|  |  |
| --- | --- |
| AutoML Modeling Report |  |

*Noureddine Ettayyeby*

Binary Classifier with Clean/Balanced Data

|  |  |
| --- | --- |
| **Train/Test Split**  How much data was used for training? How much data was used for testing? | The model used 80% (160) of pictures for training, 10% (20) for validation, and 10% (20) 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? | The first cell in the confusion matrix is the rate of the pneumonia class that was correctly predicted (TP) which 100%, the second cell represent the rate of the same class which falsely predicted to be normal (FP) in this case 0%, as for the third cell represent the rate of the normal class which was falsely predicted to have pneumonia (FN) in this instance 0%, the last cell is the rate of the normal class which was rightly predicted (TN) 100. |
| **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 measure how accurate the model when it predicts Yes, in this case, the model is 100% accurate, the recall, in this case, measures how good the model at predicting the pneumonia class which in this case 100%,  See the screenshot below. |
| **Score Threshold**  When you increase the threshold what happens to precision? What happens to recall? Why? | When the threshold is increased, the precision stays the same but the recall drops a little bit. This happens because when we increase the threshold the probability of some point being in the normal class well fell below the threshold and be miss classified even though they are in the normal class. |

Binary Classifier with Clean/Unbalanced Data

|  |  |
| --- | --- |
| **Train/Test Split**  How much data was used for training? How much data was used for testing? | The model used 80% (320 images ) of the data for training, 10% (40 images) validation, and 10% (40 images) for testing. |
| **Confusion Matrix**  How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix. | The effect of the unbalanced data on the confusion matrix was that 10% of the lesser class which in this case the normal class got wrongly classified as has pneumonia see the screenshot below. |
| **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 and recall both drop to 97.5 % from 100% compared to the previous model. |
| **Unbalanced Classes**  From what you have observed, how do unbalanced classes affect a machine learning model? | The unbalanced data cloud has a huge impact on the performance of the model, usually, the lesser class gets misclassified when the data is unbalanced which has an impact on the overall performance of a model that’s why we need to address this problem before training 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. | The dirty data had a huge impact on the confusion matrix compared to the two previous models, here we see that 20% of pneumonia got misclassified as normal and 30% of the normal class got misclassified as has pneumonia see the screenshot below. |
| **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 model and precision have dropped significantly compared to previous models, for a threshold of 0.5 we got both of them at 75%.    The first classifier has the highest recall and precision which shows the effect of clean and balanced data on the performance of a model, on the other hand the third classifier has the lowest of both recall and precision which show how bad the effect of dirty data can be for a model performance. |
| **Dirty Data**  From what you have observed, how does dirty data affect a machine learning model? | A model train with dirty data will result in a bad model. As we can see here, we get a lot of misclassified instance for both classes. And a significant drop in precision and recall that’s why we need to take a good care to clean the data and get it to the right for format before training 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. | The confusion matrix tells us that the model gets 100% of its prediction right for the bacterial class and the normal class, but preform badly on viral class by getting just around 70% right, confuse 10% with the normal class and 20% with the bacterial class. See the screenshot below.  As for remediations we need first identifies the cause of this problem maybe the data for the class is unclean so need to clean it, but the more likely cause here is that the viral class is similar to the bacterial class that’s why the model confuses these two more often one possible solution is to increase the data for building the model. |
| **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)? | Both recall and precision are at 90% for a threshold of 0.5. we calculate these by calculating recall and precision for each class and some them up then divide the some by the number of classes. |
| **F1 Score**  What is this model’s F1 score? | The formelt for calculating the F1 is:  Which equal |

Inference

I run an inference job against the imbalanced model using 1000 images, 500 from each class and this the confusion matrix I got.

