

筆記具の動きから学習者の学習状況を 推定する手法の提案と開発

安達脩太¹ 加藤直樹²

概要：本稿では、筆記具の動きから学習者の学習状況を推定する手法を提案し、それを組み込んだ学習状況可視化ツールの開発を行った。さらに学習状況可視化ツールを用いて、実際の学習を模したタスクに対して提案する手法が有用であるかについて評価実験を行った。評価実験の結果、提案するニューラルネットワークと隠れマルコフモデルを用いた手法は学習者の学習状況の推定に有用であることが示された。また、隠れマルコフモデルはある学習状況から異なる学習状況に移行する局面における推定性能の向上に有用であることが示された。一方である学習状況から異なる学習状況に移行する局面における推定性能について課題があることが示された。

キーワード：学習状況可視化、学習者センシング、筆記動作、教育支援システム

Developing a Method for Estimating Learners' Learning States Based on Writing Implement Motion Data

SHUTA ADACHI^{†1} NAOKI KATO^{†2}

1. はじめに

1.1 研究背景

授業中の児童生徒の学習状況を把握することは重要である。「児童生徒の学習評価の在り方について（平成31年中央教育審議会）」[1]では、学習評価は子供たちの学習状況を把握し、学習改善や指導改善につなげることを目的とすると記されており、教員が授業の過程で児童生徒の学習状況を捉えることの意義が示されている。また、文部科学省「豊かな学びに繋がる学習評価の在り方（令和7年）」[2]では、学習者が自らの学習過程を省察し、学びの改善に生かすことの意義が示されている。授業中の学習状況の記録は児童生徒が自らの学習方略を見直す契機となる。

GIGA スクール構想の実現により、教育現場のICT環境の整備は飛躍的に進んだ。文部科学省「令和6年度学校における教育の情報化の実態等に関する調査結果」[3]によると、令和7年3月時点で児童生徒あたりの学習者用コンピュータの台数が1.1台であり、全ての児童生徒が学習者用端末を利用できる環境の整備が完了している。文部科学省「新時代の学びを支える先端技術活用推進方策（最終まとめ）」[4]では、このように整備された環境を基盤とし、先端技術・教育ビッグデータを活用することの意義の1つとして、センシング技術の活用による児童生徒の個々の状態の客観的かつ継続的な把握により、個別に最適で効果的な学びや支援が可能になることを挙げている。

文部科学省の実証事業（令和6年3月）[5]によると、デジタル教材と紙媒体の教材が併用される授業が多いことが示されている。このような授業形態において、デジタル教科書実践事例集[6]や文部科学省のStuDX Styleの事例の1つである「取り組み状況をすばやく把握して個別サポート」[7]に示されているように、児童生徒が端末を使用しているときは学習画面の共有やデジタル教材の入力操作の取得により、授業中の児童生徒の学習状況の把握を支援することが可能である。しかし、デジタル教材を閲覧するだけで入力操作がない状態や、紙の教材を使用しているときにはこの方法は使えない。児童生徒の表情を捉えるカメラ映像を用いるなど[8]、児童生徒の生体情報を検出することで、学習状況を推定する技術も開発されている[9][10]が、運用コストが高いなど教育現場への導入は難しい面もある。

この課題に対して、デジタル教材の使用有無に関係なく、児童生徒が共通して利用する筆記具の動作から学習状況を推定することができれば、この課題を解消し、あらゆる授業においても児童生徒の学習状況の推定が可能になる。

1.2 本研究の目的

本研究では、筆記具の動きから学習者の学習状況を推定する手法を提案する。そして、それを実装したツールを用いた評価を通して、提案した手法が学習者の学習状況を推定する手法として有用であるかを検証する。

本研究では、筆記具を使用した学習において、学習者が筆記具を持ち、一画ずつ書き進めている動きを継続してい

¹ 東京学芸大学教育学部教育支援課程
Faculty of Education, Tokyo Gakugei University
² 東京学芸大学
Tokyo Gakugei University

る状況である「書いている」、筆記具を持っているが書いていない状態で、学習者が筆記具を水平方向ではない方向に継続して静止させている状況、または書いているとは異なる不規則な動きが継続させている状況である「手が止まっている」、学習者が筆記具を手を持たず机に置いている状態で、筆記具を水平方向に継続して静止させている状況である「筆記具を置いている」の3つを取り上げる。これらの学習状況の推定は、学習者の学習状況を客観的に把握することや「集中して学習している」などのより具体的な学習者の学習状況を推定することの一助になると考える。

