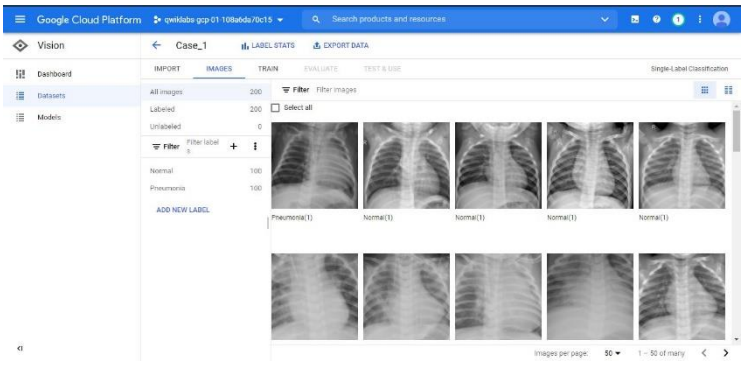
AutoML Modeling Report

Harini Pavithra Elangovan

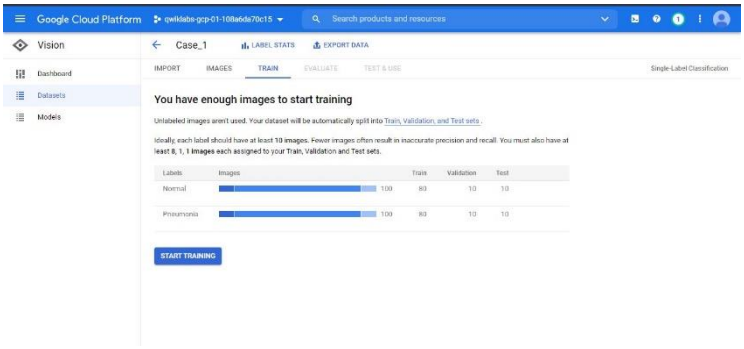
Binary Classifier with Clean/Balanced Data

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

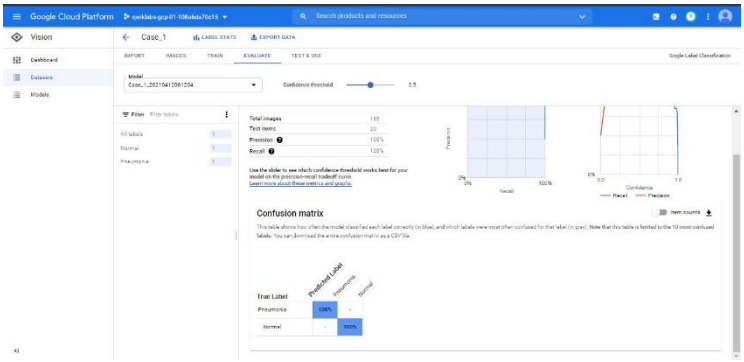
Initialy,200 images were used. Out of which 100 belongs to Normal class and remaining 100 belongs to Pneumonia class in order to create a Clean/Balanced data.



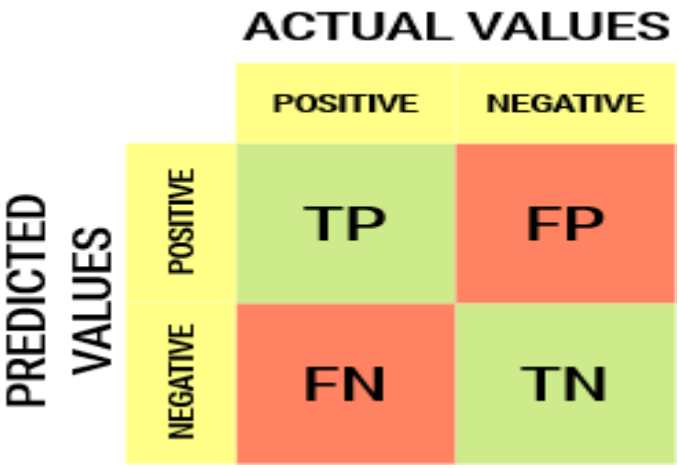
In which the 80% of the images are used for Training and 10 percent are used for Testing, remaining for Validation



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?

