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

*Ashkan Yousefi*

Binary Classifier with Clean/Balanced Data

|  |  |
| --- | --- |
| **Train/Test Split**  How much data was used for training? How much data was used for testing? | I used 271 items as normal images and also 281 items as pneumonia images to be uploaded into the Google ML. I used 30% of the training set 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? | There are two main values in the confusion matrix i.e. precision and recall. The precision describes how accurately the model can predict and the formulation is calculated as follows:  TP/(TP+FP) i.e. TP (True Positive) and FP (False Positive)  The recall is calculated as the correct number of identified items to the total number of items analyzed and the formulation will be as follows:  TP/(TP+FN)    Normal case:  Precision = 100/(100+0) = 100%  Recall = 100 / (100 + 4) = 96.1%  Pneumonia  Precision: 96/(96+4) = 96%  Recall : 96/(96+0)=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 means how accurate the model can identify the normal cases and the cases for pneumonia in the current project.  Recall means what percentage of the normal cases identified by the model and it is applicable to the pneumonia as well.  Normal case:  Precision = 100/(100+0) = 100%  Recall = 100 / (100 + 4) = 96.1%  Pneumonia  Precision: 96/(96+4) = 96%  Recall : 96/(96+0)=100% |
| **Score Threshold**  When you increase the threshold what happens to precision? What happens to recall? Why? | When we increase the threshold the precision level increases but the recall level goes down. I have compared two cases for the precision and the recall with confidence level of 0.65 and the confidence level of 0.14 |

Binary Classifier with Clean/Unbalanced Data

|  |  |
| --- | --- |
| **Train/Test Split**  How much data was used for training? How much data was used for testing? | The confidence level of 0.65:    The confidence level of 0.14: |
| **Confusion Matrix**  How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix. |  |
| **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)? |  |
| **Unbalanced Classes**  From what you have observed, how do unbalanced classed affect a machine learning 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. |  |
| **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? |  |
| **Dirty Data**  From what you have observed, how does dirty data affect a machine learning 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. |  |
| **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? |  |