なお、落書きと文字筆記の区別のような学習者が書いた内容についての推定や、集中しているなどのより具体的な学習状況の推定までは踏み込まず、筆記具の動きのみから推定できる学習者の学習状況の検出方法の確立に集中する。また、先行研究では、「手を動かしているが書いていない」状態を1つの学習状況として推定している。しかし、「手を動かしているが書いていない」状態や「筆記具を振っている」状態は、ペンの動きとして区別することが難しい一方、ペン先の状態を取得することで容易に分類することが可能である。本研究ではペンの動きだけからの推定の可能性を追求し、上記状態の分類は行わない。

本研究では、先行研究で課題として挙げられた、ある学習状況から異なる学習状況へ移行する局面における推定精度を向上させることにより、実際の学習を模したタスクに対する推定性能を向上させることを試みる。

2. 先行研究

2.1 しゅくだいやる気ペン

学習者の筆記データを基に児童生徒の学習状況を推定するツールとしてコクヨが提供するしゅくだいやる気ペンがある[11]。このツールでは、筆記具に取り付けたセンサで取得した加速度データを基に学習時間と集中度を推定し、アプリと連携して児童生徒の学習状況を管理している。通常の筆記では発生しない大きな振動では学習していると検知しない仕組みは搭載されているが、ストローク間の移動の際に加速度の特徴値が大きく異なる場合や、筆記や姿勢に特徴がある場合は適切な推定が行われないことがある。

2.2 学習者センシングのための筆記行為の検知手法と評価

江木らは、筆記具の頭部に取り付けたセンサで取得した加速度データから学習用データと評価用データを用意し、求めた振動の平均、分散、エネルギーを特徴量とし、J48 決定木アルゴリズムによるデータの分類、学習と評価を HascTool で行った。学習者の筆記具の動きから学習者の学習を「置いている」、「手が止まっている」、「手が動いている」、「書いている」、の4つの学習状況に分類した。学習用データと評価用データには、被験者12名に4つの学習状況に対してそれぞれ、筆記具を置いた状態を継続させる、

筆記具を持っているが手を動かさない、筆記具を持って動かしているが書いていない、ひらがなの五十音表が印刷されたA4の1枚の記入用紙にひらがなを書く、という30秒間のタスクを5回ずつ行わせ際に、10ミリ秒間隔でサンプリングしたデータを使用した。評価用データに対する総合識別率は90.62%であった。この結果から、4つの学習状況ごとに別々でタスクを行いサンプリングした評価用データに対してはJ48決定木アルゴリズムを用いた分類モデルは高精度であった。しかし、実際の学習におけるある学習状況から異なる学習状況に移行する局面では取得されるデータの特徴量が大きく変動する。これらの状況に対して提案された分類モデルの有用性は十分に検証されておらず、その結果として分類精度が低下する可能性が示唆された。

2.3 本研究の位置付け

前節の研究より、筆記具の動きをセンシングすることは学習者の学習状況の推定において有用であると言えるが、2.2節の研究で指摘されたように、学習状況ごとに別々でタスクを行いサンプリングした評価用データを用いた評価から提案された手法では、ある学習状況から異なる学習に移行する局面における推定性能が低くなることが示唆される。そこで、本研究では、1.2節で定義した3つの学習状況を高い精度で推定する自動推定手法を開発し、実際の学習を模したタスクを用いた実験に対する推定性能の評価、ある学習状況から異なるに移行する局面における推定性能の評価を通して、自動推定手法の有用性を検証する。

3. 推定に用いるデータの収集

本章では筆記具の動きから学習者の学習状況を推定する手法を探究する上で、推定に有用と考えられるデータ(以下、筆記データと呼ぶ)の選定、及び推定する3つの学習状況に対応する筆記データの収集について述べる。

3.1 筆記データの選定

江木ら[13]は加速度センサを用いて学習者の筆記行為を検出するために取得する加速度の取得位置として、筆記具の筆先部、クリップ部、頭部について、各位置で計測した加速度の値を比較し、筆記具の支点から最も遠く、移動量が大きい筆記具の頭部が取得位置として適していることを示した。そのため、本研究でも筆記具の頭部の加速度、角速度、角度を筆記データとして選定する。

