AutoML Modeling Report

Ahmed Abdelnasser

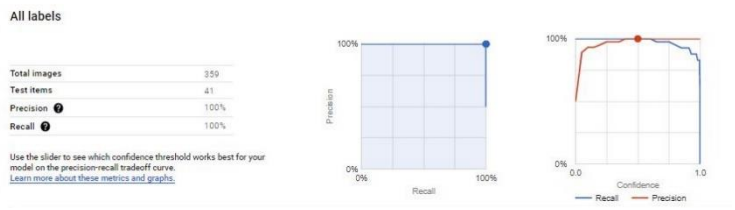
Binary Classifier with Clean/Balanced Data

Train/Test Split How much data was used for training? How much data was used for testing?	Dataset consists of 200 images, 100 for normal class and 100 for pneumonia class in order to create a Clean/Balanced data. Training Data: 80 Testing Data: 10 Validation: 10 In both Normal and Pneumonia.
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 on a set of test data for which the true values are known. The four cells represent TP, TN, FP, FN. The TP for pneumonia class is 100% and TN rate for normal class is 100% FP for Pneumonia 0 and FN for Normal 0

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. **Precision** measures the ability of a model to identify only the relevant data. **Recall** measures the ability of a model to identify and find all of the relevant data. 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?	As the threshold value increases, Precision increases but Recall decreases, such that at a confidence. because when we increase the score threshold, we want to be more confident when we make a prediction. Recall = true positives / (true positives + false negatives) Precision = true positives / (true positives + 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?	Dataset consists of 400 images, 100 for normal class and 300 for pneumonia class in order to create a Clean/Unbalanced data. **For Normal:** Training Data: 80 Testing Data: 10 Validation: 10 **For Pneumonia:** Training Data: 239 Testing Data: 31 Validation: 30
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 on a set of test data for which the true values are known. The four cells represent TP, TN, FP, FN. The TP for pneumonia class is 100% and TN rate for normal class is 100% FP for Pneumonia 0 and FN for Normal 0

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. **Precision** measures the ability of a model to identify only the relevant data. **Recall** measures the ability of a model to identify and find all of the relevant data. The model achieved a precision of 100% and recall of 100% 
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 TP and FN 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 on a set of test data for which the true values are known. The four cells represent TP, TN, FP, FN. The TP for pneumonia class is 50% and TN rate for normal class is 100% FP for Pneumonia 50% and FN for Normal 0 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. **Precision** measures the ability of a model to identify only the relevant data. **Recall** measures the ability of a model to identify and find all of the relevant data. The model achieved a precision of 75% and recall of 75%, Which decreased 25% of clean data. The highest precision founded in clean/balanced and clean unbalanced as it equal to 100% The highest recall founded in clean/balanced and clean unbalanced as it equal to 100%

Dirty Data From what you have observed, how does dirty data affect a machine learning model?	From observation, dirty data (incorrectly annotated data) can significantly affect a model's Precision and Recall negatively, especially increasing False Positives and False Negatives

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 on a set of test data for which the true values are known. The TN rate of Bacterial Pneumonia class is 70%, TN rate of Viral Pneumonia class is 30%, the TP rate of Normal class is 100%, the TN rate of Viral Pneumonia class is 18%, TN rate of Bacterial Pneumonia rate is 82%. We can attempt to remedy this situation and improve the model by adding more accurately labeled data for the bacterial and viral pneumonia classes.

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 model precision is 91.6, Which calculated by Sum TP for all classes / Sum (TP for all classes + FP for all classes) = sum precision for all classes / 3 = (100+100+75)/3= 91.6 The model recall is Which calculated by Sum TP for all classes / Sum (TP for all classes + FN for all classes) = sum recall for all classes / 3 = (100+100+75)/3= 91.6
F1 Score What is this model's F1 score?	F1 Score = (2 * Precision * Recall) / (Precision + Recall) F1 score = (2 * 0.916 * 0.916) / (0.916 + 0.916) = 0.9