

簡易脳波計で取得した脳波データのデータ量と機械学習のパフォーマンスの関係についての検討

Examination of the relationship between the amount of EEG data acquired by a Low-cost EEG and the performance of deep learning

伊藤遼哉^{*1}, 秋元頼孝^{*1}, 中平勝子^{*1}
Ryoya ITO^{*1}, Yoritaka AKIMOTO^{*1}, Katsuko T. NAKAHIRA^{*1}
^{*1}長岡技術科学大学
^{*1}Nagaoka University of Technology
Email: s191010@stn.nagaokaut.ac.jp

あらまし：脳波データは学生の思考状態を反映しており、効率良い状態判別によって教育への活用が期待される。本稿は、脳波データの効率的な状態判別のため、機械学習によるパタン分析手法比較を行なった。ノイズ除去済みの脳波データを入力とし、その状態判別を深層学習 (Deep Neural Network), Support Vector Machine, Random Forest の3つの機械学習による二値分類に対して、学習に用いるデータ量とパフォーマンスの関係性を評価する。

キーワード：簡易脳波計, 機械学習, 深層学習, 二値分類

1. はじめに

脳波データを学生の状態判別に用いることによって、問題の理解度を脳波から読み取り適切な難易度の問題を提示するなどの教育への活用が期待される。

脳波データの状態判別には、未知のデータから正解を予測可能な、機械学習によるパタン分析が有効であると考えられる。一般的に、機械学習はサンプル数が多いほど高い精度で出力できるとされているが、どの程度必要であるかはデータの性質や機械学習の手法によって異なる。また、脳波計測には手間と時間がかかるため、大量のサンプルを取得できないケースも多い。以上を踏まえ、本稿は、14chの簡易脳波計で取得したデータを入力データとし、代表的な機械学習の手法である Deep Neural Network (以下 DNN), Support Vector Machine (以下 SVM), Random Forest (以下 RF) に対して学習に用いるサンプル数の違いによりどの程度パフォーマンスが異なるのかを比較検討した。一連の基礎的検討により、取得したサンプル数によって最適な機械学習の手法を選択する際の目安が得られると考えられる。

2. 使用データ

入力に用いたデータは、8名の大学生・大学院生がそれぞれ3回(うち1人は2回)、1回あたり160試行のピアノ演奏課題を行った時の脳波データである。ピアノ演奏課題では、まず3秒間の間にC4-C5(低いドから高いドまでの黒鍵音を除いた8音)の中から、ランダムな5音が聴覚提示された(課題音区間)。その後対応する楽譜が画面上に提示され、実験参加者は対応するキーボードのキーを押した(演奏区間)。脳波は、サンプリングレートが128Hzで14chの簡易脳波計 Emotiv X を用いて計測した。

得られた脳波データについて、課題音提示後の6秒間(課題音区間3秒+演奏区間3秒)の波形を切り出し、1Hzのハイパスフィルタをかけた。次に、Artifact Subspace Reconstruction⁽¹⁾を用いてノイズの除去を行い、さらに独立成分分析を用いて目視で眼球運動などのノイズを除去した。これらによっても補正しきれない大きなノイズを含むデータを削除した結果、5738サンプルのデータが得られた。うち半分が課題音区間の脳波であり、残り半分が演奏区間の脳波である。その特徴量の数は、14個の電極から128Hzのサンプリングレートで3秒間測定しているため $128(\text{Hz}) \times 3(\text{s}) \times 14(\text{ch}) = 5376$ である。

脳波データを得る際に、条件ごとに同じサンプル数をそろえることができない可能性が考えられるため、入力データは課題音区間・演奏区間に等分せずランダムに選択した。

ピアノ演奏課題一回ごとに、ノイズによって削除されたサンプル数が異なるため、被験者毎のサンプル数も異なる。今回の実験ではサンプル数をそろえる必要があるため、被験者毎に訓練データとテストデータに分割するのではなく、ランダムに分割した。

この脳波データセットを訓練データとテストデータに7:3の割合で分割し、機械学習によってピアノを聞いている時(課題音区間)と弾いている時(演奏区間)の二値分類を行った。

3. 機械学習モデル

【DNN】実験に使用したDNNモデルはpytorchを用いて構築した、入力層(ニューロン数:5376)1層、中間層(ニューロン数:256, 512)2層、出力層(ニューロン数:2)1層の計4層から構成される。活性化関数は中間層2層ともSigmoid関数に設定し、出

力層では Softmax 関数を使用している. Sigmoid 関数は入力を 0~1 の間の値に変換して出力するという性質があり, 二値分類に適しているため中間層の活性化関数として採用した. Softmax 関数は分類問題に用いられることが多く, 出力値の合計が 1.0 (100%) になる性質があることから採用した. 学習係数は 0.0001 に設定した. 中間層は 2 層ともドロップアウト率 0.5 でドロップアウトを行っている. ドロップアウトをすることでモデルの過学習抑制が期待できる⁽²⁾.

