AutoML Modeling Report

Jude Anlasun

Binary Classifier with Clean/Balanced Data

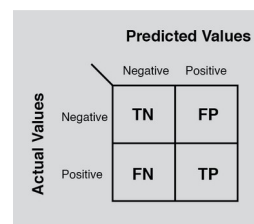|	
Train/Test Split How much data was used for training? How much data was used for testing?	There is total of 200 images of which 100 images are labeled as "normal" and the other 100 labeled as "pneumonia". 80% of the images was used for training. 10% was used for Validation and the remaining 10% was 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?	A confusion matrix is a table that describes the performance of a classificaiton model on a set of test data for which the true values are known. From the diagram above the different cells labeled (TP, FN, FP and FN) of the confusion matrix can be descriped based on the model as follows: • **TP – True Positive:** cases predicted as pneumonia and were actually pneumonia. • **TN – True Negetive:** cases predicted as normal and were actually normal. • **FP – False Positive:** cases predicted as pneumonia, but were actually normal. • **FN – False Negetive:** cases predicted as normal but were actually pneumonia. The values observed in the actual predictions can be seen in the below screenshot:

	  True Positive Rate (pneumonia) = TP/(TP+FN) = 10 /(10+0) *100 = **100%** **False Positive Rate (normal)** = FP/(FP+TN) = 0/(0+10) * 100 = **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)?	Precision measures how often the correctly predicted observation is correct. It is expressed as the ratio of the correctly predicted positive observations to the total positive observations. Recall measures how truly relevant results are return. It is expressed as a ratio of the number of positive class predictions to all positive observations in the dataset. The model at a score threshold of 0.5 recorded a Precision of 100% and a Recall at 100%.
Score Threshold When you increase the threshold what happens to precision? What happens to recall? Why?	Precision is already at 100%, increasing the score threshold decreases the recall however, the precision is still maintained at a 100%. Assuming that the precision wasn't already at 100% increasing the score threshold would have decreased recall score with an increase in precision score. This happens because when the score threshold is increased precision increases which produces fewer false positives with a high confidence of lower misclassification.

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 for the development of this model. 300 images are labeled as "pneumonia" The other 100 labeled as "normal". 80% of the images was used for training. 10% was used for Validation and the remaining 10% was used for testing.
Confusion Matrix How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix.	The confusion Matrix predicted has a True Positive rate of 100% and a True Negetive rate at 100%. Meaning there was no misclassification and that the prediction was 100% accurate.
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)?	Precision and Recall of the unballanced data is at 100% respectively. There is actually no effect on the precision and recall of the unbalanced data as compared to that of the clean/balanced data.
Unbalanced Classes From what you have observed, how do unbalanced classed affect a machine learning model?	The unbalanced data didn't have any effect on the machine learning model as the models precision and recall was at 100%. From my observation the unbalanced model is more likely to predict the majority class (pneumonia) with the highest number than the minority class. This would indicate that the unbalanced data is biased towards the majority class. If the unbalanced data did not exhibit bias towards the majority class at least there should have been a few misclassification.

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.	In the unbalanced model for the "normal" class 8 were predicted as true normal whereas 2 were misclassified as pneumonia. For the "pneumonia" class 9 were predicted as pneumonia whereas 1 was misclassified as normal. **True Positive Rate (pneumonia)** = TP/(TP+FN) = 9 /(9+1) *100 = 9**0%** **False Positive Rate (normal)** = FP/(FP+TN) = 2/(2+8) * 100 = 20**%** The dirty data introduced a bias in the model thereby increasing the False Positive rate from 0% as seen in the clean data to 20% and reducing the True Positive rate from 100% in the clean data to 90%.
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 models Precision at a score threshold of 0.5 has a precision of 85% and a recall of 85%. This may be due to the dirty data introduced. The dirty model is more likely to produce more false positives and more false negetives than the Clean/Balanced and Clean/Unbalanced models which both have a precision and recall at 100% respectively.
Dirty Data From what you have observed, how does dirty data affect a machine learning model?	It can be observed that dirty increases false positives and false negetives in a model thereby reducing the precision and recall scores.

3-Class Model

Confusion Matrix	
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.	 The 3-class confusion matrix as it can be seen above, out of 10 bacteria pneumonia cases it predicted 8 as bacteria pneumonia and misclassified 2 as normal. Out of 10 viral pneumonia cases it predicted 9 as viral pneumonia and misclassified one as bacteria pneumonia. Out of 10 normal cases it predicted all of them as normal without any pneumonia. The models confusion matrix can be remedy by increasing the number of training dataset for the model to learn from more patterns. Slightly reduction in the score threshold also increases the Precision and Recall which would definately provide some remedy for the confusion matrix. The model is most likely to get the class "normal" right.
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's precision score is 89.29% and a recal score is at 83.33%. Precision = True Positives / (True Positives + False Positives) Recall = True Positives / (True Positives + False Negetives) Precision(bacteria pneumonia) = 80/(80+10) * 100 = 88.89% Precision(viral pneumonia) = 90/(90) * 100 = 100% Precision(normal) = 10/(10+2) * 100 = 83.33% **Average model precision = (88.89+100+83.33)/3 = 90.74%** Recal(bacteria pneumonia) = 8/(8+2) * 100 = 80% Recal(viral pneumonia) = 9/(9+1) * 100 = 90% Recal(normal) = 10/10 * 100 = 100% **Average model recall = (80+90+100)/3 = 90%**
F1 Score What is this model's F1 score?	F1 Score = (2 * Precision * Recall)/(Precision+Recall) F1 Score = (2 * 89.29 * 83.33)/(89.29 + 83.33) = 86.21%

References:

[1] https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/