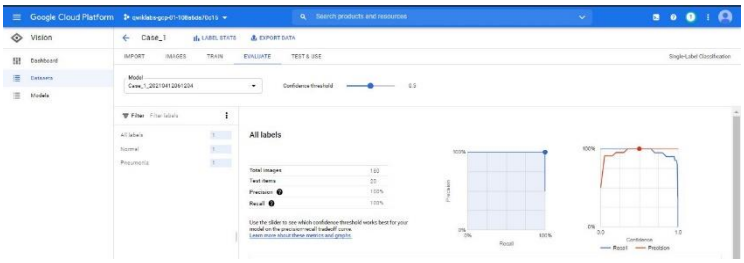


A **confusion matrix** is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.

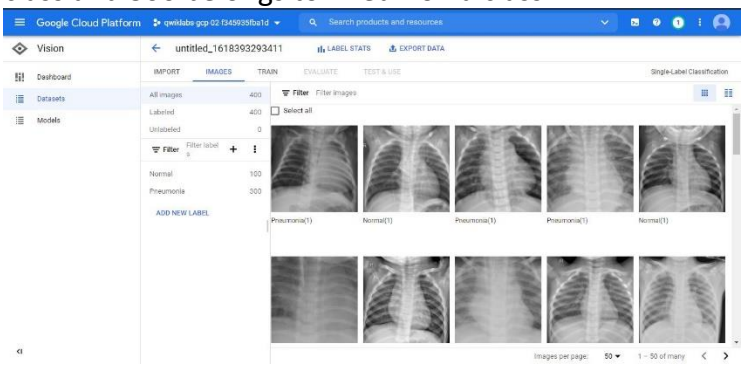
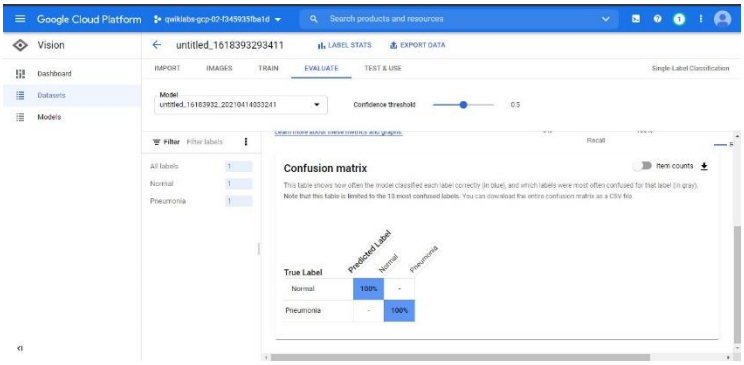


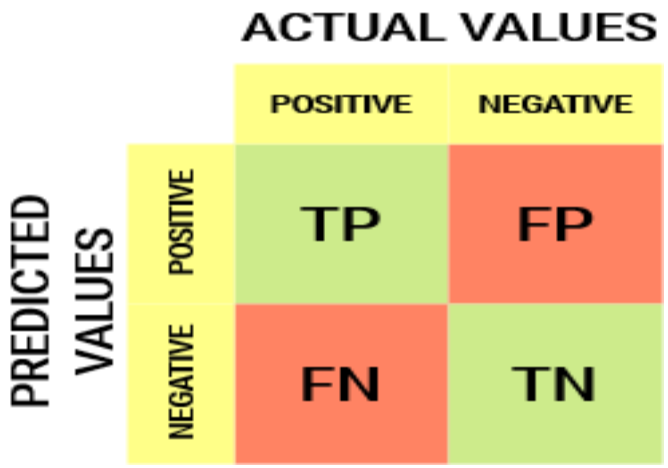
True positives (TP): These are cases in which we predicted yes for the data
True negatives (TN): We predicted no for the data
False positives (FP): We predicted yes, but it’s not true (Also known as a "Type I error.")
False negatives (FN): We predicted no, but it’s true (Also known as a "Type II error.")

