

AutoML Modeling Report



AYŞE DEMİREL

Binary Classifier with Clean/Balanced Data

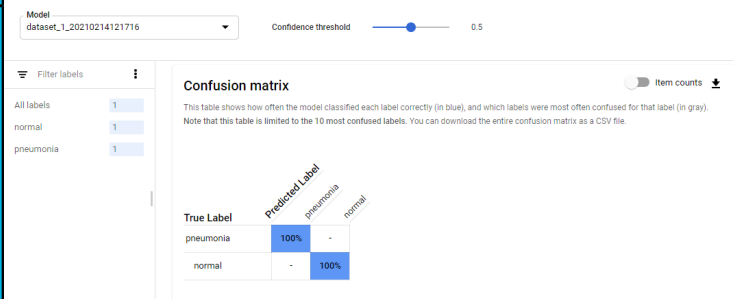
Train/Test Split

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

Totally, there was 200 images, 180 images were used for training and 20 images were used for testing.

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?



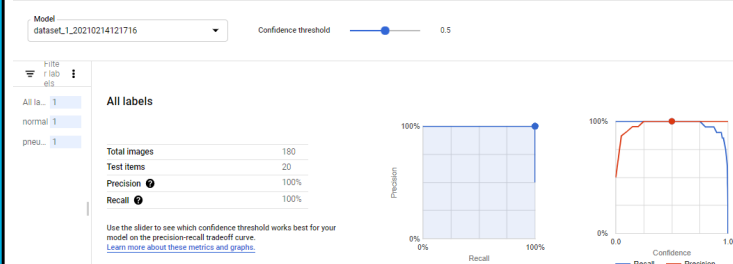
Label : Pneumonia, Normal

TP = %100

FP = %100

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 ratio of TP to predicted positives. Recall measures the ratio of TP to actual positives. Precision and recall values are %100 for the score threshold of 0.5

Score Threshold

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

When I increased the threshold until 1, the recall value decreased(%65) and the precision value unchanged. The recall value decreased because the expectation of the truth increased so the TP value decreased.

Binary Classifier with Clean/Unbalanced Data

Train/Test Split

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

Totally, there were 400 images, 100 images were normal label, 300 images were pneumonia label. 40 items used for testing and 360 items used for training.

Confusion Matrix

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

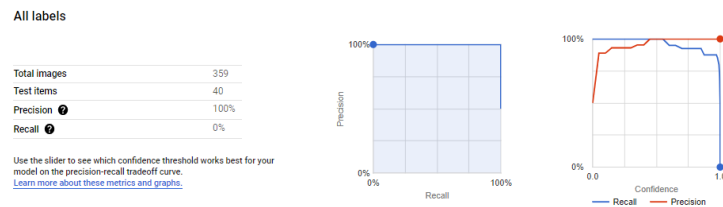
True Label	Predicted Label	
	pneumonia	normal
pneumonia	100%	-
normal	-	100%

True Label	Predicted Label	
	pneumonia	normal
pneumonia	30	-
normal	-	10

Not affected. Actually, it was an unexpected result for me. Test data is also unbalanced, so the result is same.

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)?



Same.

Unbalanced Classes

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

When threshold value increased, the recall value decreased. When threshold value is 1, recall value %0. Accuracy affected from unbalanced classes.

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.

True Label	Predicted Label	
	pneumonia	normal
pneumonia	80%	20%
normal	10%	90%

Totally, there were 200 images. 100 normal and 100 pneumonia. Switched the labels of 30 images in each class.

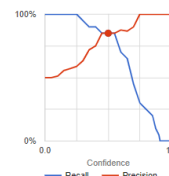
Precision and Recall

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?

All labels

Total Images	180
Test Items	20
Precision	85%
Recall	85%

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



Normal -> precision: 81,82% , recall: 90%

Pneumonia -> precision: 88,89% , recall: 80%

Highest precision: pneumonia

Highest recall: normal

Dirty Data

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

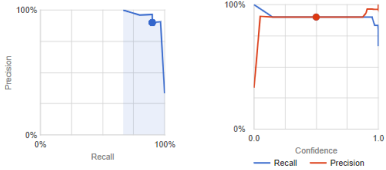
Precision and recall values decreased so accuracy of model decreased.

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.

True Label	Predicted Label		
	bacterial	normal	virus
bacterial	90%	10%	-
normal	-	90%	10%
virus	10%	-	90%

	<p>There were 100 "normal" images, 100 "bacterial pneumonia" images, and 100 "virus pneumonia" images. Bacterial and virus pneumonia classes are the model most likely to confuse. Normal class is the model likely to get right. Model needs more data to solve the confusion.</p>								
<p>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)?</p>	<p>All labels</p> <table><tr><td>Total Images</td><td>270</td></tr><tr><td>Test Items</td><td>30</td></tr><tr><td>Precision ①</td><td>90%</td></tr><tr><td>Recall ②</td><td>90%</td></tr></table> <p>Use the slider to see which confidence threshold works best for your model on the precision-recall tradeoff curve. Learn more about these metrics and graphs.</p>  <p>The figure contains two plots. The left plot is a Precision-Recall tradeoff curve with 'Recall' on the x-axis (0% to 100%) and 'Precision' on the y-axis (0% to 100%). A blue curve starts at (0,0) and rises to a point at approximately (90%, 90%) before dropping. The right plot is a Precision-Confidence tradeoff curve with 'Confidence' on the x-axis (0.0 to 1.0) and 'Precision' on the y-axis (0% to 100%). It shows two curves: a blue curve for Recall and a red curve for Precision. The Recall curve starts at (0,100%) and drops to (1,0%). The Precision curve starts at (0,0%), rises sharply to (0.5, 100%), and then stays at 100% until confidence reaches 1.0.</p>	Total Images	270	Test Items	30	Precision ①	90%	Recall ②	90%
Total Images	270								
Test Items	30								
Precision ①	90%								
Recall ②	90%								
<p>F1 Score What is this model's F1 score?</p>	<p>$F1 = (2 * 9/10 * 9/20) / (9/10 + 9/10) = 0,9$</p>								