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

*Clarissa Watson*

Binary Classifier with Clean/Balanced Data

|  |  |
| --- | --- |
| **Train/Test Split**  How much data was used for training? How much data 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? | The above confusion matrix displays the number of true positives, true negatives, false positives, and false negatives given some number of input data points n.  True positive rate for pneumonia class is 86%  False positive rate for normal class is 14% |
| **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 indicates how often a model accurately predicts a positive label, formula: True positives/ Total predicted positives  Recall is the number of positives the model predicted, formula = True positives/ Total actual positives |
| **Score Threshold**  When you increase the threshold what happens to precision? What happens to recall? Why? | The score threshold is the level of confidence the model must have to assign a category to an image.  When I increase the threshold to .75 from .5 the precision level increase and the recall remains the same. This is because a higher threshold allows for higher chances the model will make accurate predictions because, as we increase the threshold there will be fewer false positives. The model is more confident in its prediction. |

Binary Classifier with Clean/Unbalanced Data

|  |  |
| --- | --- |
| **Train/Test Split**  How much data was used for training? How much data 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. | I am not sure if there has been an error, but it appears the confusion matrix has been impacted by the unbalanced data (ie: 100 normal and 300 pneumonia), because both the precision and recall at 100% confident at a .5 threshold. |
| **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 model’s precision and recall are showing 100% accuracy which does not seem realistic. |
| **Unbalanced Classes**  From what you have observed, how do unbalanced classed affect a machine learning model? | After testing and model by uploading a few images not from the training data, it labeled the images correctly. It seems unbalanced data impacts the model by making it over-confident. When I increased the threshold to .75, the precision remained at 100%, while the recall dropped to 92.5%. |

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 confusion matrix was affected by the dirty data because the precision and recall are 97.5%, meaning the model is not totally confident in its precision to label the images. |
| **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’s precision and recall were affected by the data dirty because neither are 100% confident in their ability to classify images as both are 97.5% when the threshold value if 0.5%.    The normal classifier has the highest precision and recall. |
| **Dirty Data**  From what you have observed, how does dirty data affect a machine learning model? | The dirty data affects the model’s ability to be 100% confident in its prediction. When testing this model with images, the model successfully tested the image but with a .97 score. |

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. |  |
| **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)? |  |
| **F1 Score**  What is this model’s F1 score? |  |