

# AutoML Modeling Report



Magido Mascate

## Binary Classifier with Clean/Balanced Data

### Train/Test Split

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

There were used 180 for Training and 20 for Testing (10 in each class – normal and pneumonia, respectively).

Total Images	180
Test Items	20
Precision	100%
Recall	100%

### 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?

The cells in confusion matrix (CM) describe the relation between the predicted and current classification of chest x-ray images in two classes (normal and pneumonia). From CM we found 10 out of 20 chest images were correctly classified as Normal and other 10 as Pneumonia.

True Label	Predicted Label	
	NORMAL	PNEUMONIA
NORMAL	10	-
PNEUMONIA	-	10

### 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 measures the percentage of correctly (True positive) predicted data out of Total (sum of correctly classified and wrongly classified data from other class, in this specific class). Recall measures the percentage of correct (True positive) prediction out of Total of data of same class.

The model achieved (In threshold of 0.5):

- Precision = 100%
- Recall = 100%

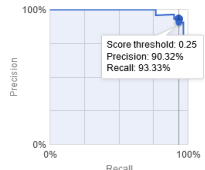
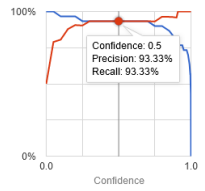


### Score Threshold

When you increase the threshold what happens to precision? What happens to recall? Why?

No change occurs when I increase the threshold, neither for precision nor recall. The percentages remain the same. It is due the balanced data and small dataset.

## Binary Classifier with Clean/Unbalanced Data

<h3>Train/Test Split</h3> <p>How much data was used for training? How much data was used for testing?</p>	<p>Training = 270 images Testing = 30 images</p> <table><tr><td>Total images</td><td>270</td></tr><tr><td>Test items</td><td>30</td></tr><tr><td>Precision ?</td><td>93.33%</td></tr><tr><td>Recall ?</td><td>93.33%</td></tr></table>	Total images	270	Test items	30	Precision ?	93.33%	Recall ?	93.33%			
Total images	270											
Test items	30											
Precision ?	93.33%											
Recall ?	93.33%											
<h3>Confusion Matrix</h3> <p>How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix.</p>	<p>The confusion matrix shows a drop of two (2) wrong prediction, one for normal class and one for pneumonia.</p> <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>9</td><td>1</td></tr><tr><th>PNEUMONIA</th><td>1</td><td>19</td></tr></table>	True Label	Predicted Label		NORMAL	PNEUMONIA	NORMAL	9	1	PNEUMONIA	1	19
True Label	Predicted Label											
	NORMAL	PNEUMONIA										
NORMAL	9	1										
PNEUMONIA	1	19										
<h3>Precision and Recall</h3> <p>How have the model's precision and recall been affected by the unbalanced data (report the values for a score threshold of 0.5)?</p>	<p>Both precision and recall present a small drop around (-6.7 to -9.7)%.</p> 											
<h3>Unbalanced Classes</h3> <p>From what you have observed, how do unbalanced classes affect a machine learning model?</p>	<p>Associated with the drop of model precision and recall, unbalanced dataset affects the model prediction confidence in terms of precision and recall by -6.7%.</p> 											

## 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.

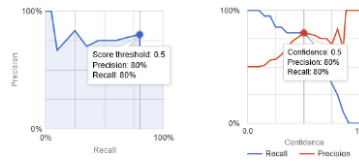
The confusion matrix was drastically affected as it shows an increase on the False Positives (wrong prediction).

True Label	Predicted Label	
	NORMAL	PNEUMONIA
NORMAL	8	2
PNEUMONIA	2	8

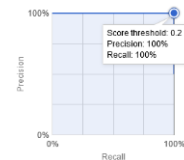
### Precision and Recall

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?

In this model we found precision and recall as equal to 80% respectively.



For all binary classifiers, the one with highest precision has highest recall too, account to 100%. And it is the face **Clean/Balanced Dataset (100 / 100)**.



### Dirty Data

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

It has dropped about -20% of precision and recall and affects the ML model accuracy.

## 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.

The model is likely to confuse the Bacterial and Viral class together. On-the-other-hand, it is likely to get right the Normal class.

True Label	Predicted Label		
	normal	viral	bacterial
normal	9	-	1
viral	-	8	2
bacterial	-	3	7

I might try to retrain the model with relatively big dataset and adopt different train/test split proportion.

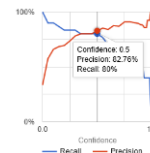
### Precision and Recall

What are the model's precision and recall? How are these values calculated (report the values for a score threshold of 0.5)?

The model's precision is 82.76% and recall is 80%.

- Precision (%)** =  $[(\text{normal prec.} + \text{bacterial prec.} + \text{viral prec.}) / 3] \times 100$
- Recall (%)** =  $[(\text{normal rec.} + \text{bacterial rec.} + \text{viral rec.}) / 3] \times 100$

$\text{Prec.} = (\text{True Positive}) / (\text{True Positive} + \text{False Positive})$   
 $\text{Rec.} = (\text{True Positive}) / (\text{True Positive} + \text{False Negative})$



### F1 Score

What is this model's F1 score?

F1 score = **0.851**