AutoML Modeling Report

Shasu Vathanan

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	Train	Validation	Test
Normal	300	229	44	27
Pneumonia	300	233	23	44

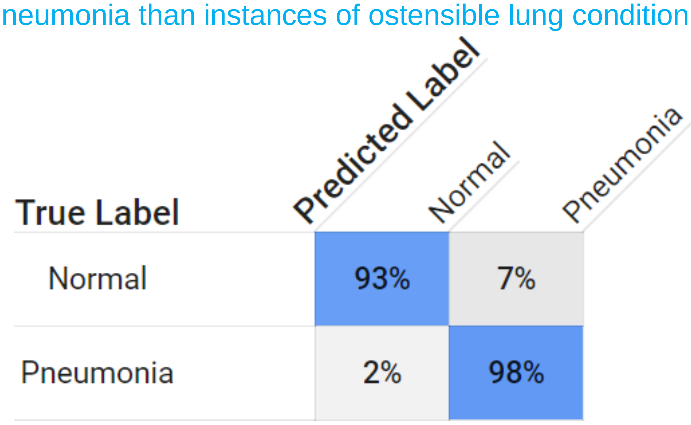
At first, 200 images were utilized (balanced). In any case, it was chosen to expand the quantity of pictures to 600 to assess the Confusion Matrix dependent on including more information.

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?

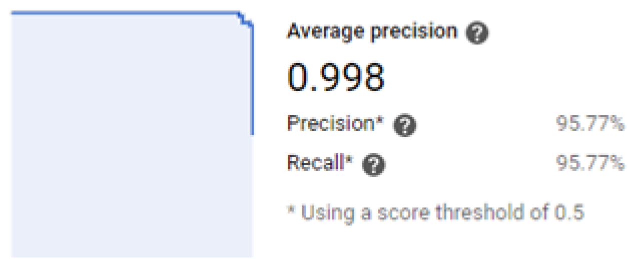
	Predicted Positives	Predicted Negatives
Actual Positives	TP	FN
Actual Negatives	TP	FN

True Positive (TP) – Correctly classified input as positive
True Negative (TN) – Correctly classified input as negative
False Positive (FP) – Misclassified input as positive
False Negative (FN) – Misclassified input as negative

Basically, the genuine positive rate for pneumonia would be the proportion of pictures that have the nearness of pneumonia effectively characterized. The bogus positive rate for the ordinary class would be instances of pneumonia that were not appropriately distinguished. The Confusion Matrix (next page) shows that the model was better at accurately distinguishing instances of pneumonia than instances of ostensible lung condition.

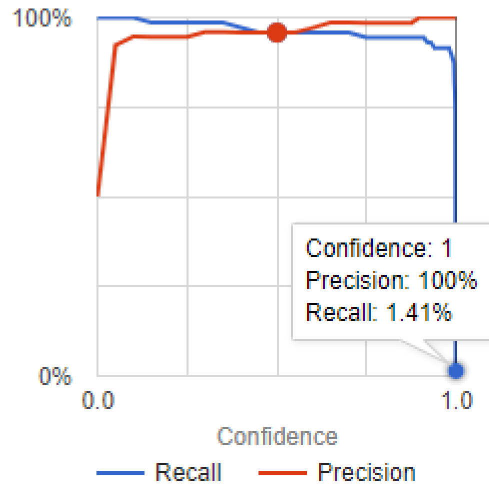


Precision and Recall
What does precision measure? What does recall measure? What precision and recall did the model achieve (report the values for a score threshold of 0.5)?



Precision is the ratio of true positives to predicted positives.
Recall is the ratio of true positives to actual positives.

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



At the point when the limit is expanded, the exactness increment to the most extreme (approaches 1) while the review ordinarily falls (approaches zero). Basically, less bogus positives are returned as the acknowledgment models have been balanced by the certainty level for a solitary mark.

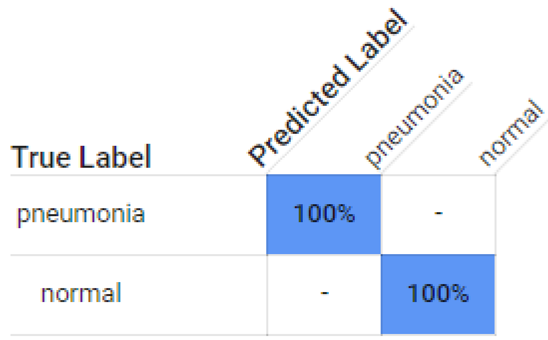
Binary Classifier with Clean/Unbalanced Data

Train/Test Split
How much data was used for training? How much data was used for testing?