加速度は筆記具の移動速度の変化を表すデータで、筆記動作における方向転換で生じる急激な移動速度の変化や線を書く際の瞬間的な加速、筆記具を動かしていない場合の移動速度の僅かな変化などを捉える。角度は筆記具の空間的な傾きを表すデータで、3つの学習状況の空間的特徴を捉える。角速度は筆記具の回転運動の速さを表すデータで、筆記動作における手首や指のひねり動作による筆記具の傾きの変化や筆記の際の筆記具上部の回転の速さを捉える。

3.2 筆記データの定義

3.2.1 加速度

時刻 t における XYZ 軸それぞれの加速度 $(\alpha_{x_t}, \alpha_{y_t}, \alpha_{z_t})$ はセンサ内で 0.1 秒前の速度 $(v_{x_{t-0.1}}, v_{y_{t-0.1}}, v_{z_{t-0.1}})$ とその 0.1 秒後の時刻 t の速度 $(v_{x_t}, v_{y_t}, v_{z_t})$ の差をその 2 点の取得時間差 Δt (0.1 秒) で割ることで求める。

$$\alpha_{x_t} = \frac{v_{x_t} - v_{x_{t-0.1}}}{\Delta t}, \quad \alpha_{y_t} = \frac{v_{y_t} - v_{y_{t-0.1}}}{\Delta t}, \quad \alpha_{z_t} = \frac{v_{z_t} - v_{z_{t-0.1}}}{\Delta t}$$

3.2.2 角度

時刻 t における XYZ 軸それぞれの角度 $(d_{x_t}, d_{y_t}, d_{z_t})$ はセンサ内で筆記具の上部の先端を中心として求める (図 1)。

3.2.3 角速度

時刻 t における XYZ 軸それぞれの角速度 $(\omega_{x_t}, \omega_{y_t}, \omega_{z_t})$ はセンサ内で 0.1 秒前の角度 $(d_{x_{t-0.1}}, d_{y_{t-0.1}}, d_{z_{t-0.1}})$ とその 0.1 秒後の時刻 t の角度 $(d_{x_t}, d_{y_t}, d_{z_t})$ の差をその 2 点の取得時間差 Δt (0.1 秒) で割ることで求める。

$$\omega_{x_t} = \frac{d_{x_t} - d_{x_{t-0.1}}}{\Delta t}, \quad \omega_{y_t} = \frac{d_{y_t} - d_{y_{t-0.1}}}{\Delta t}, \quad \omega_{z_t} = \frac{d_{z_t} - d_{z_{t-0.1}}}{\Delta t}$$

3.3 筆記データの環境

3.3.1 ハードウェア環境

筆記具の頭部にセンサを固定して筆記データを取得する。筆記具の軸と並行な方向を X 軸、筆記具の軸に対して直角方向を Y 軸、筆記具と鉛直方向を Z 軸として固定したセンサ付き筆記具を図 1 に示す。センサは Witmotion 社の WT901BLECL [14] を使用し、気温、磁場、直線的な直交軸 XYZ 軸それぞれの加速度、角速度、角度を 0.1 秒ごとに、加速度は $\pm 16G$ 、角速度は $\pm 200^\circ/s$ 、XZ 軸の角度は $\pm 180^\circ$ 、Y 軸の角度は $\pm 90^\circ$ の範囲で測定する。重量は 18g で、学習者の筆記動作に与える重量の影響は小さいと考えられる。筆記データの通信には BLE5.0 を使用し、壁など障害物のない環境において 50m 以内でデータ通信が可能である。

3.3.2 ソフトウェア環境

筆記データの収集は開発したスクリプトを用いて行う。Bleak ライブラリを使用し、センサから送信される筆記データを bytearray 形式で取得し、保存する。非同期で BLE (Bluetooth Low Energy) による通信を確立し、筆記データの取得、保存をループ処理で 0.1 秒ごとに実行する構成にし、センサから筆記データを受信しつつ、同時にデータを保存する。

データ収集の手間を極力減らすために、センサの電源を入れると接続が開始され、筆記データの取得を開始し、センサの電源を切ると接続が解除され、筆記データを格納したファイルを指定したディレクトリに自動で保存する。

