

IoT機械学習システム開発レポート

学生番号: 35714121

氏名: 福富隆大

総作業時間: 6時間

自己評価: A

● タスク1：擬似データセット改良

自己評価: A

■ 作業時間

3時間

■ 課題の概要

元の理想的な擬似データセット（00_sample_pseudoDataset.py）を改良し、現実的で高難易度のデータセットを生成する。環境相関、ノイズ、外れ値等を導入して実用的な機械学習環境を再現し、分類性能への影響を評価する。

■ 課題の結果（考察を含む）

改良版データセットの現実化手法

- **不均衡クラス分布:** comfortable=400, hot=200, cold=200
- **環境相関の導入:** 温度と湿度の逆相関（温度1°C上昇で湿度1.5%低下）
- **測定ノイズ:** 温度±0.3°C、湿度±1.5%、気圧±1.0hPa
- **外れ値:** 2%のサンプルで極端値（センサー異常・極端環境を模擬）
- **ラベルノイズ:** 8%の誤ラベル（主観的判断の個人差を再現）

分類精度比較

- **元データセット:** 98.9%（理想的だが非現実的）
- **改良版データセット:** 57.6%（現実的な困難度）
- **ランダム分類:** 33.3%（3クラス均等）

考察

改良版は現実のIoT環境の複雑さを反映し、実用的なモデル性能を正確に評価可能。精度低下は現実の不確実性を表しており、実際の展開時の期待性能により近い。

● タスク2：ハイパーパラメータ最適化

自己評価: A

■ 作業時間

3時間

■ 課題の概要

改良されたデータセットに対して、MLPの構造（層数、ニューロン数、エポック、バッチサイズ）を体系的に最適化し、M5GO制約内で最高性能を達成する構成を発見する。グリッドサーチによる全探索を実施し、元モデルとの性能比較を行う。

■ 課題の結果（考察を含む）

ハイパーパラメータ探索方策

- **探索手法:** グリッドサーチ (63組み合わせ)
- **探索範囲:** 層構成7種類、エポック[25,35,45]、バッチサイズ[8,16,32]
- **制約条件:** M5GO TensorFlow Lite対応 (<50KB)

最適パラメータ構成

- **最適構成:** 隠れ層[12]、エポック45、バッチサイズ8
- **精度結果:** テスト精度99.44%、検証精度100%
- **モデルサイズ:** 87パラメータ、TFLiteファイル2.04KB

精度比較結果

- **元モデル（標準化後データ）:** 98.9%
- **最適化モデル:** 99.44%
- **改善効果:** +0.54ポイント（標準化データ比較）

考察

1. **最適化効果:** ベースライン比+0.54ポイントの微改善を達成
2. **効率性向上:** パラメータ数19%削減でより軽量なモデルを実現
3. **前処理の重要性:** データ標準化が最も重要な性能改善要因であることが判明
4. **実用性:** M5GO制約内で高精度・軽量化を両立