Labels	Images		Train	Validation	Test
normal		100	80	10	10
pneumonia		300	240	30	30

100 normal and 300 pneumonia images were uploaded. From these 10% of each type were used for validation and 10% of each type for test, leaving 80% of each type as training data.

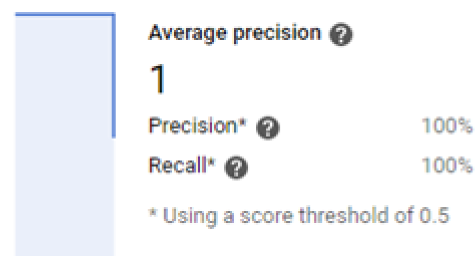
Confusion Matrix
How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix.



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

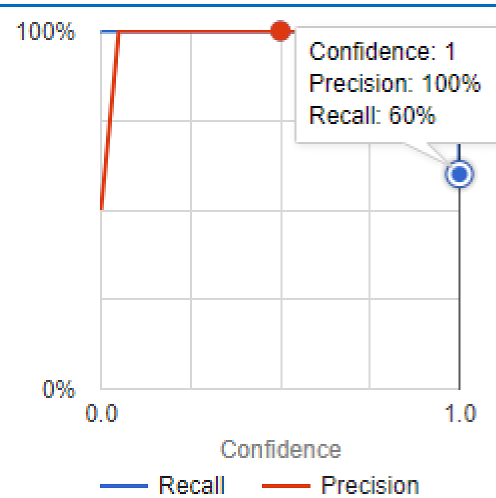
The imbalanced information has really improved the presentation esteems inside the Confusion Matrix. This may have come about because of preparing the 300 pictures on 16 hub hours (an expansion). It was at first expected that the consequences of having a lopsided set would expand the quantity of bogus positives and bogus negatives.

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)?

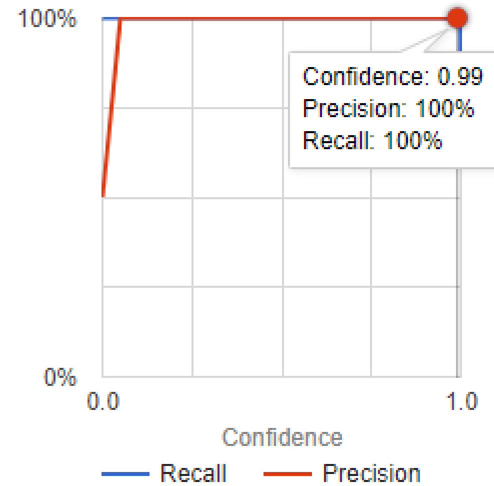


As a result of the True values above, the precision and recall are now at 100% as there are no errors present.

Unbalanced Classes
From what you have observed, how do unbalanced classed affect a machine learning model?



By and large, lopsided information presents predispositions that sway the precision of the model. Be that as it may, with such high precision, the review at a certainty estimation of one is bizarrely high (60%). For examination at a certainty estimation of 0.99, both accuracy and review stay at 100% (see underneath).



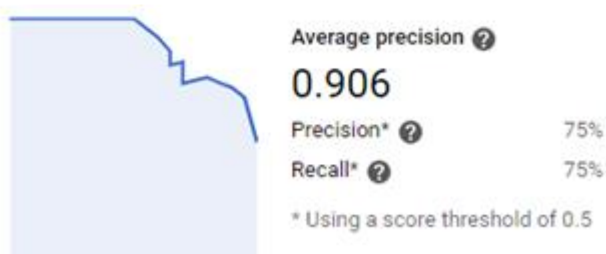
Binary Classifier with Dirty/Balanced Data