3.4 データ収集

筆記具の動きから学習者の学習状況を推定する手法を設計するため、各学習状況における筆記データの収集を行った。

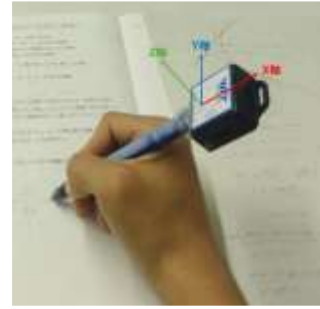


図 1 センサ付き筆記具

3.4.1 タスク

各学習状況に対応するタスクは次の通りとした。

- (ア) 筆記具で書く
- (イ) 筆記具を持っているが書かない
- (ウ) 筆記具を置く

タスク (ア) では、図 1 の向きで筆記具を持ち、コクヨのキャンパスノート A 罫 [15] の 7mm 四方のマスの任意の平仮名、片仮名、アルファベット、漢字、記号を書く動作を 30 秒間継続するタスクを行う。タスク (イ) では、図 1 の向きで筆記具を持ち、筆記していない状態を 30 秒間継続するタスクを行う。タスク (ウ) では、机上の任意の場所に筆記具を置き、静止させている状態を 30 秒間継続させるタスクを行う。

3.4.2 データ収集手順

被験者は筆者が所属する大学の学生 10 名である。

被験者に対してセンサ付き筆記具の使い方とタスクの説明を行った後、タスク (ア)、タスク (イ)、タスク (ウ) の順でそれぞれ 4 回ずつ間隔をあけて行ってもらった。

4. 自動推定手法の設計と構築

本章では、学習者の学習状況を推定するための自動推定手法について、自動推定手法の設計と前章で収集した筆記データを用いた構築について述べる。

4.1 学習用データと評価用データ

自動推定手法の設計を行うため、推定に用いる特徴量を抽出するための学習用データと設計した推定手法の推定性能を評価するための評価用データを用意する。3.4.2 節で収集した被験者 10 人分の筆記データの XYZ 軸それぞれの加速度、角速度、角度の最小値から最大値を範囲として -1 から 1 に正規化を行い、「筆記具を置いている」を 0、「手が止まっている」を 1、「書いている」を 2 とした正解ラベルを加え、作成した。前 36000 個のデータを学習用として 30000、評価用として 6000 を用いた。

4.2 自動推定手法の設計

筆記具の動きから 3 つの学習状況を推定するには収集した筆記データから特徴量を抽出し、推定を行うモデルが必要である。先行研究では閾値を使用した推定が行われている。しかし、閾値を使用する手法は、筆記データから抽出

する特徴量が大きく変動する、ある学習状況から異なる学習状況に移行する局面の推定性能の低下を招く課題がある。そこで、筆記データのような非線形データの分類や学習、特徴量の抽出が可能なニューラルネットワークを用いて3つの学習状況を推定することを試みる。ニューラルネットワークは筆記データに対して高曲率の分類境界を学習し、各学習状況に対して推定確率を取得する。

筆記データは0.1秒ごとに取得する時系列データであり、RNN (Recurrent Neural Network) などの再帰構造により時系列要素を考慮した推定を行うニューラルネットワークを用いた推定が、先に記したある学習状況から異なる学習状況に移行する局面での推定に有用と考えられる。

探索的にネットワーク構成を検討し、スライド幅、各レイヤーから出力されるニューロン数、重みなどのパラメータを変更し、再帰構造を含むニューラルネットワークを構築し、評価用データを用いて3つの学習状況に対する再現率、適合率、F1値を評価した結果、「書いている」、「手が止まっている」の学習状況の再現率、適合率、F1値が0.5未満であり、3つの学習状況の推定において十分な推定性能を確保できなかった。学習者によって各学習状況における筆記具の動きに要する時間の個人差が大きく、推定に影響させる時系列要素の範囲を適切に設定できないためだと考えられる。そこで、再帰構造を用いず、入力層、全結合層、出力層で構成される多層パーセプトロンを用いたニューラルネットワーク（以下、ニューラルネットワークと記す）を使用する。

