

機械学習を用いた学習者の生体情報と心的状態の 関係性抽出の試み

田和辻 可昌*、宇野 達郎**、岡崎 桂太***、松居 辰則*

* 早稲田大学 人間科学学術院

** 早稲田大学 人間科学部

*** 早稲田大学 総合研究機構

Trial Study on Extraction of Relationship between Learner's Physiological Indexes and Mental States by Machine Learning

Yoshimasa TAWATSUJI*, Tatsuro UNO**, Keita OKAZAKI*, Tatsunori MATSUI*

* Faculty of Human Sciences, Waseda University

** School of Human Sciences, Waseda University

*** Comprehensive Research Organization, Waseda University

* y.tawatsuji@aoni.waseda.jp

概要: 教授・学習過程において学習者の心的状態を把握することは教育効果・学習効果の観点から極めて重要である。本研究では、機械学習を用いて、教師の発話および学習者の生理指標と心的状態の関係の抽出を試みた。学習者の生理指標については、NIRS、脳波、呼吸・皮膚コンダクタンス・容積脈波計を取得した。一方、学習者の心的状態は Achievement Emotions Questionnaire: AEQ を用いて、9つの感情状態で抽出された。本シミュレーションから、中間層を二層とする深層学習を用いることで、高い精度で教師の発話と学習者の生体情報から心的状態の推定が可能であることが示唆された。

Abstract: Estimation of learners' mental states during the interaction between teacher and learners is very important issues for teacher from quality of learning environment point of view. In this experimental study, relationship between teacher's utterances, behaviors, learner's physiological indexes and mental states were tried to be detected by Machine Learning Method. As for learner's physiological indexes, Near-Infrared Spectroscopy: NIRS, Electroencephalogram: EEG, Ventilatory Frequency, Skin Conductance Activity: SCA and photoplethysmogram: PTG were measured during her learning activities. On the other hand learner's mental states were reported by introspection method using the Achievement Emotions Questionnaire: AEQ constructed by some related scales on 9 emotions as Enjoy, Hope, Pride, Anger, Anxiety, Shame, Hopelessness, Boredom, and Other.

キーワード: IMS、生体情報、深層学習、感情推定

Keywords: Intelligent Mentoring System, physiological data, deep learning, estimation of learner's emotion

1. はじめに

1.1 背景と目的

教授・学習過程において学習者の心的状態を把握することは教育効果・学習効果の観点から極めて重要であ

る。教育工学研究においても、学習者の眼球運動や発汗量など生体情報を学習行為や心理状態と関係付けるための基礎的な研究は多くの知見を蓄積している。そして、昨今の計算機や生体計測機器の高機能化と低廉化によって、生体計測機から得られるリアルタイムかつ大量のデータを高速に処理することにより、生体情報や行動

情報を用いた学習者の心的状態の計算機による自動推定と教育支援への試みが盛んに行われている。一方、教育実践研究においては、教授・学習過程における教師と学習者のインタラクションは学習者の心的状態の変化に影響を及ぼし、学習効果の決定要因として重要であることは広く共有されているところである。したがって、学習時における教師の行動や発言と学習者の心的状態、あるいは心的状態の変化に関する要因との関係の形式化は重要な課題であり、その成果は学習支援システムへの学習者の心的状態の推定機能の実装のための基礎的な知見を与えることも期待される。

著者らは、教師と学習者のインタラクションにおいて教師の発話と学習者の生理データ、および学習者の心的状態との関係の形式化を試みてきている (Takehana 2016)。そこでは、生体情報、発話情報、内省報告等、学習に関わる情報を多面的に用い、相関ルールにより関係の形式化を試みている。一方、機械学習アルゴリズムの学習支援を含めて様々な分野への応用可能性も示されてきている (藤好 2015) (堀口 2010) (小島 2014)。そこで、本研究では、上記と同様の学習に関わる多面的情報から Deep Learning (以下、DL) を用いて学習者の心的状態の推定の可能性を検討した。

