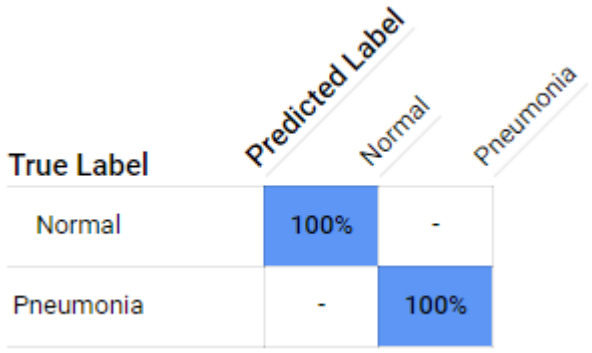
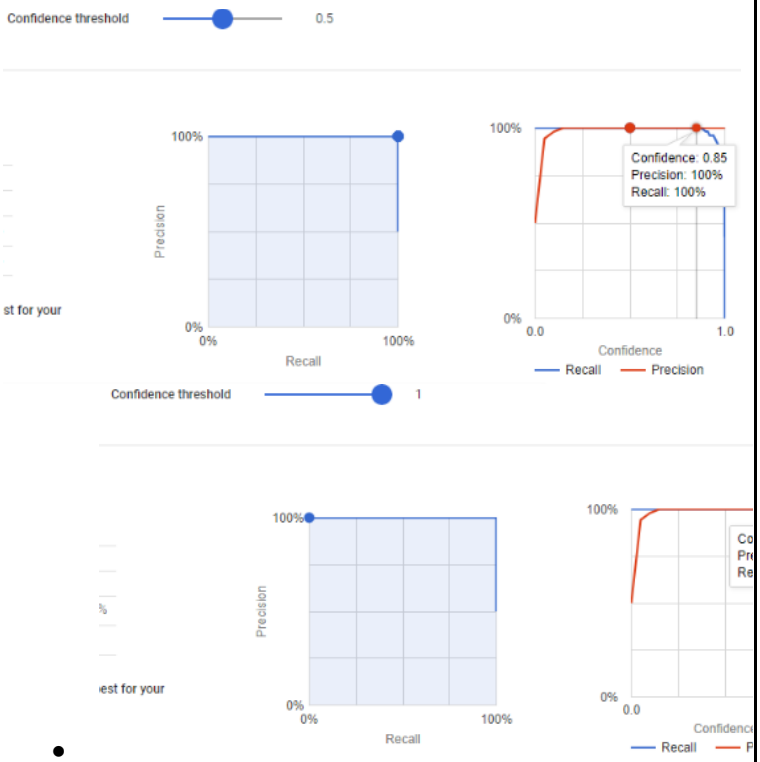
AutoML Modeling Report

Ekta Bharti

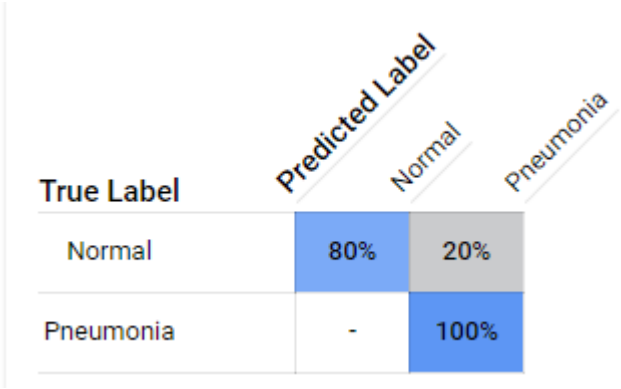
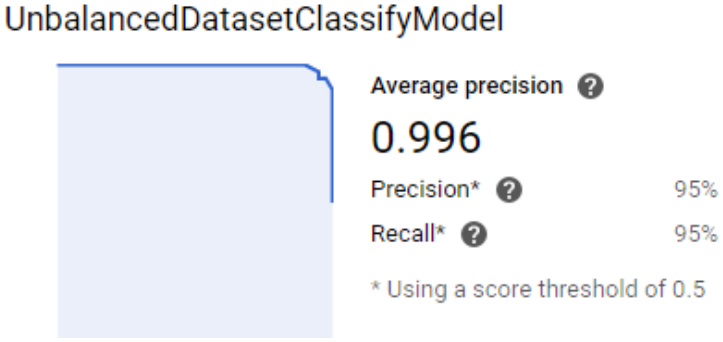
Binary Classifier with Clean/Balanced Data