True positive rate for Pneumonia class is 100%(all the cases are predicted accurately), False positive rate for Normal class is 0(not even a single case was predicted wrongly)

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)?	 The **precision** is the proportion of relevant results in the list of all returned search results. The **recall** is the ratio of the relevant results returned by the search engine to the total number of the relevant results that could have been returned. The model achieved a precision of 100% and recall of 100%.
Score Threshold When you increase the threshold what happens to precision? What happens to recall? Why?	When you increase the score threshold, the precision goes up and recall seems to decrease. This is because when you increase the score threshold you want to be more confident when you make a prediction. Hence by increasing the score threshold, your will classify fewer images but it will have lower risk of misclassifying the images.

Binary Classifier with Clean/Unbalanced Data

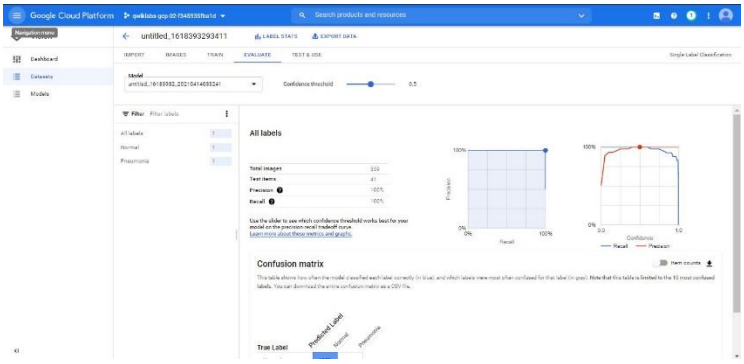
Train/Test Split How much data was used for training? How much data was used for testing?	400 images were used, in which 100 belongs to Normal class and 300 belongs to Pneumonia class  80% of the images are used for Training, 10% used for Testing and the remaining for Validation
Confusion Matrix How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix.	 A **confusion matrix** is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.



True positives (TP): These are cases in which we predicted yes for the data
True negatives (TN): We predicted no for the data
False positives (FP): We predicted yes, but it's not true (Also known as a "Type I error.")
False negatives (FN): We predicted no, but it's true (Also known as a "Type II error.")

True positive rate for Normal class is 100%(all the cases are predicted accurately), True Negative positive rate for Pneumonia class is 0%(not even a single case was predicted wrongly). There is not much difference between the Balanced and unbalanced data.

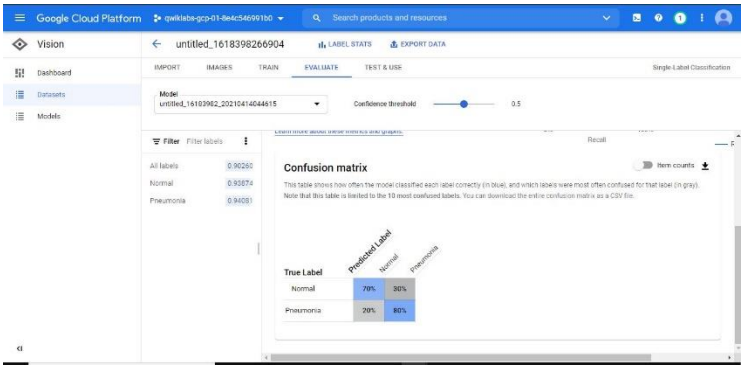
Precision and Recall
How have the model's precision and recall been affected by the unbalanced data (report the values for a score threshold of 0.5)?