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

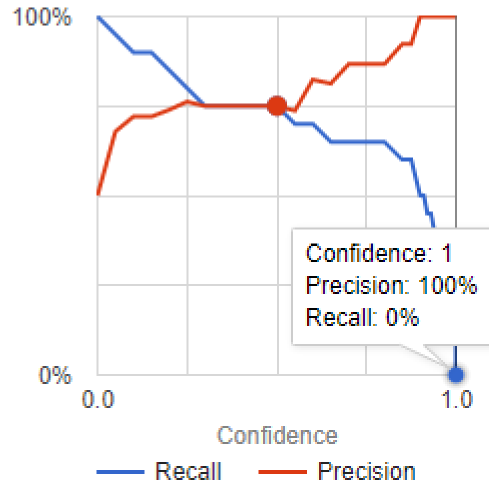
True Label \ Predicted Label	normal	pneumonia
normal	60%	40%
pneumonia	10%	90%

Watching the Confusion Matrix, the "dirty data" as unfathomably diminished the genuine positives for the ordinary picture set. Therefore, it gets hard for the model to appropriately assess pictures that are taken care of into it. With a proportion of 70:30 (clean versus dirty), the model gets infective at characterizing pictures while reaffirms the significance of guaranteeing inappropriate information is rarely blended.

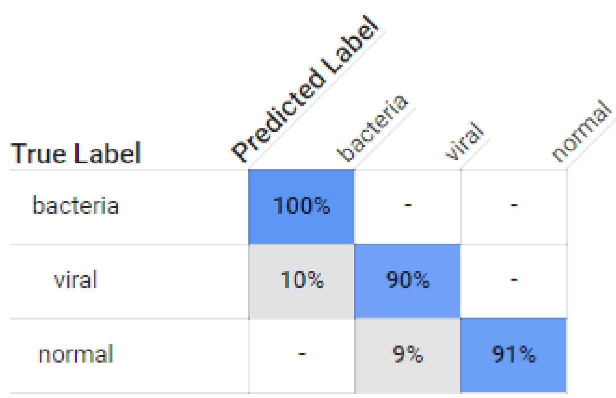
Precision and Recall
How have the model's precision and recall been affected by the dirty data (report the values for a score threshold of 0.5)? Of the binary classifiers, which has the highest precision? Which has the highest recall?

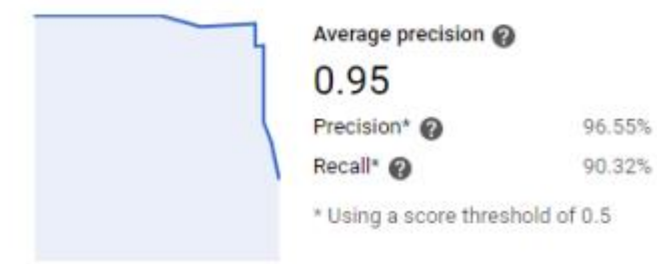


Therefore, the exactness and review have diminished however are falsely high. That is, utilizing such a model without knowing, that the information has issues will neglect to satisfy quality guidelines and would be infective in viable use. So far, the Clean-Unbalanced

	picture set has had the best exactness and review however this could be expected to overtraining (node hours) and "luck" in picture arranging taking during preprocessing (rather than choosing on the initial 100 images, for instance).
Dirty Data From what you have observed, how does dirty data affect a machine learning model?	 Investigating the traverse point, the effect of the "dirty data" can be believed to affect the quantity of right expectations dependent on the certainty level. In that capacity, cleaning information and appropriate distribution is a need for any model.

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.	 Instinctively, the bacterial and viral pneumonia cases are the most comparative which prompts mistakes in appropriate forecast. In any event, for human spectators, the distinctions can be hard to watch (which is the reason paired order prompts better outcomes while anticipating the presence of an item or illness when contrasted with legitimate recognizable proof and arrangement). To cure this, I would prepare with substantially more information just as change the naming plan since the watchword "pneumonia" is in the

	metadata for the two sorts of pneumonia (which is a legitimate decision that inability specialists could without much of a stretch make).
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)?	 $P_{model} = \frac{\sum_{i=1}^{n} P_i}{n}$ $R_{model} = \frac{\sum_{i=1}^{n} R_i}{n}$
F1 Score What is this model's F1 score?	$F1 = \frac{2*Precision*Recall}{(Precision+Recall)} = \frac{2*\Pi\left(\frac{TP}{(TP+FP)}\right)*\Pi\left(\frac{TP}{(TP+FN)}\right)}{\Pi\left(\frac{TP}{(TP+FP)}\right)+\Pi\left(\frac{TP}{(TP+FN)}\right)}$ $F1 = 93.33\%$