1.2 本論文の構成

まず、第二節では、本研究で試みた機械学習の学習用データの取得について行った実験について述べる。実験では、学習者の生体情報、教師による教授行動(発話行為)、および学習者の心的状態をそれぞれ取得した。第三節では、取得したデータセットを用いて、学習者の生体情報および教授行為を入力とした、学習者の心的状態を出力とする三層の古典的ニューラルネットワークを構築し、それぞれの写像関係に関する学習をおこなった。第四節では、第三節では検討されなかった各入力情報と出力情報に関する時間的観点を導入した上で、中間層をさらに増やした多層のニューラルネットワークを構築して、同様の写像関係に関する学習をおこなった。第五節では、それぞれの結果の考察を行い、第六節ではまとめと今後の展望について述べる。

2. 学習に関わる多面的情報の取得

学習に関わる多面的情報の取得を目的として生体計測機器を用いた計測を中心に実環境での実験を行った。被験者は個別指導塾(教師1名、学習者1名の個別学習)に通う中学生1名であった(本実験の実施にあたっては当該塾の講師を通して被験者(生徒)および保護者の許可を得ている)。使用した機材は Near-Infrared Spectroscopy (NIRS) (日立 WOT-100 製)、呼吸・皮膚コンダクタンス・容積脈波計 (NeXus) であった。被験者には上記の機材を全て装着してもらい、通常通りの授業を受けてもらった。各計測機器は計測時間を対応させるために計測開始、終了時にマーカーを付与した。得られたデータのうち、NIRS のデータに対して、大域平均基準化 (野澤 2009) (平山 2012) を施した。

実験中の様子は3か所から3台のビデオカメラで撮影した。この映像から、著者らが教師の教授行動、特に発話行為に焦点を当て、「説明」「指示・確認」など9種類のカテゴリに分類した。教師の発話を表すカテゴリは、先行研究 (藤江 2000) (岸 2006) (清水 2001) で使用されていたカテゴリを一部修正した9種類のカテゴリ (1:説明, 2:発問, 3:指示確認, 4:復唱, 5:感情受容, 6:応答, 7:注意, 8:雑談, 9:その他) を用いた。教師の発話へのカテゴリの付与に関しては、分析者が授業映像を見ながら分析者の視点で行った。また、被験者には後日実験で得られた映像を見ながら学習時の心的状態の内省報告を求めた。心的状態を表すカテゴリは、Achievement Emotions Questionnaire (AEQ) (Perkun 2011) で使用されている9感情についての尺度 (Enjoy, Hope, Pride, Anger, Anxiety, Shame, Hopelessness, Boredom, Other) を用いた。被験者にはアノテーション付与のための自作アプリケーション (自身の動画を再生しながら該当する心的状態のボタンを押すことによってその時の心理状態を内省報告する) を用いて授業時の心的状態の内省報告を求めた。

3. 三層ニューラルネットワークを用いた心的状態推定の試み

2節で取得したデータを用いて三層ニューラルネットワ

ークを用いた生体情報からの心的状態の推定を試みた。今回分析の対象としたのは、約 60 分授業の中で教師と学習者のインタラクションが比較的多く確認できた 63 秒（開始後 19 分 37 秒から 20 分 40 秒まで）であった。

3.1 データ構造とネットワーク構造

入力データは、①NIRS によって取得した脳血流量 (5Hz)、②呼吸 (32Hz)、③皮膚コンダクタンス (32Hz)、④教師の発話 (上記の 9 カテゴリのうち「説明」「発問」「指示・確認」「注意」「その他」) であった。出力データは、⑤内省報告 (上記の 9 カテゴリ) であった。なお、①②③は粒度を統一するために、粒度の荒いデータに対して、粒度の最も細かいデータの粒度に合わせて線形補完を施した。その結果、最終的な対象データ数は、2024 行×5 (「NIRS」「皮膚コンダクタンス」「呼吸」「教師の発話」「心的状態」) となった。また、入力データのうち、①②③の数値データは、それぞれカテゴリ内で標準化 (平均 0、分散 1) をおこなった。

ネットワーク構造は、入力層・中間層 (1 層)・出力層の 3 層ニューラルネットワークとした。中間層のユニット数はシミュレーションの結果 19 とした。これは、中間層のユニット数を 1 から 25 まで変化させたところ、ユニット数 19 において、学習の Loss 値が最小、Accuracy 値が最大となったためである。この結果を図 1 に示す。