特徴量の抽出は0.1秒ごとに行われ、短時間で変化する動きの特徴を正確に捉えることが可能である一方で、「書いている」から「手が止まっている」、そして「筆記具を置いている」に状態が移る際、学習者の筆記具の持ち方によっては「手が止まっている」が0.1秒より短くなり、0.1秒ごとに出力する推定結果が、「書いている」から「手が止まっている」状態を経由しない「筆記具を置いている」への変化になるなど、学習状況の変化として不自然な推定となることがある。そこで、ある学習状況から異なる学習状況に移行する局面において、学習状況の継続性や変化の妥当性を考慮した推定を可能にするため、隠れマルコフモデルのforward アルゴリズムを、ニューラルネットワークが出力する各学習状況の推定確率に適用する（以下、自動推定モデルと呼ぶ）。

4.3 自動推定モデルの設計

自動推定モデルに使用するニューラルネットワークの選定と構築の際に、ネットワーク構成やパラメータの値の検討や評価用データを用いた推定性能の比較を行う必要がある。そのため、推定モデルの構築、学習、評価には、複数の推定モデルとの比較が容易であり視覚的な操作で構造

やパラメータの値を操作できる Sony の Neural Network Console[16]を使用し、自動推定モデルを設計する。

4.3.1 ニューラルネットワークの選定

学習用データの XYZ 軸の加速度、角速度、角度の九つのパラメータは相互に関係しているため、ベクトルデータとしてニューラルネットワークに入力することで、各パラメータ間の相関関係や複合的な特徴を捉えることができる。

学習用データを学習させたニューラルネットワークのネットワーク構成やスライド幅、各レイヤーから出力されるニューロン数、重みなどのパラメータを探索的に変更し、それぞれの適合率、再現率、F1 値、正解率を求め、自動推定モデルに使用するニューラルネットワークを F1 値が一番高いニューラルネットワーク（図 2）とする。ニューラルネットワークは入力層、全結合層、出力層から構成される。入力された9次元ベクトルデータの特徴量を全結合層で100次元にし、活性化関数に ReLu 関数を適用することで非線形性を導入する。過学習を抑制するために Dropout を適用した後、全結合層で特徴量を3次元に変換する。「筆記具を置いている」を0、「手が止まっている」を1、「書いている」を2とし、Softmax 関数を用いて各学習状況に対する確率を0から1の範囲で算出する。そして、確率が一番高い学習状況をニューラルネットワークの出力結果とする。なお、各全結合層の重みの初期値にはネットワーク中のニューロン数によって決まるある範囲の一様分布に従って初期パラメータを生成する手法で生成した値である Xavier initialization を使用する。

選定したニューラルネットワークの正解率、適合率、再現率、F1 値を表 1 に示す。適合率、再現率、F1 値は、各学習状況をデータ数によらずに同等に扱って評価を行うため、マクロ平均を用いて算出する。評価用データを用いて評価を行った前 2.2 節の先行研究では、正解率が 0.906 である。本研究の評価用データに対する正解率は 0.968 であり（表 1）、同等以上の結果を得られた。そのため、このニューラルネットワークを自動推定モデルに使用する。

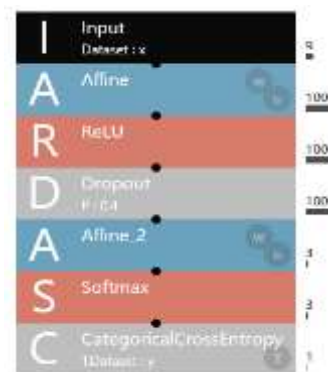


図 2 ニューラルネットワークの構成

表 1 ニューラルネットワークの評価

| 評価指標 | 評価結果 |
|------|-------|
| 正解率 | 0.968 |
| 適合率 | 0.969 |
| 再現率 | 0.968 |
| F1 値 | 0.967 |

4.3.2 自動推定モデルの詳細設計

前節で選定したニューラルネットワークの Softmax 関数による各学習状況に対する 0 から 1 の推定確率を隠れマルコフモデルの forward アルゴリズムに入力して処理を行い、自動推定モデルの推定結果を求める。

初期分布 a_0 を次の定数とする。

$$a_0 = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right]$$

遷移行列 A を次の定数とする。

$$A = \begin{bmatrix} 0.5 & 0.5 & 0.0 \\ 0.3 & 0.3 & 0.3 \\ 0.0 & 0.5 & 0.5 \end{bmatrix}$$

