AutoML Modeling Report

#### *<Your Name Here>*



# Binary Classifier with Clean/Balanced Data

|  |  |
| --- | --- |
| **Train/Test Split** How much data was used for  training? How much data was used  for testing? | <your answer text here> |
| **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? | <your answer text here> |
| **Precision & 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)? | <your answer text here> |
| **Score Threshold** When you increase the score threshold, what happens to precision? What happens to recall? Why? | <your answer text here> |

# Binary Classifier with Clean/Unbalanced Data

|  |  |
| --- | --- |
| **Train/Test Split** How much data was used for  training? How much data was used  for testing? | <your answer text here> |
| **Confusion Matrix** How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix. | <your answer text here> |
| **Precision & 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.) | <your answer text here> |
| **Unbalanced classes** From what you’ve observed, how do unbalanced classes affect a machine learning model? | <your answer text here> |

# 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. | <your answer text here> |
| **Precision & 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? | <your answer text here> |
| **Dirty Data** From what you’ve observed, how do dirty data affect a machine learning model? | <your answer text here> |

# 3-Class Model

|  |  |
| --- | --- |
| **Confusion Matrix** Summarize the 3-class confusion matrix. What classes are the model most likely to confuse? What class(es) is the model most likely to get right? What might you do to try to remedy the model’s “confusion”?  Include a screenshot of the new confusion matrix. | <your answer text here> |
| **Precision & Recall** What are the model’s precision and  recall? How are these values  calculated? (Report the values for a  score threshold of 0.5.) | <your answer text here> |
| **F1 Score** What is this model’s F1 score? | <your answer text here> |