図1 中間層ユニット数による Loss 値、Accuracy 値の変化

3.2 シミュレーション方法と結果

Python3.5, Tensorflow (ver 0.12.1) (Tensorflow) で実装した。活性化関数は中間層では tanh 関数を、出力層ではソフトマックス関数を用いた。また、損失関数にはクロスエントロピー誤差関数、オプティマイザには Gradient Descent (最急降下法) を用いた。学習率は 0.05 とした。対象データ (2049 データ) を 6:4 で分割し、6 割のデータ

を学習データ、4 割のデータを評価データとして 10 回の交差検定 (Cross-Validation) を行った。また、学習にあたっては 5000 回繰り返し学習させた。学習では交差エントロピーが最小になるように学習させた。その一部を図 2 に示す。早い段階で学習が収束していることがわかる。

また、10 回の交差検定において Accuracy を算出 (小数点第 4 位で四捨五入) した結果、「0.900, 0.917, 0.910, 0.917, 0.910, 0.914, 0.915, 0.899, 0.912, 0.905」となった。このことから、非常に高い精度で生体情報から心的状態の推定が行われていることがわかる。

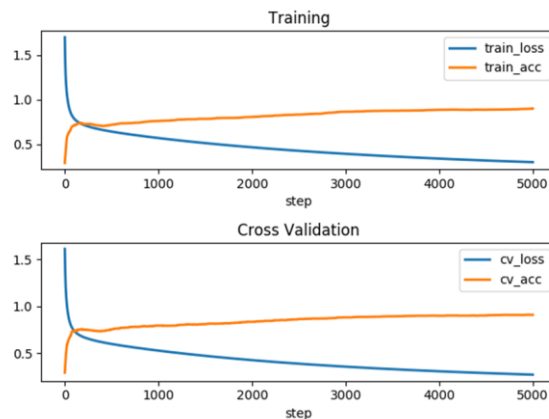


図2 三層 NN におけるシミュレーション結果

4. DL を用いた心的状態推定の試み

3 節で説明した三層ニューラルネットワークでは、時間的側面が考慮されておらず、ある一つの単位時刻における生体情報と心的状態の関連性について議論されてきた。そこで、この点を考慮したうえで、中間層を多層化することで推定の精度向上が認められるかを検討した。

4.1 リアルタイム性処理と時間的側面の検討

前節では、ある一つの単位時刻における生体情報と心的状態に関する入出力の写像関係を三層のニューラルネットワークを用いて学習させた。ところが、このニューラルネットワークでは、以下の二点において課題が残る。一点目は、応用的側面を考慮すると入力データの標準化を行なうことが難しいという点である。より具体的には、IMS などのリアルタイム性の高い処理が求められる状況では、NIRS などから得られた情報に対して大域平均基準化や、生体情報に対する標準化を適用することが難しい。したがって、IMS を検討する上では、リアルタイムに

入力される NIRS などの生体情報データをそのまま使用する必要がある。二点目は、生体情報の表出および教師の発話があった時刻から心的状態の認知に至るまでの時刻に存在すると考えられる時間遅れが検討されていない点である。そこで、入力から出力に至るまでの時間遅れを考慮した入出力の関係を学習する必要がある。

4.2 データ構造とネットワーク構造

入力データは 3 節と同様、インタラクションが多く確認できた 63 秒間に存在する①NIRS、②呼吸(32Hz)、③皮膚コンダクタンス(32Hz)、④教師の発話であった。ただし、DL での学習においては、生体情報に関するデータの標準化は行われていない。また、出力データは 3 節と同様に⑤内省報告であった。

