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

*Yasin Hassan*

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 110 normal images and 121 pneumonia images for training, so 231 images in total. There were 208 images 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? | Chart, waterfall chart  Description automatically generated  The cells in the confusion matrix state the true negatives, true positives, false negatives, and false positives. The 91% refers to the true negatives. The 9% is the false positive. The 8% is the false negatives. The 92% is the true positives. |
| **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 is the “number of true positives over all positives” (Precision and Recall). The recall is the “number of true positives over true positives plus false negatives” (Precision and Recall). Both the precision and recall were 91.3% for 0.5 threshold. |
| **Score Threshold**  When you increase the threshold what happens to precision? What happens to recall? Why? | If you increase the threshold, the precision increases while the recall decreases. This is because for the graph of recall-precision graph, after a period of stillness, the precision decreases as the recall increases.  f1 = (2)(.913)(.913))/(.913+.913)=.913=91.3%  Chart, line chart  Description automatically generated |

Binary Classifier with Clean/Unbalanced Data

|  |  |
| --- | --- |
| **Train/Test Split**  How much data was used for training? How much data was used for testing? | I had 100 normal images and 299 pneumonia images. 359 images were used for testing. |
| **Confusion Matrix**  How has the confusion matrix been affected by the unbalanced data? Include a screenshot of the new confusion matrix. | Chart, waterfall chart  Description automatically generated  The true negative is now 90%, the false positive is 10%, the 3% is the false negative, the 97% is the true positive. The true negative and the false negative decreased while the true positive and the false positive increased. |
| **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 are now 95% for 0,5 threshold, so it increased compared to the clean-balanced. |
| **Unbalanced Classes**  From what you have observed, how do unbalanced classed affect a machine learning model? | F1=2(.95)(.95)/(.95+.95)=.95=95%, so the unbalanced model gave a higher f1 score. Since the f1 score is the “overall measure of the model’s performance” the model has performed better with clean, unbalanced data (F1 Score). |

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. | A picture containing chart  Description automatically generated  The true negative is 77%, false negative is 15%, false positive is 23%, the true positive is 85%. The true negative and true positive have decreased, and the false negative and positive have increased from the clean, balanced version. |
| **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 precision and recall is now 80.77%.. The clean-unbalanced binary classifier has the highest precision and recall. |
| **Dirty Data**  From what you have observed, how does dirty data affect a machine learning model? | The f1=2(.8077)(.8077)/(.8077+.8077)=.8077, so the dirty data decreased the effectiveness of the 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. | Table  Description automatically generated  This matrix predicts the behavior of the model. There will likely be more confusion since pneumonia is now split up into bacterial and viral pneumonia. It should correctly predict the normal pictures behavior. To remedy the confusion, I could train the model for a longer period of time, or change the names of the parameters. |
| **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 precision is 80% and the recall is 72.73%. The precision is the “number of true positives over true positives plus false positives”, and the recall is the “number of true positives over true positives plus false negatives” (Model Evaluation). |
| **F1 Score**  What is this model’s F1 score? | F1 = 2(.8)(.7273)/(.8+.7273) = .7619197276 |

References

F1 Score. Udacity. Retrieved December 14, 2020 from

<https://classroom.udacity.com/nanodegrees/nd088/parts/1bcfdbea-d4b3-45ba-bcee-f6532d0f83f8/modules/9ff8da38-5b15-4e31-96fe-fb5df5ff85e9/lessons/efc39a55-6ffd-4fb1-a07c-1a2f7b6ff2c7/concepts/ed39c34a-c21e-455f-a17e-215f2dc9ecf9>

Model Evaluation. Udacity. Retrieved December 14, 2020 from

<https://classroom.udacity.com/nanodegrees/nd088/parts/1bcfdbea-d4b3-45ba-bcee-f6532d0f83f8/modules/9ff8da38-5b15-4e31-96fe-fb5df5ff85e9/lessons/efc39a55-6ffd-4fb1-a07c-1a2f7b6ff2c7/concepts/ed39c34a-c21e-455f-a17e-215f2dc9ecf9>

Precision and Recall. Udacity. Retrieved December 14, 2020 from

<https://classroom.udacity.com/nanodegrees/nd088/parts/c4b3af94-0d10-4106-a44e-f3015a582b8b/modules/66bec2b6-8cf8-49db-8817-4cade52cccae/lessons/14cec49a-6372-4999-bbb3-1dec3f311937/concepts/416ad453-2af6-4479-b6d9-521600303a82>