時刻 t で多層パーセプトロンから出力された分類確率を e_t として次の式で表す。

$$e_t = [e_t(0), e_t(1), e_t(2)]$$

forward アルゴリズムで 0.1 秒ごとに更新される状態分布 a_t を次の計算式で表す。

$$a_t = \frac{(a_{t-0.1}A) \odot e_t}{\sum_{k=0}^2 [(a_{t-0.1}A)[k] \cdot e_t(k)]}$$

a_t の各要素 0, 1, 2 のうち、一番確率が大きい要素をその時刻における学習状況の推定結果として出力する。

4.4 自動推定モデルの構築

前 4.3 節で設計したニューラルネットワークと隠れマルコフモデルの forward アルゴリズムを用いた自動推定モデルを構築する。Neural Network Console で作成したニューラルネットワークについて、Neural Network Console のエクスポート機能を用い、ニューラルネットワークの構造が定義された Python コードと、学習設定、データ設定、学習済み重みの情報を格納した構造化されたデータをそれぞれ Python ファイル、NNP 形式のファイルとして取得する。ニューラルネットワークの推定結果を前 4.3.2 節で設計した隠れマルコフモデルの forward アルゴリズムの計算手法で処理するプログラムを作成し自動推定モデルを構築する。

5. 学習状況可視化ツールの設計と開発

本章では、自動推定モデルの有用性を評価するために自動推定モデルによる学習状況可視化機能を実装した学習状況可視化ツールの設計と開発について述べる。



図 3 リアルタイムでの学習状況可視化機能

5.1 開発環境

本ツールは、Windows11 が搭載された PC を用い、Visual Studio Code を用いて開発を行った。動作環境は仮想環境の venv で、動作確認は Google Chrome を使用して行った。開発言語は Python, HTML, CSS, JavaScript, フレームワークは Django を使用した。データはローカルサーバー上に構築したデータベースに保存した。

5.2 利用した API とライブラリ

筆記具に取り付けたセンサとの接続は Bluetooth Low Energy を Web Bluetooth API を用いて実現した。センサと GATT 通信をする際は Bleak ライブラリを用いて mac アドレス, UUID, Characteristic UUID を取得して行う。また、構築した自動推定モデルを学習状況推定ツールに組み込むために、ツール上のニューラルネットワークの推論実行エンジンには Neural Network Libraries を使用し、自動推定モデルとの連携には NNabla Python API を使用した。

5.3 機能の設計と実装

5.3.1 リアルタイムでの学習状況可視化機能

学習中の学習者の学習状況を把握するため、リアルタイムで筆記データを取得し、学習状況を推定する必要がある。そのため、筆記具に取り付けたセンサを Bluetooth で端末と接続し、筆記データを取得する。取得した筆記データを -1 から 1 に正規化された 9 次元ベクトルデータに変換し、自動推定モデルに入力として与える。組み込まれた自動推定モデルを実行し、推定結果を座席表に表示する (図 3)。

5.3.2 ログ収集機能

自動推定モデルを組み込んだツールにおいて学習者の学習ログを用いて、自動推定モデルの有用性を評価するため、組み込んだ自動推定モデルが推定した学習状況を記録したログを収集する必要がある。リアルタイムでの学習状況可視化機能の使用時のセンサ付き筆記具と端末の接続が完了した時点からセンサ付き筆記具と端末の接続が切れる時点までをログの収集範囲とする。記録する項目は推定結果を出力した時刻、自動推定モデルの推定結果、ニューラルネットワークのみの推定結果である。収集したログはローカル上の指定したディレクトリに保存される。

6. 評価実験

6.1 目的

開発したツールを用いて、実際の学習を模したタスクにおける自動推定モデルの推定性能の評価を通して、提案する自動推定モデルの有用性を検証する。

6.2 実験手順

前 3.4.2 項で筆記データを収集した被験者とは異なる大学生 10 名を被験者とし、センサ付き筆記具の使用方法について説明した後に、令和 6 年度全国学力状況調査小学校第 6 学年国語[17]の大問 2 の文章題を提示し、センサ付き筆記具を使用し 4 分間の制限時間を設けて解答させるタスクを遂行させた。

