dirty data affect a machine learning model?

The model clearly struggled to find patterns among classes as the labels/classes are mixed up. Model sees same patterns in both labels and hence performs poorly.

## 3-Class Model

### Confusion Matrix

Summarize the 3-class confusion

By intuition, the bacterial and viral classes are likely to be confused. But the model was able to classify the viral

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.

pneumonia class accurately with false positives in bacterial class and the normal class.

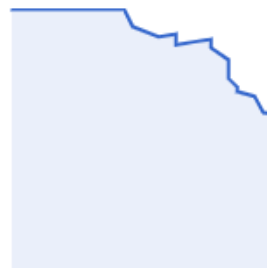
To remedy this, I would change the labelling to exclude the pneumonia metadata.

True Label	Predicted Label		
	bacterial	viral	normal
bacterial	60%	40%	-
viral	-	100%	-
normal	10%	10%	80%

### 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)?

Class\_20210706015017



Average precision ?

**0.905**

Precision\* ?

85.19%

Recall\* ?

76.67%

\* Using a score threshold of 0.5

### F1 Score

What is this model's F1 score?

F1 = 80.701%