

6T-07

精神課題における脳波による機械学習モデルの データ前処理手法の比較～感情と認知～

金井健太郎^{†1} 鈴木圭^{†1} 菅谷みどり^{†1}
芝浦工業大学^{†1}

1. はじめに

近年、感情推定や認知能力推定などを目的とし、生体情報を用いた機械学習モデルの構築の関心が高まっている[1]。これらのモデル構築においては、感情や認知の精神課題への回答や、診断結果を正解とし、脳波や心拍などの生体情報を機械学習させることで、推定や診断支援を行うことを目的としている[2][3]。こうした精神課題を用いた機械学習モデルの応用範囲は広く、医療診断の補助やヘルスケアのような医療分野等のみならず、ロボット分野やマーケティング分野等にも適応されており、その重要性はますます高くなっている。こうした生体計測から得られる値の多くは、人が恣意的に変更することが困難であり、リアルタイムに情報を取得できるといった性質から客観性が高く、応用性も高いことが特徴である[4]。

生体計測機器の中でも、Electroencephalogram(EEG)は、簡単に装着することができ、かつ人の精神状態を直接的に把握することができることから、汎用的な用途である感情分析や認知能力の推定に用いられている[4]。EEG から取得されるデータを用いた機械学習モデルの例として、感情推定[5]や認知能力推定[6]等が挙げられる。こうした精神課題では、それぞれの判定指標を目的変数とし、EEG の出力データを説明変数として用いることでモデルの構築を行う。このように、正解ラベルである目的変数を変更することで多様な精神状態に関するモデルを構築が可能である。しかし、同一の EEG データを用いる場合でも、モデルの構築手法が同一で良いとは限らない。例えば、認知レベルの推定と感情推定課題では、異なる脳の特性を用いることが予想される。このことから、同一データを用いた場合でも、特徴量の選択方法や、正規化などの方法が異なることが予測される。しかし、正解ラベルがモデルの精度にどの程度影響を与えているか、また、正規化などのデータの前処理手法の違いにより、重要となる特徴量の違いや、機械学習モデルの精度への影響は明らかになっておらず、どこを変更すべきか統一的で汎用的な議論は行われていない。

そこで本研究の目的を、脳波を用いて感情推定と認知レベル推定の正解ラベルが異なる精神課題におけるデータの処理手法による重要となる特徴量や機械学習モデルに与える影響の比較・検討を目的とする。

目的を実現するための提案として、本研究では説明変数に使用する脳波指標を複数の処理方法で感情推定モデルと認知レベル推定モデルの二つの機械学習モデルの構築を行う。具体的な方法として、複数種類の正規化を行い、それぞれの正規化手法と、正規化を行っていないもので精度の比較を行う。

2. 機械学習に用いるデータ・機械学習手法

2.1 データ取得に用いる機器

本研究では、脳波を取得する機器として Emotiv 社の EPOC X を用いた。EPOC X は 14 チャンネルから脳波の取得が可能であるため脳の広範囲をカバーすることが可能である。EPOC X で計測可能なチャンネルは、前頭葉 6 チャンネル、側頭葉 2 チャンネル、頭頂葉 4 チャンネル、後頭葉 2 チャンネルである。これらのチャンネルは、脳の意識的な思考や聴覚などの反応を測定することに役立つとされる[4][7]。

2.2 感情推定モデルのデータ取得

初めに、感情推定モデル構築のためのデータ取得を目的とし、20 代の男女 4 名を対象として実験を行った。実験手順は、鈴木らの手法に従った[2]。椅子に腰かけた状態で脳波計とイヤホンの装着を行う。次に実験前安静として 2 分間安静状態をとってもらい。その後、1 分 30 秒間の安静時間、2 分の音楽聴取、音楽聴取時の感情の主観評価を 1 試行として計 8 試行を行う。音楽聴取時による感情喚起では音楽はラッセルの円環モデルの各象限に属する音楽を 2 種類、ユヴァキュラス大学の音楽データベース[8]より選択し聴取してもらう。聴取の順番はランダムに決めた。安静時、音楽聴取時には外からの刺激の軽減のためにディスプレイ上に灰色の背景の中心に配置した黒い十字がある画面を見てもらうようにした。主観評価では、Self-Assessment Manikin[9]により 4 感情を取得した。

2.3 認知レベル推定モデルのデータ取得

次に、認知レベル推定のモデル構築のためのデータ取得を目的とし、20 代の男女 5 名を対象として実験を行った。実験手順では、認知レベルを計測するためのトレイルメイキングテスト(TMT:Trail Making Test)の練習を行なってもらう。練習の終了後、脳波計の装着を行い、安静時間として 5 分、TMT-A,B 認知負荷をかけるためのストループ課題 10 分、TMT-A,B 安静時間 5 分の順番で実験を行った。TMT とは決

^{†1} KENTAROU KANAI, KEI SUZUKI, MIDORI SUGAYA, Shibaura Institute of Technology

められた順番に数字や文字をなぞっていき、なぞり終わるまでの時間を計測し、3段階で評価する検査である。TMT-Aでは数字のみを「1→2→3→…」の順番でなぞっていき、TMT-Bでは数字と文字を交互に「1→あ→2→い→3→…」の順番でなぞる。ストループ課題とは認知負荷の一つであり、文字の意味と異なる文字の色で書かれた文字の意味を答える課題である。例えば赤色で「あお」と書かれていた場合、「あか」と答える。