Train/Test Split How much data was used for training? How much data was used for testing?	448 images used for Training 50 images 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?	• Each cells represent the following:- • True positive: Images which were actually of pneumonia and where predicted as pneumonia. • True Negative: Images which were actually of normal and where predicted as normal. • False positive: Images which were actually of normal but predicted as pneumonia. • False negative: Images which were actually of pneumonia but predicted as normal. Total images 448 Test items 50 Precision 100% Recall 100% TPR for pneumonia=1 FPR for normal= 0 
Precision and Recall	• Precision measures out of all predictions made

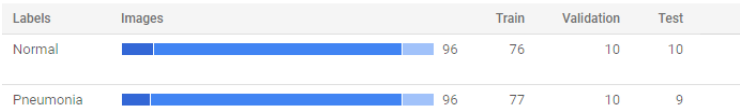
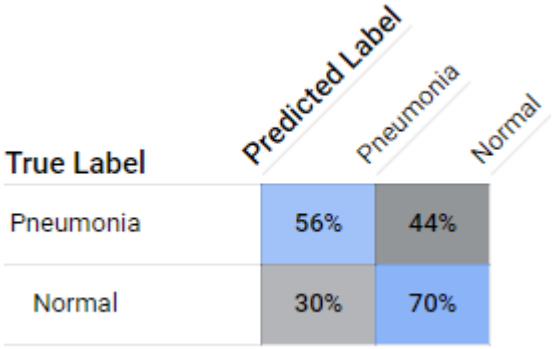
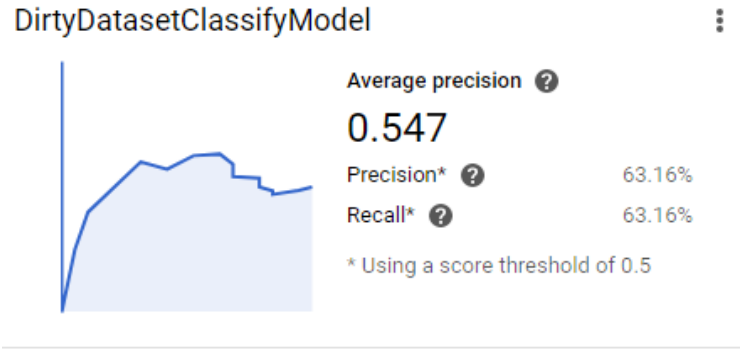
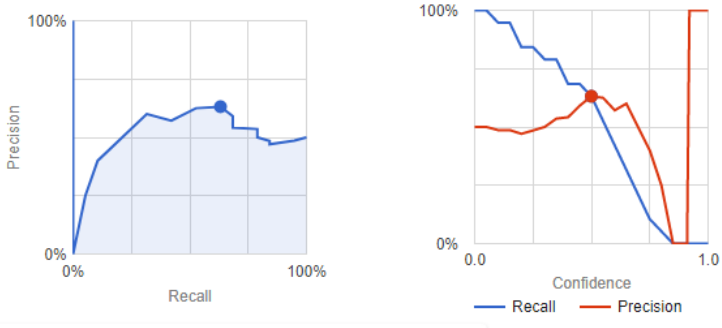
What does precision measure? What does recall measure? What precision and recall did the model achieve (report the values for a score threshold of 0.5)?	as positive how many are actually positive. • Recall measures out of all actual positives, how many were predicted as positive. • At confidence threshold of 0.5, overall model's precision reached 100% and over model's recall reached 100%.
Score Threshold When you increase the threshold what happens to precision? What happens to recall? Why?	 • But here Precision remains the same and recall decreases to 98% because Precision is already at 100%. • When we increase the threshold, precision probably should increase. This is because, Threshold increase reduces False positives.