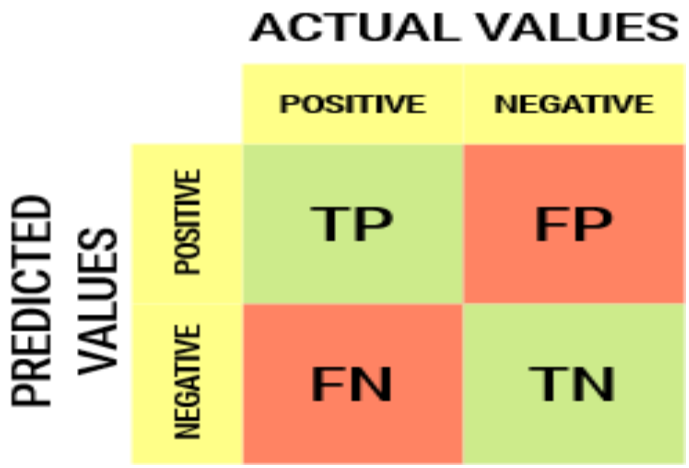


The **precision** is the proportion of relevant results in the list of all returned search results. The **recall** is the ratio of the relevant results returned by the search engine to the total number of the relevant results that could have been returned.

	The model achieved a precision of 100% and recall of 100%. The unbalanced data doesn't affected the model's precision and recall much.
Unbalanced Classes From what you have observed, how do unbalanced classed affect a machine learning model?	Unbalanced data didn't introduce any bias in this case, the True positive and False negative rates were accurate and even the Precision, Recall of the model.

Binary Classifier with Dirty/Balanced Data

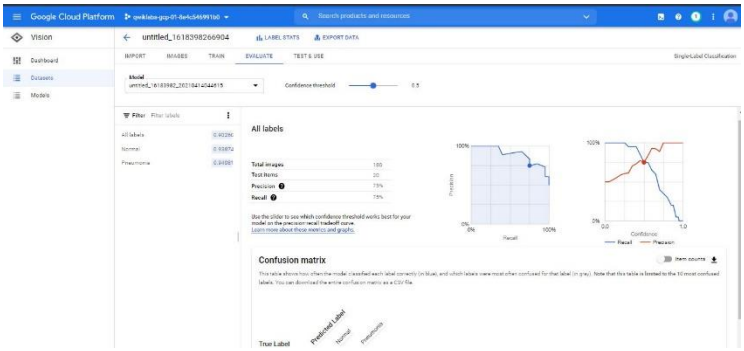
Confusion Matrix How has the confusion matrix been affected by the dirty data? Include a screenshot of the new confusion matrix.	 A **confusion matrix** is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.



True positives (TP): These are cases in which we predicted yes for the data
True negatives (TN): We predicted no for the data
False positives (FP): We predicted yes, but it's not true (Also known as a "Type I error.")
False negatives (FN): We predicted no, but it's true (Also known as a "Type II error.")

The True positive rate of Normal is 70%, False positive rate of Pneumonia class is 30%. Similarly, the False negative rate of Normal Class is 20% and True negative rate of Pneumonia class is 80%
Dirty data introduces bias. Model will have a bias towards predicting the label, in which some images are exchanged in both the classes.

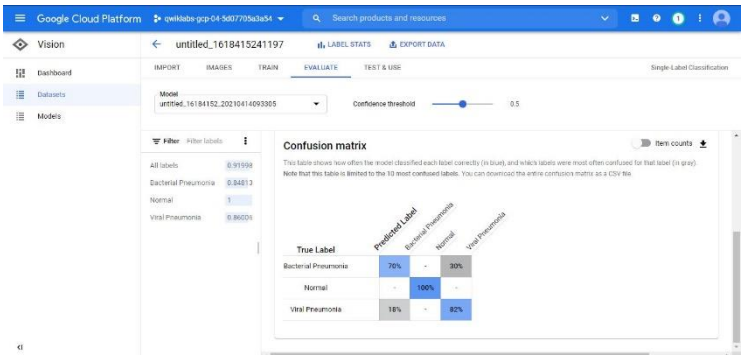
Precision and Recall
How have the model's precision and recall been affected by the dirty data (report the values for a score threshold of 0.5)? Of the binary classifiers, which has the highest precision? Which has the highest recall?