【SVM】実験に使用した SVM モデルは sklearn を用いて構築した. rbf カーネル, $C = 1$, $\text{gamma} = 1/5376$ に設定した.

【RF】実験に使用した RF モデルは sklearn を用いて構築した. $n_estimators$ (決定木の数) は入力データ数と同じに設定し, $random_state$ を 0 とした.

4. 実験方法

入力データは sklearn の StandardScaler によって平均 0, 分散 1 の範囲に標準化した. 5738 サンプルの中から, 入力データをランダムに取得し, 100, 200, 500, 1000, 2500, 5738 の 6 パターンのデータ数 (データ取得後 7 : 3 の割合で訓練データとテストデータに分割) で各 10 回ずつ結果を出力した. 各サンプル数で DNN, SVM, RF のテストデータの正解率と F 値の平均を比較した. F 値は

$$F = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

(但し, *Recall* は再現率, *Precision* は適合率)

で定義される.

5. 結果

入力データ数による各手法の正解率の平均を図 1 に示す. データ数が 100, 200 と少ない場合は DNN の正解率が最も高く, 最も低い SVM より 12.3% 高い. データ数が多くなると逆に DNN の正解率が最も低く, 最も高い RF より 2.0% 低い.

入力データ数による各手法の F 値の平均を図 2 に示す. データ数が 100, 200 と少ない場合は DNN の F 値が最も高く, 最も低い SVM より 27.2% 高い. データ数が多くなると正解率と同様に DNN の F 値が最も低くなり, 最も高い RF より 2.1% 低い.

6. 考察

取得できたデータ数によって機械学習の手法を使い分けることによって学生が問題を理解しているか等の状態判別の精度が上がり, 問題の理解度を脳波から読み取り適切な難易度の問題を提示するというような教育への活用へつながるのではないかと考えた. このとき, 状態判別の精度がデータ数が小さい時の 60% 程度では, 適切な難易度の問題を提示

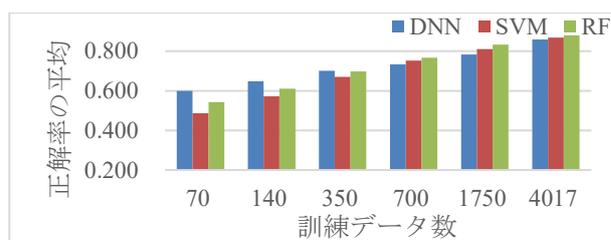


図 1 入力データ数による正解率の平均

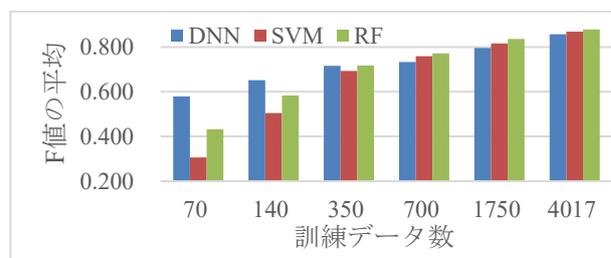


図 2 入力データ数による F 値の平均

するためのものとして不十分であると考えられる. 状態判別の精度が約 80% 程度であれば, 適切な難易度の問題を提示する上での一つの指標として扱えると考えられるため, 訓練データ数は少なくとも 1750 必要であると考えられる.

SVM と RF は訓練データ数が 70, 140, 350 の時に, パフォーマンスが DNN より低くなっているが, それが単に訓練データが少ないことによるものか, パラメータチューニングの不足によるものかは不明である. 今後はパラメータチューニングを行った場合のパフォーマンスの変化について検証する必要がある.

学習に用いた入力データは全データを一つにまとめた後ランダムにデータを取り出したものであるが, 実際にはほかの人のデータで学習し, 対象の人に適用する場面が多いと考えられる. その場合のパフォーマンスの変化について今後検証する必要がある. 今回の実験では 14ch の脳波データを入力としたが, 電極数を多少減らしても機械学習の精度を維持できる可能性も考えられる. 今後は電極数を減らすことによるパフォーマンスの変化も検証して行きたい.

謝辞: 本研究は科研費(19K03026, 19K12246, 20H04731)の助成を受けたものである.

参考文献

- (1) Chang, C. -Y., Hsu, S. -H., Pion-Tonachini, L, Jung, T. -P. : "Evaluation of Artifact Subspace Reconstruction for Automatic EEG Artifact Components Removal in Multi-channel EEG Recordings", IEEE Transactions on Biomedical Engineering (2019)
- (2) Nitish Srivastava Geoffrey Hinton Alex Krizhevsky Ilya Sutskever Ruslan Salakhutdinov: Dropout: A Simple Way to Prevent Neural Networks from Overfitting Journal of Machine Learning Research 15 (2014) 1929-1958