# CONCLUSION

In this work, the detection of electricity theft in smart grids was investigated using time-domain and frequency-domain features in a DNN-based classification approach. Isolated classification tasks based on the time-domain, frequency domain and combined domains features were investigated on the same DNN network. Widely accepted performance metrics such as recall, precision, F1-score, accuracy, AUCROC and MCC were used to measure the performance of the model. We observed that classification done with frequency-domain features outperforms classification done with time-domain features, which in turn is outperformed by classification done with features from both domains.

The classifier was able to achieve 87.3% accuracy and 93% AUC-ROC when tested. We used PCA for feature reduction. With 7 out of 20 components used, the classifier was able to achieve 85.8% accuracy and 92% AUC-ROC when

tested. We further analyzed individual features' contribution to the classification task and confirmed with the MRMR algorithm the importance of frequency-domain features over time-domain features towards a successful classification task. For better performance, a Bayesian optimizer was also used to optimize hyper parameters, which realized accuracy improvement close to 1%, on validation. Adam optimizer was incorporated and optimal values of key parameters were investigated.

In comparison with other data-driven methods evaluated on the same dataset, we obtained 97% AUC which is 1% higher than the best AUC in existing works, and 91.8% accuracy, which is the second-best on the benchmark. The method used here utilizes consumption data patterns. Apart from its application in power distribution networks, it can be used in anomaly detection applications in any field. Our work brings a small contribution towards accurately detecting energy theft as we detect theft that only took place over time. We wish to extend our method to detect real-time electricity theft in the future. Since this method was evaluated based on consumption patterns of SGCC customers, it can further be validated against datasets from different areas to ensure its applicability anywhere.