心的状態の時間遅れと持続モデルを考慮した生体情報からの学習者の心 的状態推定の試み

Study on Estimation of Learner's Mental States from Physiological Indexes Considering Time Dilation and Persistent Model of Mental States

松居 辰則*1	宇野 達朗*2	田和辻 可昌*1
Tatsunori Matsui	Tatsuro Uno	Yoshjimasa Tawatsuji

*1 早稲田大学人間科学学術院 Faculty of Human Sciences, Waseda University #1 *² 早稲田大学人間科学部 School of Human Sciences, Waseda University #2

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. Particularly, in this study, we tried on estimation of learner's mental states from physiological indexes considering time dilation and persistent model of mental states. As a result, the effectiveness of considering persistent model and time delay of physiological indexes were suggested.

1. はじめに

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理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[Tatkehana 16]. 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[Fujiyoshi 15][Horiguchi 10][Kojima 14]. 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深 層ニューラルネットワーク(Deep Neural Network(以下, DNN)) を用いて学習者の心的状態の推定の可能性を検討した. 特に、時限解析による時間遅れ、心的状態の持続モデルを考慮する ことによる推定精度の向上の可能性を検討した.

本発表は,第80回人工知能学会先進的学習科学と工学研 究会において発表したものを編集したものである.

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

連絡先:松居辰則, 早稲田大学人間科学学術院, 〒359-1192 所沢市三ケ島 2-579-15, E-mail: matsui-t@waseda.jp は NIRS(日立 WOT-100),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.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. 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.

3.3 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 DL 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.心的状態を表すカテゴリは, Achievement Emotions Questionnaire(AEQ)[Perkun 11]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.教師の発話を表すカテゴリは、先行研究で使用 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. なお, NIRS データには大域平均基準化[野澤 09][平山 12]を施した.

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 カリ)であった.なお、①②③は粒度を統一するために、粒度の 荒いデータに関しては、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.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(3.2で詳述).

3.2 シミュレーション

Python3.5, Tensorflow (ver 0.12.1) [Tensorflow]で実装した. 中間層での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.

(1) 中間層のユニット数

中間層の数を 1 から 25 まで変化させ、中間層数 19 で Loss 値が最小, Accuracy 値が最大となったため、中間層のユニット 数は 19 とした(図 2 参照).



図2 中間層のユニット数による Loss 値, Accuracy 値の変化



図3 シミュレーション結果

(2) シミュレーション方法と結果

中間層のユニット数 19 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 3 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.

4. DNNを用いた心的状態推定の試み

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 化は行わっていない.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-て*から時 刻tまでの①②③④のデータを、また、出力層では時刻 t-τ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 値が、 $\tau = 9$ 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.





4.2 シミュレーション方法と結果

Python3.5, Tensorflow (ver 0.12.1)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. 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 値を算出(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 0.800, 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 0.800 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.







図5 5000 回学習時の Loss 値と Accuracy 値の変化

5. 次元解析による時間遅れと次元の決定

これまでは、時間遅れ τ (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)については、経験的に決定してい た ($\tau=3\sim10$). そこで、次元解析の手法を用いて、時 間遅れ τ とその自由度(次元)を求めた. 具体的には、各 生体情報を時系列データとして捉え、その自己相関性が最 も高くなる場合の時間遅れ τ と、システムとしてもつ次元 (システムの自由度:そのシステムをモデル化する際に必 要となる最低限の局所的な独立変数の数度)を求めた. そ の結果を表 1 に示す. この結果から、①NIRS は 2 次元、 ②呼吸は 1 次元、③皮膚コンダクタンスは 4 次元、④教師 の発話を DL の入力データとした(合計 1530 データ). 出力データは⑤内省報告であった. なお、時間遅れを考慮

した結果、推定される感情状態(内省報告)は3種類とな った(これまでは5種類).ネットワークの構造は4節と 同様であり、中間層1のユニット数は19、中間層2のユ ニット数は 20 であった. 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. 対象データ (1530 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 6(上段 Loss 値,下段 Accuracy 値) に示す. 今回も4節の場合と同様に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 値を算出(小数点第4位で四捨五入)した結果, 平均 0.894, 標準偏差.0087 となった. このことから, 次 元解析による時間遅れとシステムの次元を考慮した結果, 4 節の場合と比較してより高い精度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.

表1 次元解析の結果(時間遅れてと次元m)

	NIRS	呼吸	皮膚コンダクタンス
τ	400	155	173
(Step)			
次元 <i>m</i>	1.86	1.18	3.81



図 6 5000 回学習時の Loss 値と Accuracy 値の変化(次 元解析による時間遅れとシステムの次元を考慮)

6. 心的状態の持続モデルの導入

一般に人間の心的状態は時間経過に従って変化するとさ れている.特に,ある刺激に対して生起した感情(心的状態)は,時間の経過に対して指数関数的に減衰するとの知 見が報告されている[Steephan 13][金子 16].

6.1 心的状態の持続モデル(感情強度関数)

そこで、本研究においては、④教師による発話と⑤学習 者の心的状態に対して、以下のような減衰関数を適用する ことにした.

$$f(t) = f(T_0) \times \exp\left(-\frac{t - T_o}{T_1 - T_0}\right) \qquad (T_0 \le t \le T_1)$$

ただし、 T_0 , T_1 はそれぞれ、教師のある発言、および学習者のある心的状態の開始時刻と終了時刻を表す. f(t)は emotional intensity(感情強度)を表現する関数である と考えられる. なお、本研究においてはデータの性質上、 $f(T_0)=1$ としている.

