

水使用量データのディープラーニング手法による解析

研究分担者 大塚玲 情報セキュリティ大学院大学・教授

複数の異なる家庭の単一区間波形(開栓から閉栓までの流量波形)について用途を普遍的に識別する 1d-CNN に基づく機械学習モデル(単一区間波形識別器)を開発した。単一区間波形の特徴は、個人差、家庭差が大きいため、訓練データの過学習が起きやすく、一つの家庭を対象にしたときのような精度を出すことが難しい。アルゴリズムの改良を重ねたが、単一区間波形識別器は 9 クラス分類の rank1 識別精度で 60%程度の精度にとどまった(図1、図2)。更なる精度向上のためには、誤りデータの特徴を解析し、更なるアルゴリズムの改良が必要である。

また、社会実装のためには、現在のような 1 秒のサンプリング周期では、電池容量やデータ転送の問題が生じるため、現実的ではない。粗い粒度でも実用的な精度が得られれば活用可能性が広がるため、流量センサーのサンプリング周期を 1 秒から 20 秒までを想定してシミュレーション実験により、精度の劣化度を測定した。その結果、粒度に対して指数関数的に劣化する結果が得られた(図3)。社会実装のための現実的な粒度で一定の精度を実現するには、技術的に越えなければならない課題がまだ多く残る結果となった。

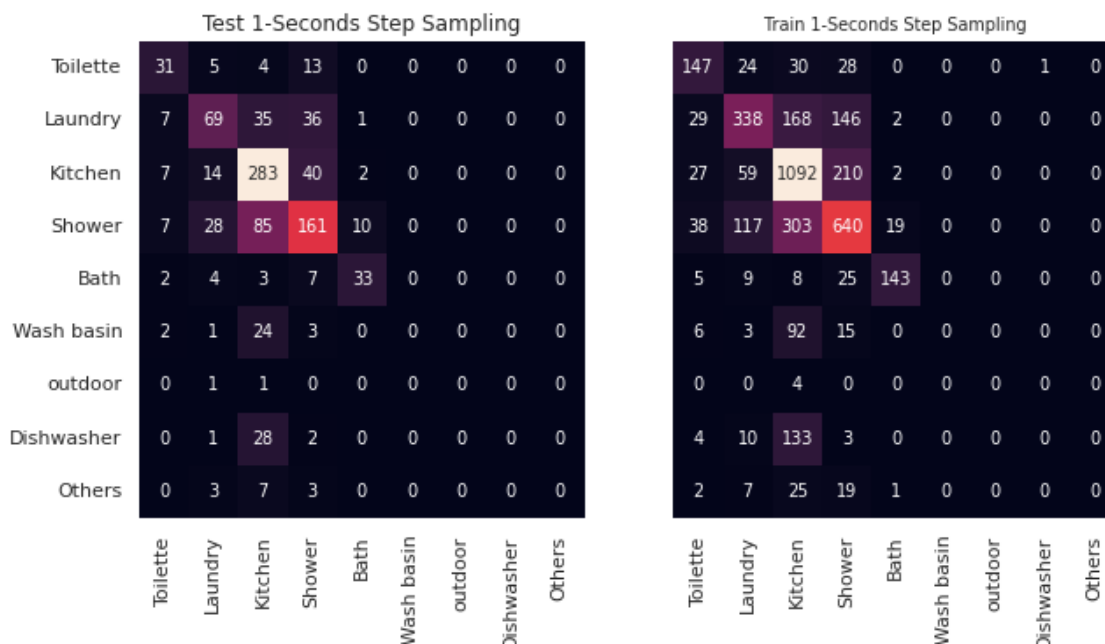


図1 用途別識別器の Confusion Matrix

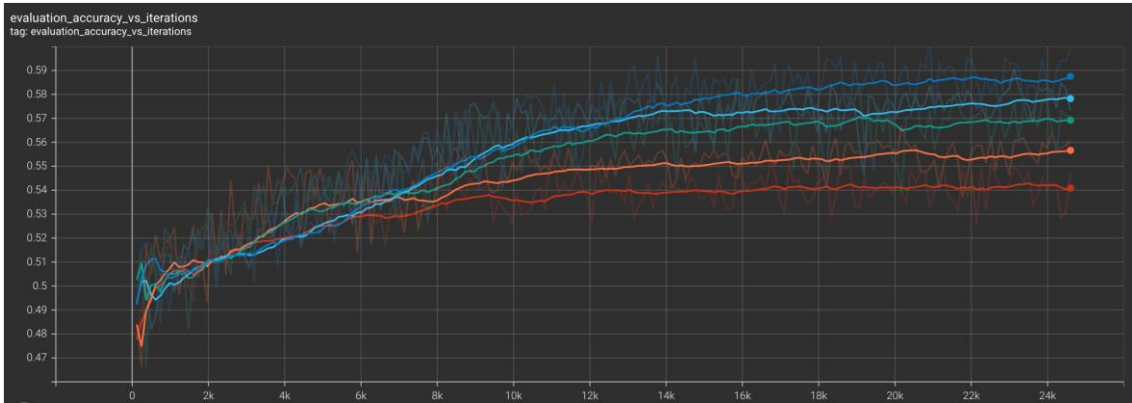


図2 単一区間波形識別器の精度

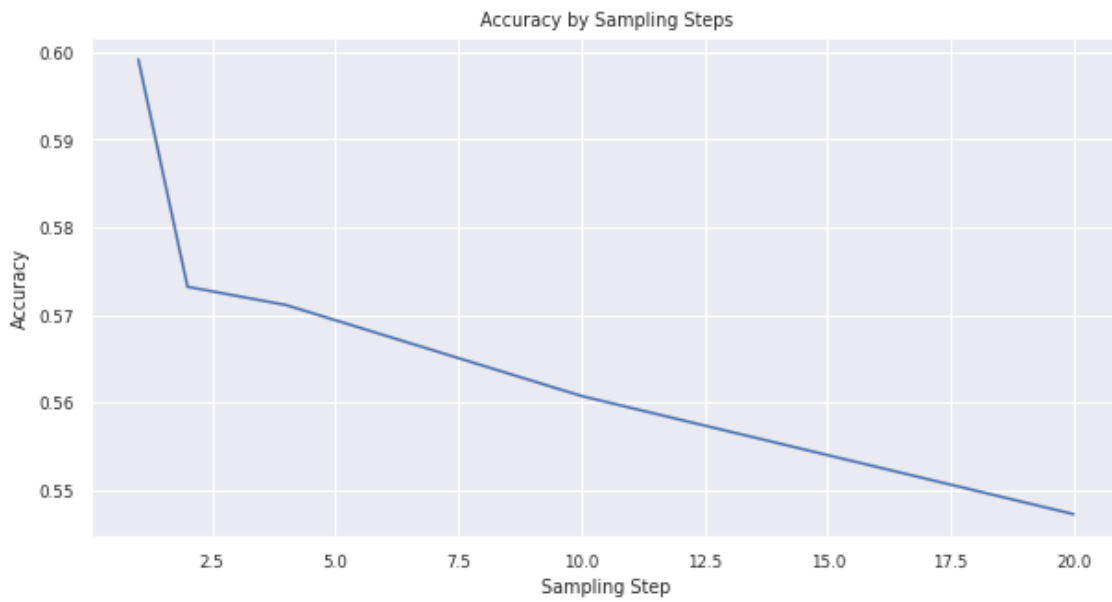


図3 精度の劣化度のシミュレーション