タスクの遂行中には学習状況可視化ツールのログ収集機能を使用し、学習状況可視化ツールに組み込んだ自動推定モデルの推定結果とニューラルネットワークのみの推定モデルの推定結果を取得した。推定結果は 0.1 秒ごとに出力されるため、0.1 秒間隔でログを取得しようと試みたが、センサデバイスおよびプログラムの処理時間による遅延が発生し、記録される時刻の間隔にずれが生じた。そこで、サンプリング間隔を 0.2 秒に設定し、ログを取得した。また、実際の学習状況を記録するため、解答中の被験者の様子とその時刻を表す時計を録画した。

6.3 評価方法

タスク中の 4 分間の録画映像を 0.2 秒ごとに分割し、「筆記具を置いている」を 0、「手が止まっている」を 1、「書いている」を 2 として手でラベル付けを行い、その時刻における正しい学習状況とする。各時刻における自動推定モデルによる推定結果と正しい学習状況を比較する。

自動推定モデルは「筆記具を置いている」を 0、「手が止まっている」を 1、「書いている」を 2 と分類する三値分類であるため、ある学習状況を評価対象とする場合、その学習状況を陽性 (Positive)、それ以外の二つの学習状況を陰性 (Negative) として混合行列を定義する。たとえば、評価対象が「筆記具を置いている」の場合は、実際の学習状況が 0 であり、自動推定モデルの推定結果が 0 であった場合を True Positive (TP_0)、自動推定モデルの推定結果が 1 または 2 であった場合を False Negative (FN_0) とし、実際の学習状況が 1 または 2 であるが、自動推定モデルの推定結果は 0 であった場合を False Positive (FP_0)、自動推定モデルの推定結果は 1 または 2 であった場合を True Negative (TN_0) と定義する。「手が止まっている」および「書いている」を評価対象とする場合も同様に混合行列を定義する。

各学習状況 ($k \in \{0, 1, 2\}$) に対する再現率 (k)、適合率 (k)、F 値 (k) を次の式で求める。

$$\text{再現率}(k) = \frac{TP_k}{TP_k + FN_k}$$

$$\text{適合率}(k) = \frac{TP_k}{TP_k + FP_k}$$

$$F1 \text{ 値}(k) = \frac{2 \times \text{適合率}(k) \times \text{再現率}(k)}{\text{適合率}(k) + \text{再現率}(k)}$$

さらに、自動推定モデル全体の正解率、再現率、適合率、F1 値を算出し、自動推定モデルの有用性を評価する。全データ数 N に対して正しく推定できた割合である正解率は次の式で求める。

$$\text{正解率} = \frac{TP_0 + TP_1 + TP_2}{N}$$

なお、自動推定モデル全体の再現率、適合率、F1 値をマイクロ平均で求めると、データ数の多い学習状況における再現率や適合率、F1 値に大きく依存するため、マクロ平均を用いて次の式で求める。

$$\text{再現率} = \frac{1}{3} \sum_{k=0}^2 \text{再現率}(k)$$

$$\text{適合率} = \frac{1}{3} \sum_{k=0}^2 \text{適合率}(k)$$

$$F1 \text{ 値} = \frac{1}{3} \sum_{k=0}^2 F \text{ 値}(k)$$

また、ある学習状況から異なる学習状況に移行する局面の時刻を正しい学習状況から取得し、その前後 0.2 秒の時刻の正しい学習状況を、自動推定モデルの推定結果、ニューラルネットワークのみの推定モデルの推定結果とそれぞれ比較する。このとき、ある学習状況から異なる学習状況に変化する時刻の前後 0.2 秒の推定結果の両方が正しい学習状況と一致している場合を「正しい推定」、それ以外の場合を「誤った推定」とする。学習状況が変化する回数と正しい推定の発生回数、誤った推定の発生回数を集計する。自動推定モデルとニューラルネットワークのみの推定モデルを比較し、ある学習状況から異なる学習状況に移行する局面における自動推定モデルの有用性について評価する。

6.4 結果

6.4.1 自動推定モデルの推定結果

実際の学習状況と自動推定モデルの推定結果の対応と 3 つの学習状況の再現率、適合率を表 2 に、F1 値を表 3 に、自動推定モデル全体の性能評価を表 4 に示す。