Binary Classifier with Clean/Unbalanced Data

Train/Test Split How much data was used for training? How much data was used for testing?	Number of images used for training: 359 Number of images used for testing: 40 Average Precision

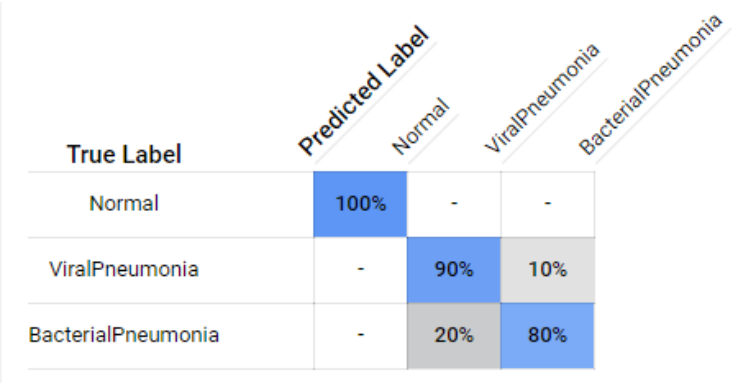
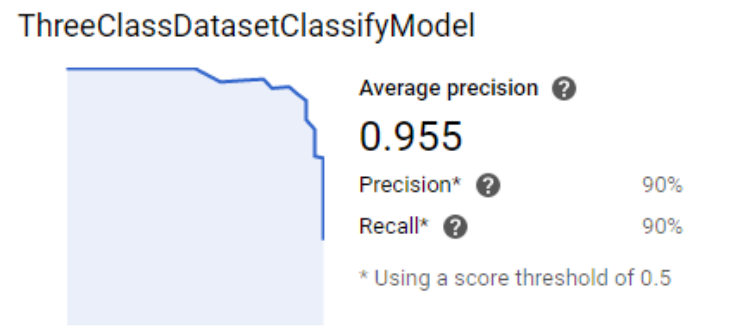
	95% Average Recall 95% For a threshold of 0.5
Confusion Matrix How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix.	 Confusion matrix is showing that it has predicted 20% of normal test data, as pneumonia. While 100% of the actual pneumonia cases are predicted correctly.
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)?	 • Overall Precision and recall of the model has gone down. • When talking about Precision of individual classifiers for Normal (or the minority class) , precision became 80% since the model is now biased towards the majority class.
Unbalanced Classes From what you have observed, how do unbalanced classed affect a machine learning model?	Unbalanced training data, tends to create bias behavior for the majority class and ends up predicting some of the minority class as majority

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.	96-actual normal 96- actual pneumonia 28- mislabeled in each normal and pneumonia classes  Confusion matrix has ended up predicting wrong labels both for pneumonia and normal labels. Precision and recall for both classifiers has fallen considerably. 
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?	 Recall for pneumonia=56% Recall for normal=70% ⇨ Hence normal has higher recall Precision for pneumonia=(56%0f 96)/{(56%of 96)+(30% of 96)}=65% Precision for normal=44/(44+70)=38.5% ⇨ Pneumonia has a higher precision
Dirty Data From what you have observed, how does dirty data affect a machine learning model?	

	Dirty data decreases the performance of the model considerably. Both precision and recall take a hit.

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.	 => Normal has 100% precision. 10% of Viral pneumonia images got predicted as Bacterial, while 20% of Viral pneumonia images are predicted as Viral. =>Viral and Bacterial pneumonia images are the most confusing classes. =>Normal class if the most likely to get right. I may try to add more images data to improve the confusion matrix
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)?	 Model's precision is 90% Mode's recall is 90% Model's precision is average of precision of individual classes. Model's recall is average of recall of individual classes.

F1 Score What is this model's F1 score?	F_1 = 2*90*90/(90+90)= 90%