2.4 データセット・機械学習手法

データセットは各実験で得られた脳波にノイズ処理を施した後、周波数解析によりチャンネルごとに指標の作成を行った。作成した指標を表1に示す。

表1 作成した脳波指標

指標	説明
δ	1-3Hzのパワーバンド
θ	4-7Hzのパワーバンド
α	8-12Hzのパワーバンド
β	13-30Hzのパワーバンド
γ	30-100 Hzのパワーバンド
Low- α	8-9 Hzのパワーバンド
High- α	10-12 Hzのパワーバンド
Low- β	13-17 Hzのパワーバンド
High- β	18-30 Hzのパワーバンド
Low- γ	31-40 Hzのパワーバンド
Mid- γ	41-50 Hzのパワーバンド
High- γ	51-100 Hzのパワーバンド
Total_Power(1-30)	1-30 Hzのパワーバンド
Total_Power(1-40)	1-40 Hzのパワーバンド
Total_Power(1-100)	1-100 Hzのパワーバンド

これらの210個の指標をそのまま使用したデータセット、正規化手法としてデータを0-1のスケールにするmin-Max正規化、平均を0、標準偏差を1にするZ-score正規化の2種類を行った。また、正規化についても個人ごとに正規化を行なったデータセットと、全部のデータをまとめた後に正規化を行なったデータセットの2パターンを用意した。

モデルについては、感情推定には音楽聴取時の脳波を説明変数、その際の主観評価を目的変数としてモデルを構築した。認知レベル推定には安静時の脳波を説明変数、TMT-Aの結果とTMT-Bの結果をそれぞれ目的変数とした2つのモデルを構築した。

機械学習については、ディープラーニングにてモデルの構築を行い、精度評価手法として層化10分割交差検証を使用した。手順としては、最初に、説明変数の比率が全体の比率と同じになるようにデータを10分割し、そのうちの1つをテストデータ、その他を訓練データとしてテストデータを入れ替えながら学習を行い、精度を検証する手法である。

3. 結果

感情と認知レベルそれぞれのモデルの精度を表2に示す。

表2 各手法の機械学習モデルの精度

	感情推定(%)	TMT-A(%)	TMT-B(%)
正規化なし	92.53	100.00	100.00
min-Max(Individual)	99.93	99.92	100.00
min-Max(all)	92.92	100.00	100.00

Z-score(Individual)	99.70	99.92	100.00
Z-score(all)	93.34	100.00	100.00

認知レベル推定モデルについてはTMT-A、B両方とも全ての手法において99%以上の結果となった。TMT-Aにて、「個人ごとに正規化を行なった場合は他の手法に比べてmin-MaxとZ-scoreの両方とも低い結果となった。しかし、感情推定においてはデータ全体を正規化するよりも個人ごとに正規化を行なった方が精度は高い結果となった。これは感情が喚起された際の個人差をデータのスケールを揃えることである程度抑えることができたと考えた。

4. 結論

本研究では、多チャンネル脳波計を用いて2つの精神課題を用いた機械学習モデルのデータの前処理手法の精度への影響を比較することを目的とし、データの取得、各手法のモデルの構築を行った。その結果、感情ではデータを個人ごとに正規化を行なった方が約7%精度は高くなった。理由として認知機能の測定では個人差の影響が少ないと考えられる。一方で認知レベルの推定においては正規化をしないでモデルの構築を行う手法や、データ全体の正規化を行う手法の方がモデルの精度は高くなった。今後は、次元圧縮や移動平均など今回試していない手法を使用した場合の精度検証を行う予定である。また、今回は実験協力者の数が少なかったため、この人数が増えた際にモデルの精度がどのように変化するのか調査していく。さらに、これらの手法を使用したときに各課題や前処理手法で重要となる特徴量がどのように異なってくるのか、それぞれの課題に適した前処理手法は何かあるのか検討をしていく。これらを明らかにすることにより、脳波データの活用を促進し、モデル構築の指針を明確にする。

参考文献

- [1] 厚生労働省.AI懇談会報告書.<https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkankouseikagaku-Kouseikagaku/0000169230.pdf>,2021年7月24日参照.
- [2] Suzuki Kei, Laohakangvalvit Tipporn, Matsubara Ryota, Sugaya Midori, "Constructing an Emotion Estimation Model Based on EEG/HRV Indexes Using Feature Extraction and Feature Selection Algorithms", Sensors, 21(9), 2910, 2021.
- [3] M. de Bardeci, C. T. Ip, and S. Olbrich, "Deep learning applied to electroencephalogram data in mental disorders: A systematic review," Biol. Psychol., vol. 162, p. 108117, May 2021.
- [4] Soraia M, et al, Emotions Recognition Using EEG Signals A Survey, IEEE Transactions on Affective Computing, pp 1- 12 June 2017.
- [5] Salma Alhagry, Aly Aly Fahmy, Reda A. El-Khoribi. Emotion Recognition based on EEG using LSTM Recurrent Neural Network. IJACSA. vol. 8, No. 10, 2017.
- [6] 柳澤 琢史,畑 真弘,福岡 良平,石井 良平,吉山 顕次,原田 達也,池田 学, 貴島 晴彦.安静時脳波・脳磁図に深層学習を用いた認知症診断.日本脳神経外科認知症学会誌,2021年1巻1号p1-6.
- [7] Konstantine K.Zakzanis,Richard Mraz,Simon j.Graham.An fMRI study of the Trail Making Test.Neuropsychologia,Volume 43,Issue13,2005,P1878-1886.
- [8] Tuomas Eurola, Jonna K. Vuoskoski. A comparison of the discrete and dimensional models of emotion in music. 2011.
- [9] Narumon Jadram, Tipporn Laohakangvalvit and Midori Sugaya, "A Preliminary Evaluation of Comfortable Arousal using Biological Information Measurement for Autonomous Driving", Asia Pacific Conference on Robot IoT System Development and Platform (APRIS2020).