ネットワーク構造は、入力層・中間層(2 層)・出力層の 4 層ニューラルネットワークとした。時間遅れを表現するために、入力層では時刻 $t - \tau$ から時刻 t までの①②③④のデータを、また、出力層では時刻 $t - \tau$ の心的状態を出力することで時間遅れを表現した(図 3)。尚、中間層のユニット数は中間層 1 と中間層 2 のユニット数を変化させながらシミュレーションを行い、Loss 値と Accuracy 値を比較することで、一層目 19 次元、二層目を 20 次元と決定した。図 4 に、中間層 2 のユニット数を 20 として、中間層 1 のユニット数を段階的に変化(10~19)させながら、シミュレーションを行った際の Loss 値と Accuracy 値を示す。また、本シミュレーションでは、3 および 10 ステップ、つまり $\tau = 2, 9$ でシミュレーションを行ったが、 $\tau = 2$ のときは、中間層 1 のユニット数を変化させても、Accuracy が、 $\tau = 9$ のときを超えることがなかった。したがって、以降本シミュレーションでは、 $\tau = 9$ の場合について述べる。

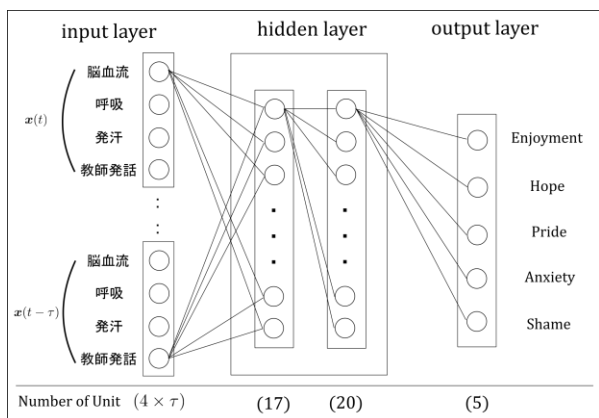


図 3 ネットワーク構造図

4.3 シミュレーション方法と結果

Python3.5、Tensorflow(ver 0.12.1) (Tensorflow) で実装した。中間層の活性化関数は二層とも tanh 関数を、出力層ではソフトマックス関数を用いた。また、損失関数にはクロスエントロピー誤差関数、オプティマイザには Gradient Descent (最急降下法) を用いた。学習率は 0.05 とした。

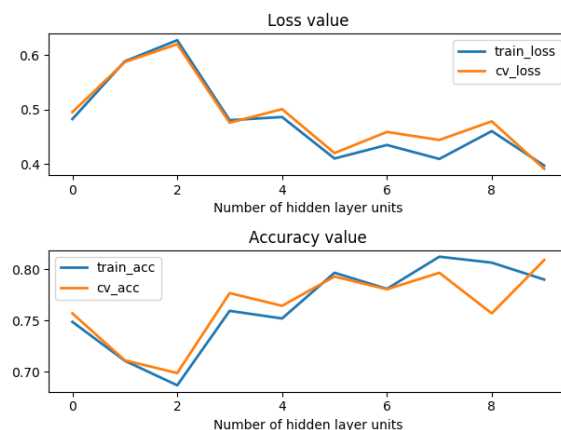


図 4 中間層 1 のユニット数を 10 から 19 まで段階的に変化させたときの Loss 値と Accuracy 値

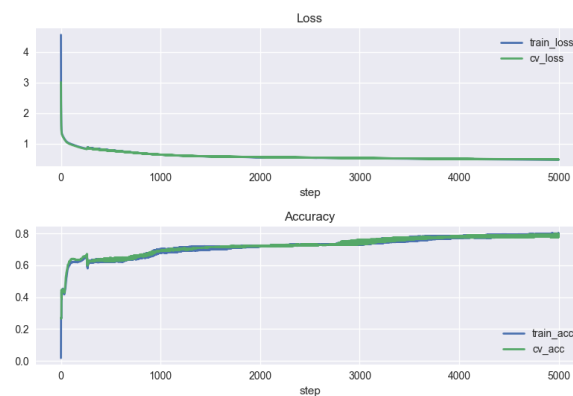


図 5 5000 回学習時の Loss 値と Accuracy 値の変化

中間層 1 のユニット数 19、中間層 2 のユニット数 20 でシミュレーションを行った。対象データ(2049 データ)を 6:4 で分割し、6 割のデータを学習データ、4 割のデータを評価データとして 10 回の交差検定(Cross-Validation)を行った。また、学習にあたっては 5000 回繰り返し学習させた。学習では交差エントロピーが最小になるように学習させた。その一部を図 5 に示す。早い段階で学習が収束していることがわかる。

