

# AutoML Modeling Report



<B. Akintade>

## Binary Classifier with Clean/Unbalanced Data

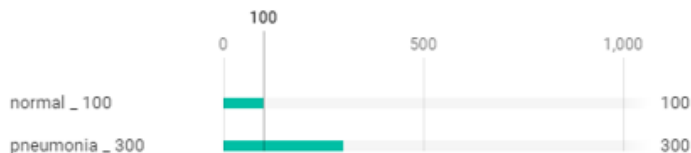
### Train/Test Split

317 data images used for training  
41 data images used for testing  
42 data images used for validation

IMAGES TRAIN EVALUATE PREDICT

**You have enough images to start training**

At least 100 images are currently assigned to each label. [Learn more](#)



Your images will be automatically split into [training and test sets](#), so you can evaluate your model's performance. Unlabeled images will not be used.

Training images	317
Validation images	42
Test images	41

### Confusion Matrix

The confusion matrix shows that 100% of the “normal” labelled class was correctly predicted/classified 100% of the time in the “normal” category during evaluation; and the “pneumonia” labelled class was classified correctly 94.1% times in the “pneumonia” category - this shows that for 5.9% of the evaluation time, 'normal' labelled images were often incorrectly predicted and confused with 'pneumonia' labels.

Effect of unbalanced data in the confusion matrix shows that both values in the blue diagonal (94.1% and 100.0%) is high compared to 5.9% confused/incorrect predictions, and the classifier is still identifying categories correctly to the extent indicated.

#### Confusion matrix

This table shows how often the model classified each label correctly (in blue), and which labels were most often confused for that label (in orange)

True label	Predicted label	
	pneumonia _ 300	normal _ 100
pneumonia _ 300	94.1%	5.9%
normal _ 100	-	100.0%

<b>Precision &amp; Recall</b>	<p>For a score threshold of 0.5, the unbalanced data caused the model to achieve the same precision and recall of 95.1% each (see image below) for every labelled 'pneumonia' and 'normal' images – which is reasonably high but not perfect (if 100%); and the model will produce fewer false positives and fewer false negatives.</p> <p>Score threshold ⓘ <span style="display: inline-block; width: 100px; border-bottom: 1px solid #ccc; position: relative; top: -5px;"><span style="position: absolute; left: 50%; transform: translate(-50%, -50%); background-color: #007bff; border-radius: 50%; width: 10px; height: 10px;"></span></span> 0.50</p> <table> <tr> <td>Total images</td><td>400</td></tr> <tr> <td>Precision ⓘ</td><td>95.1%</td></tr> <tr> <td>Recall ⓘ</td><td>95.1%</td></tr> </table>	Total images	400	Precision ⓘ	95.1%	Recall ⓘ	95.1%
Total images	400						
Precision ⓘ	95.1%						
Recall ⓘ	95.1%						
<b>Unbalanced Classes</b>	<p>My observation with model outcomes of unbalanced classes is that outcome can be biased and misleading; also accurate data balance and effective class distribution are not accounted for – this might be due to the fact that the main purpose of designing ML Algorithms is to improve accuracy by reducing (outcome/result) errors.</p>						

## EVALUATE

IMAGES
TRAIN
**EVALUATE**
PREDICT

---

**Model**  
AutoMLDataset\_2\_v20190922144428

**Created**  
Sep 22, 2019  
1 compute hour

**Analyzed**  
400 images  
2 labels, 41 test images

**Avg precision ⓘ**  
0.989

**Precision ⓘ**  
95.122%

**Recall ⓘ**  
95.122%

Precision and recall are based on a score threshold of 0.5

Type to filter labels...

### All labels

normal \_ 100

pneumonia \_ 300

## All labels

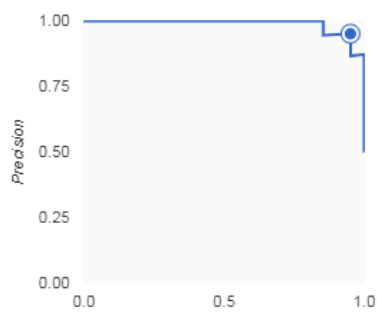
Score threshold ?  0.50

Total images 400

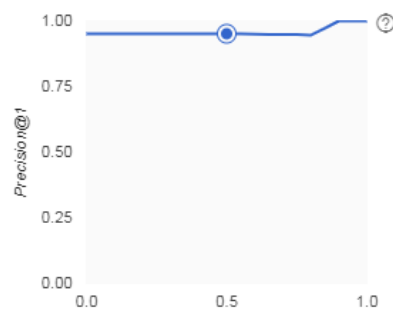
Precision ? 95.1%

Recall ? 95.1%

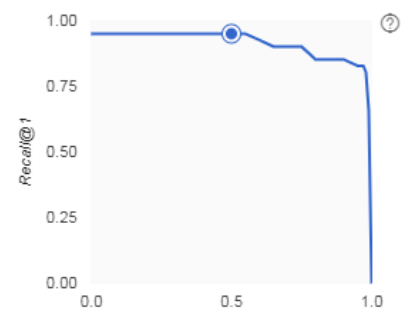
Use the slider to see which score threshold works best for your model on the precision-recall tradeoff curve. [Learn more about these metrics and graphs](#)



Recall



Score threshold



Score threshold

## Confusion matrix

This table shows how often the model classified each label correctly (in blue), and which labels were most often confused for that label (in orange).

True label	Predicted label	
	pneumonia _ 300	normal _ 100
pneumonia _ 300	94.1%	5.9%
normal _ 100	-	100.0%