RESULTS

Each of the three models were optimized using 100 trials on a specified search space and determined hyperparameters are shown in *Table 1.1*. Statistics for the optimized runs are shown in *Tables 1.2* and *1.3*. Tests regarding the 20-fold cross validation runs are displayed in *Tables 2.1-2.4*.Tests regarding the packet selection are displayed in *Tables 3.1-3.y*. *Figures 1.1-3.4* demonstrate the ROC and Precision-Recall curves of each of the models. *Figures 4.1-4.3* visually display the relationships between accuracy, loss, and training time against number of packets the model was trained on.

*Run-Time*

Most prominent is the difference in the time to optimize, with the MW-1D-CNN taking 15 hours longer than the 2D-CNN and 16 hours longer than the 1D-CNN. This difference was tested when performing 49 training trials on the models at 298 packets and was determined as significant by both one-way ANOVA (), the post hoc T-Test run between the 1D-CNN and the MW-1D-CNN (, , ), and the post hoc T-Test run between the 2D-CNN and the MW-1D-CNN (, , ).

*Model Performance*

Surprisingly, while the Multi Window model takes significantly longer to run than the sequential 1D Convolutional model, the sequential 1D model seemingly performs just as well in not only the optimized runs, but in the 20-fold cross validation tests (, , , see *Table 2.3*) as well. In the data restriction tests at all packet levels, the 1D-CNN outperforms the MW-1D-CNN slightly, but none the less significantly (, , , see *Table 2.2*). Compared to the 2D model, both of the 1D based models outperform in classification of IoT devices and discrimination against non-IoT devices across all packet lengths as determined by one-way ANOVA (), the post hoc T-Test between the 2D-CNN and 1D-CNN (, , ), and the post hoc T-Test between the 2D-CNN and MW-1D-CNN (, , ). This result correlates with the differences in, , and (*Table 1.3*) between the optimized models. The higher loss measure across all packet lengths tested implies that the 2D-CNN yields the most deviation between the predicted and actual result as determined by one-way ANOVA () and the post hoc T-Test between the 2D-CNN and 1D-CNN (, , ) This development is further supported by the lowest accuracy in the 20-fold cross validation test as determined by one-way ANOVA (, see *Table 2.1*), the post hoc T-Test between the 2D-CNN and 1D-CNN (, , , see *Table 2.4*), and the post hoc T-Test between the 2D-CNN and MW-1D-CNN (, , , see *Table 2.3*).

*Packet Restriction*

While model performance is important to the scope of this work, there is also a need for a focus on efficiency in the real world. By limiting the data used to train the models, run time is decreased (see *Figure 4.3*), yet there is no significant difference between accuracies on all differing packet lengths on the 2D-CNN () and MW-1D-CNN () as determined by one-way ANOVA. While one-way ANOVA determines there is a significant difference across all packet levels for the 1D-CNN (), there is no significant difference across packet levels 230 and above (). *Figure 4.1* shows the relationship between accuracy and number of packets used for training each model. These findings are validated by using the same comparative techniques with loss, which measures error, in lieu of accuracy. No significant difference was found between losses on all differing packet lengths on the 2D-CNN () and MW-1D-CNN () as determined by one-way ANOVA. While again one-way ANOVA determines there was a significant difference across all packet levels for the 1D-CNN (), there is no significant difference across packet levels above 230 (). *Figure 4.2* shows the relationship between loss and number of packets used for training each model