

モデルの信頼性検
証と学習挙動の
解析

モデルの信頼性検
証と学習挙動の
解析

検証手法の採用
理由

k-分割交差検証と
グリッドサーチの
組み込みフロー

検証結果

進捗報告

佐藤 力

富山県立大学
u220029@st.pu-toyama.ac.jp

December 26, 2025

検証手法の強化：K-分割交差検証

データを k 個に分割し、学習と検証を k 回繰り返して精度の平均を算出する機械学習のモデル評価手法

k 分割交差検証 (k=5 の場合)

	[0]	[1]	[2]	[3]	[4]
1 回目	V	T	T	T	T
2 回目	T	V	T	T	T
3 回目	T	T	V	T	T
4 回目	T	T	T	V	T
5 回目	T	T	T	T	V

凡例：

V 検証データ T 訓練データ

ハイパーパラメータ

学習を始める前に人間が手動で設定しなければならない値

ハイパーパラメータグリッドサーチ

ハイパーパラメータの最適な組み合わせを見つけるための最も基本的かつ
確実な手法

	プルーニング率 0.05	プルーニング率 0.1
$\alpha = 0.01$	実験①	実験②
$\alpha = 0.1$	実験③	実験④

図 2: グリッド検証.png

研究的意義

たまたま設定した値で実験するだけでは、その手法の本当の実力がわからないため。グリッドサーチの結果を比較することでパラメータの影響度合いについて考察できる。

「不安定な手法」の安定性を証明できる

「ターミナルアトラクタ」と「プルーニング」は、どちらも数学的・構造的に不安定になりやすい性質を持っているため。

限られたデータセット（WikiText-2）の有効活用

比較的小規模なデータセットを使用している。データを回転させて使うため、最終的にはデータセット内のすべてのデータを「評価」として使用することになります。これにより、限られたデータ数の中で、統計的に最も信頼性の高い評価スコア（平均値と分散）を得ることができる。

k-分割交差検証とグリッドサーチの組み込み フロー

5/6

学習フロー

- ステップ 1: パラメータの組み合わせ作成
ターミナルアトラクタの係数とプルーニング率のリストの作製「外回りのループ」
- ステップ 2: データの K 分割
学習データを 3 つに分割「内回りのループ」
- ステップ 3: 反復学習
- ステップ 4: 最終ランキングと選定
全ての組み合わせの「平均スコア」を比較し、最も精度の高かったパラメータ設定を特定

モデルの信頼性検証と学習挙動の解析

モデルの信頼性検証と学習挙動の解析

検証手法の採用理由

k-分割交差検証とグリッドサーチの組み込みフロー

検証結果