また、10 回の交差検定において Accuracy を算出(小

数点第4位で四捨五入した結果,「0.799, 0.842, 0.797, 0.793, 0.764, 0.811, 0.789, 0.829, 0.768, 0.808」となり、平均0.8、標準偏差±0.023となった。このことから、入力データを標準化していなくとも8割程度の精度で生体情報から心的状態の推定が可能であることが示唆された。

5. 考察

DLを用いた本シミュレーションから、標準化していない生体情報から8割近くの心的状態が推定で可能であることが実験的に確認された。このような高い推定可能性の要因として、本シミュレーションで取り扱った該当シーンに存在する感情状態が5つの状態と、推定する対象が少なかったこと、が考えられる。実際に、シミュレーションの学習に伴うAccuracyの値は極めて速く収束している。すなわち、今回のシミュレーションでは、過学習が起きている可能性が否定できない。

6. まとめと今後の課題

本研究では、学習に関わる多面的情報から学習者の心的状態の推定を試みた。具体的には、DLを用いて学習時の生体情報と心的状態の関係のシミュレーションを行った。この結果として、計測機器から得られるデータに対して標準化などの処理を行わずとも、8割程度心的状態が推定されることがシミュレーションから示唆された。一方で、本シミュレーションでは、学習データにおける感情状態の少なさから、過学習が起きていることが示唆された。したがって、今後は今回の63秒のデータに加え、実験によって取得した全体のデータを使用することも検討する必要があると考えられる。また、今回のDLによる学習では中間層の層数とユニット数が深く吟味されていない。今後は様々な中間層のパターンを検討した上でネットワーク構築が望まれる。さらに、今回の結果から機械学習による心的状態の推定の可能性が示唆されたが、学習・教育文脈でのデータの解釈には因果関係に関する説明が必要である。この観点からは、DLの中間層や重みの状態の可視化と解釈が重要な課題である。

参考文献

- 岸俊行 & 野嶋栄一郎. (2006). 小学校国語科授業における教師発話・児童発話に基づく授業実践の構造分析, 教育心理学研究, 54(3), 322-333.
- 小島一晃、村松慶一 & 松居辰則. (2014). 多肢選択問題の回答における視線の選択肢走査の実験的検討, 教育システム情報学会誌, 31(2), 197-202.
- 藤江康彦. (2000). 一斉授業における教師の「復唱」の機能: 小学5年の社会科授業における教室談話の分析, 日本教育工学雑誌, 23(4), 201-212.
- 藤好宏樹, 吉村和代, K. Kunze & 黄瀬浩一. (2015). 英文問題解答時の視点情報を用いた英語能力推定法. 信学技報, 115(24), 49-54.
- 平山健太, 綿貫啓一, 楓和憲. (2012). NIRSを用いた随意運動および他動運動の脳活動分析, 日本機械学会論文集C編, 78(795), 3803-3811.
- 堀口祐樹、小島一晃 & 松居辰則. (2010, March). MRAを用いた学習者のLow-Level Interaction特徴からの行き詰まりの推定手法, 第58回人工知能学会先進的学習科学と工学研究会資料, A903, 1-6.
- 野澤孝之 & 近藤敏之. (2009, September). NIRS脳計測データのオンライン分析のためのアーティファクト除去手法の比較, 計測自動制御学会生体・生理工学シンポジウム論文集, 24, 381-384.
- Pekrun, R., Goetz, Frenzel, A. C., Barchfeld, P. & Perry, R. P. (2011). Measuring Emotions in Students' Learning and Performance: The Achievement Emotions Questionnaire (AEQ), Contemporary Educational Psychology, 36(1), 36-48.
- 清水由紀 & 内田伸子. (2001). 子どもは教育のディスコースにどのように適応するか: 小学1年生の朝の会における教師と児童の発話の量的・質的分析より, 教育心理学研究, 49(3), 314-325.
- Takehana, K. & Matsui, T. (2016, July). Association Rules on Relationships Between Learner's Physiological Information and Mental States During Learning Process, In Proceedings of HCI International (HCI2016), LNCS 9735, 209-219.
- Tensorflow, <https://www.tensorflow.org> (2016.3.7 参照)