DISCUSSION

*Run Time*

This aforementioned difference in run time was expected due to the complexity of the models. As all models trained have identical linear layers applied after the convolutional layers, therefore the difference herein lies solely in the complexity and structure of the convolutional layers. The maximum convolutional layers of the 1D and 2D sequential models is limited to 4 layers with one convolutional operation per layer. The 1D Multi Window Model, while only having 1 convolutional layer, has at least 40 convolutional operations running parallel to each other, leading to a drastic increase in run time. Optimizing the model resulted in it having 80 convolutional operations. Compared to the 2 convolutional operations of the 2D-CNN and the 3 operations of the 1D-CNN determined by the hyperparameter optimization, it is suggested by the significant results that run time is strongly influenced by the complexity and number of convolutional operations performed while training the model.

*Model Performance*

While the MW-1D-CNN is the most complex model and takes the longest by far to train, it is incorrect to automatically deem it the highest performing model due to that factor. More important to model performance is the structure of the convolutional kernel used, whether its 1D or 2D. Different kernel structures look at the data differently when determining correlations to learn. The 1D kernel solely cares about the data within each feature of the packet separately, foregoing position of the data. The 2D kernel factors in position of the data as well, leading to potential missed correlations and learning opportunities due to the position of the data. The presented results showing both models based on the 1D convolutional kernel outperforming the 2D model suggest that the 1D kernels used effectively capture the dynamics of the IoT devices and can confidently discriminate them from non-IoT devices. These results also suggest the 2D architecture falters at both discriminating against non-IoT devices (more false negatives) and classifying IoT devices (less true positives). Between models on the 1D architecture, the less complex sequential model performs similarly compared to the multi-window model in the 20-fold cross validation test. In the packet restriction test, the sequential model outperforms the multi-window model across all packets tested. This strongly suggests that kernel structure factors into model performance more than model complexity, which the more complex counterintuitively led to a small decrease in performance in one case.

*Packet Restriction*

Real world implementation of an IoT classifier requires not only high accuracy, but also necessitates the most efficient training time in order to adapt to changes in the IoT landscape occurring each day. By restricting the data provided to the model, run time is reduced. While reducing data given to the model logically seems to carry a tradeoff in reducing accuracy, that tradeoff does not occur within the tested packet range for the 2D-CNN and MW-1D-CNN, and only occurs in the 1D-CNN once packets are reduced to below 230. By retaining performance while reducing packets, it is suggested that there is no tradeoff for accuracy when reducing packets until a certain critical value is reached. While testing did not reveal the critical value for the 2D-CNN and the MW-1D-CNN, the critical value for the 1D-CNN was determined to be 230, meaning that the 1D-CNN retains its max performance until it is restricted to less than 230 packets. As these results were validated by analyzing the loss (error) of the model, it is strongly suggested that the model’s efficiency (shortest training time yielding maximum model performance) is maximized at the critical value.

CONCLUSIONS AND IMPLICATIONS

With the continuous adoption of the IoT paradigm ongoing in this world, security and privacy concerns can lead to devastating consequences. This work complements ongoing research on IoT devices by optimizing efficiency of a generic, passive method to classify internet wide IoT devices. By optimizing multiple binary classifiers and cross evaluating them, the work is able to provide a pathway to find the most efficient model for classification. This pathway is not only created by this work’s suggestion of the most efficient model, but also key discoveries regarding the behavior of models regarding IoT-centric data.

Cross evaluation of the accuracies of the models following analysis of their run time revealed the structure of the model’s convolutional kernel impacts model performance significantly while model complexity has no effect on model. This reasoning was suggested due to, the 2D-CNN being outperformed by both 1D models, both 1D models performing similarly, and the fact that 2D structure factors in position of the data. This seemingly suggests the 2D structure miss correlations leading to a lower accuracy and higher error when predicting IoT devices, in line with my initial assumptions. This development contributes to the aforementioned pathway by suggesting the best kernel structure to use for internet wide IoT classification. The 1D kernel is strongly suggested as it outperforms models based on the 2D architecture and throws out position of the data, only factoring the data for each packet fed to the model.

Building on model performances, efficiency was also found to be a critical factor in determining the best overall model to use for IoT classification. It was first determined that reducing packets would decrease run time, yet the initial assumption that there would be a tradeoff of accuracy was only partially true. While a tradeoff does exist when decreasing packets given to the model, it does not take effect until a certain critical value is reached and packets are reduced past that value. Due to this critical value, it is suggested that a model performs at its maximum performance until data is reduced past the critical value and efficiency is maximized at the critical value. This allows for models’ data input to be reduced to a point where it still performs at its maximum while taking the shortest time to train. This contributes to the discussed pathway by suggesting a baseline critical value for the 1D-CNN at 230 packets and showing the critical value for the 2D-CNN and MW-1D-CNN lies below the 200 packet minimum tested.

By analyzing accuracy and efficiency of all models, it can be determined the 1D-CNN is the most efficient model to implement for IoT classification. As accuracy is significantly important in IoT classification, the 2D-CNN was removed from consideration as it performs significantly worse than the 1D-CNN and MW-1D-CNN. While the 1D-CNN outperforms the MW-1D-CNN across all packet levels tested, there was no significant difference in performance in the 20-fold cross evaluation test. Due to the likelihood there was no difference in performance between both 1D models, efficiency and run time of the model was also tested. While the 1D-CNN has a higher critical value than the MW-1D-CNN, the 1D-CNN takes significantly less time to train, meaning it achieves the same performance as the MW-1D-CNN while taking less time, suggesting the 1D-CNN has the highest efficiency. While it was initially assumed the MW-1D-CNN would be the best model due to its complexity intuitively leading to a higher accuracy, the suggestion that kernel structure affects model performance more than complexity explains why the 1D-CNN performed as well as its more complex counterpart. As there was no difference in accuracy and performance, the less complex model is the most efficient as a result of its lesser training time, further supporting the final suggestion of the 1D-CNN being the best and most efficient internet wide IoT classifier.