6.23層 NN, DNN の感情強度関数の適用

5 節までに用いたデータに、感情強度関数を適用して、 3 層 NN, DNN による、生体情報と教師の発話からの学

習者の心的状態の推定を試みた.5節までと同様に、入力 データは①NIRS, ②呼吸, ③皮膚コンダクタンス, ④教 師の発話であった.また、出力データも同様に⑤学習者の 心的状態であった. ネットワークの構造は4節と同様で あり、中間層 1 のユニット数は 19、中間層 2 のユニット 数は 20 であった.中間層の活性化関数は二層とも tanh 関 数を、出力層では出力ベクトルの max 値以外を 0 として 最小二乗法を用いた. 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(1530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.時間遅れを考慮したモ デル(3 層 NN,中間ユニット数 12)の Loss 値の変化と 時間遅れを考慮したモデル(DNN,中間層層数:3(ユニ ット数(20,3,12))のLoss 値の変化を表 2,表 3 に示す.

表 2 5000 回学習後の Loss 値 (3 層 NN)

Alpha	最適化手法	Gradient Descent	Adam
0.05		0.0526	0.0526
0.01		0.0526	0.0526
0.005		0.0526	0.0526
0.001		0.0526	0.0526
0.0005		0.0526	0.0526
0.0001		0.0555	0.0549

表3 5000 回学習後の Loss 値 (DNN)

Alpha 最適化手法	Gradient Descent	Adam
0.05	0.0526	0.0526
0.01	0.0526	0.0526
0.005	0.0526	0.0526
0.001	0.0526	0.0526
0.0005	0.0526	0.0526
0.0001	0.0557	0.0526



図 7 Loss 値の変化(学習係数:0.0001/オプティマイ ザ: Gradient Descend)

いずれの場合も, Accuracy 値は 0.950 程度の(Loss 値 は 0.050 程度)高い値に収束している.この結果は,心低 状態の持続モデルを考慮しない場合の値(0.890 程度)と 比較しても高い値になっている.

さらに、図7に3層NNとDNNとの場合のLoss 値の 変化の様子を示す(学習係数: 0.0001/オプティマイザ: Gradient Descend). この結果,DNNの方が3層NNに 比較して,Loss 値の収束が顕著に早いことがわかる.こ の結果は、アプティマイザー(例えばAdam等)によらず 同様の結果となった.また、学習係数の値の増加に伴って、 この傾向はより顕著になるとの結果も得た.

7. 考察

DNN 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約 0.800 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.今回,生体情報における時間 遅れの程度を次元解析の手法で求め,DNN への入力デー タに反映させた.その結果,より高い Accuracy 値での心的 状態の推定が可能となった.さらに,心的状態の持続に関 するモデルを反映させた結果,さらに高い Accuracy 値で心 的状態の推定が可能になった.つまり,計測データをより 実現象を近似するデータに加工することの有効性を知見と して得ることができた.

8. 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、DNN 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. 特に、次元解析による生体情報の時間遅れと自由度(シス テムの次元)、心的状態に持続性に関するモデル(感情強 度関数)を導入することにより、より高い Accuracy 値で 心的状態の推定が可能となるとの知見を得た.

ー方で、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、過学習が起こっていることが示唆 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 DNN 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のネッ 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、DNN 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.著者らは先行研究に おいて、学習に関わる多面的情報、つまり、教師の発話、 学習者の心的状態、学習者の生体情報との関係性の形式化 を相関ルール抽出の手法を用いて試みている.今後は、こ れらの知見を融合させることで DNN の中間層の解釈の可 能性が高まるものと考えている.

参考文献

- [Fujiyoshi 15] 藤好宏樹, 吉村和代, K. Kunze, 黄瀬浩一: 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, Vol.115, No.24, pp.49-54, PRMU2015-10 (2015)
- [平山 12] 平山健太,綿貫啓一, 楓和憲: NIRSを用いた随意運動および他動運動の脳賦活分析,日本機械学会論文集C編, Vol.78, No.795, pp.3803-3811 (2012)
- [堀口 10] 堀口祐樹,小島一晃,松居辰則: 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, SIG-ALST-A903, pp.1-6 (2010)
- [小島 14] 小島一晃, 村松慶一, 松居辰則: 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, Vol.31, No.2, pp.197-202 (2014)
- [Pekrun 11] Pekrun, R., Goetz, Frenzel, A. C., Barchfeld, P. and Perry, R. P.: Measuring Emotions in Students' Learning and Performance: The Achievement Emotions Questionnaire (AEQ), Contemporary Educational Psychology, Vol.36, No.1, pp.36-48 (2011)
- [Takehana 16] Kazuma TAKEHANA, Tatsunori MATSUI: Association Rules on Relationships Between Learner's Physiological Information and Mental States During Learning Process, In Proceedings of HCI International (HCII2016), LNCS Vol. 9735, pp.209-219 (2016)
- [Tesorflow] https://www.tensorflow.org (2016.03.07 参照)
- [Steephan 13] Steephan, J.: A computational model of affective adaptation and emotion dynamics, IEEE transactions on affective computing, 4(2), pp.197-210 (2013).
- [金子 16] 金子迪大, 堀毛一也: ポジティブ情動喚起刺激の不 確実性と重要性が情動持続に与える影響, 感情心理学研 究, 24(1), pp.33-41 (2016).