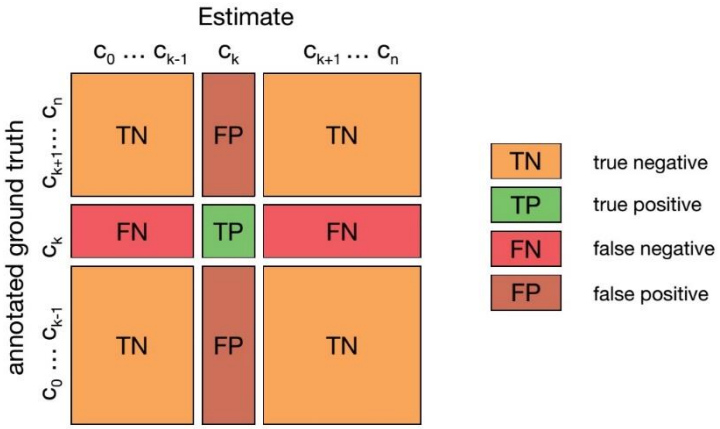


The **precision** is the proportion of relevant results in the list of all returned search results. The **recall** is the ratio of the relevant results returned by the search engine to

	the total number of the relevant results that could have been returned. The Precision and Recall of the model went down to 75%. Of the binary classifiers, the balanced, unbalanced data has the highest Precision and recall of 100%
Dirty Data From what you have observed, how does dirty data affect a machine learning model?	Machine learning model clears struggles to find patterns among classes as the data is mixed up. Model sees same patterns in both labels and hence performs poorly.

3-Class Model

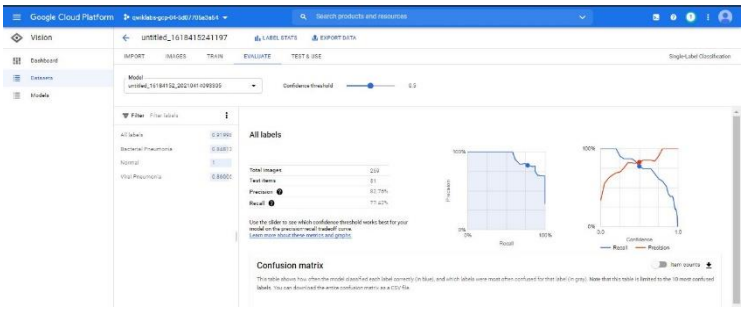
Confusion Matrix Summarize the 3-class confusion matrix. Which classes is the model most likely to confuse? Which class(es) is the model most likely to get right? Why might you do to try to remedy the model's "confusion"? Include a screenshot of the new confusion matrix.	 A **confusion matrix** is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.



True positives (TP): These are cases in which we predicted yes for the data
True negatives (TN): We predicted no for the data
False positives (FP): We predicted yes, but it's not true (Also known as a "Type I error.")
False negatives (FN): We predicted no, but it's true (Also known as a "Type II error.")

The True Negative rate of Bacterial Pneumonia class is 70%, True Negative rate of Viral Pneumonia class is 30%, the True Positive rate of Normal class is 100%,the True Negative rate of Viral Pneumonia class is 18%,True Negative rate of Bacterial Pneumonia rate is 82%.

The model which is like to confuse is Bacterial Pneumonia, Viral Pneumonia class. Normal class is likely to get right (True Positive rate – 100%). We can add more images to each class as there only 100 images for each class now.

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 Precision and Recall for the model is 82.76%, 77.42%.Precision is calculated with true positives by sum of true positives, false positives whereas Recall is calculated with true positives by sum of true positives and false negatives
F1 Score What is this model's F1 score?	The F1 score of the model is 0.80