また、実際の学習状況とニューラルネットワークのみの推定モデルの推定結果の対応と 3 つの学習状況の再現率、適合率を表 5 に、F1 値を表 6 に、ニューラルネットワークのみの推定モデル全体の再現率、適合率、F1 値を表 7 に示す。

表 2 自動推定モデルの推定結果と実際の学習状況との対応

| | | 自動推定モデルの推定結果 | | | 再現率 |
|-----------------|---|--------------|-------|-------|-------|
| | | 0 | 1 | 2 | |
| 実際の 学習状 況 | 0 | 771 | 12 | 2 | 0.982 |
| | 1 | 507 | 6252 | 776 | 0.831 |
| | 2 | 107 | 500 | 3083 | 0.836 |
| 適合率 | | 0.557 | 0.924 | 0.801 | |

表 3 学習状況ごとの F1 値 (自動推定モデル)

| 学習状況 | F1 値 |
|------|-------|
| 0 | 0.711 |
| 1 | 0.876 |
| 2 | 0.817 |

表 4 自動推定モデル全体の評価

| 評価指標 | 評価結果 |
|------|-------|
| 正解率 | 0.842 |
| 適合率 | 0.761 |
| 再現率 | 0.883 |
| F1 値 | 0.802 |

表 5 ニューラルネットワークのみの推定モデルの推定結果と実際の学習状況との対応

| | | 自動推定モデルの推定結果 | | | 再現率 |
|-------------|---|--------------|-------|-------|-------|
| | | 0 | 1 | 2 | |
| 実際の 学習状況 | 0 | 771 | 4 | 8 | 0.985 |
| | 1 | 570 | 6154 | 881 | 0.810 |
| | 2 | 128 | 485 | 2999 | 0.830 |
| 適合率 | | 0.525 | 0.926 | 0.771 | |

表 6 学習状況ごとの F1 値
(ニューラルネットワークのみの推定モデル)

| 学習状況 | F1 値 |
|------|-------|
| 0 | 0.685 |
| 1 | 0.864 |
| 2 | 0.800 |

表 7 ニューラルネットワークのみの推定モデル全体の評価

| 評価指標 | 評価結果 |
|------|-------|
| 正解率 | 0.827 |
| 適合率 | 0.741 |
| 再現率 | 0.875 |
| F1 値 | 0.783 |

表 8 ある学習状況から異なる学習状況に移行する局面における正解率

| | 正 | 誤 | 正解率 |
|-------------|----|----|-------|
| 自動推定モデル | 56 | 70 | 0.540 |
| ニューラルネットワーク | 27 | 99 | 0.262 |

6.4.2 ある学習状況から異なる学習状況に移行する局面の推定結果

ある学習状況から異なる学習状況に移行する局面での、自動推定モデルとニューラルネットワークのみの推定モデルそれぞれの正しい推定の発生回数、誤った推定の発生回数、正解率を表 8 に示す。

ある学習状況から異なる学習状況に移行する局面において、ニューラルネットワークのみの推定モデルで正しく推定した時刻においては、自動推定モデルでも必ず正しく推定された。

6.5 考察

6.5.1 自動推定モデル全体の有用性の考察

自動推定モデルは実際の学習状況を模したタスクに対して、正解率が 0.842、F1 値が 0.802 であることから、特定の学習状況に偏ることなく総合的に良好な精度で推定できることがわかる。さらに、再現率が 0.883 と高いことから、実際の学習状況を正しく推定する性能が高いことがわかる。一方で、適合率が 0.761 であり、再現率に比べてやや低く、自動推定モデルの推定結果の中には一定数の誤った推定が含まれていることがわかる。

自動推定モデルの正解率はニューラルネットワークのみの推定モデルの正解率より 0.015 高いことから、隠れマルコフモデルは推定性能の向上に寄与することがわかる。

また、自動推定モデルの F1 値が、ニューラルネットワークのみの推定モデルより 0.019 高く、適合率は 0.020 高いことから、隠れマルコフモデルの forward アルゴリズムは特定の学習状況に偏ることなく総合的に推定精度を向上させること、実際の学習状況と異なる学習状況に誤って推定されにくくなったことがわかる。

一方、自動推定モデルの再現率がニューラルネットワークのみの推定モデルとほぼ等しいことから、実際の学習状況を正しく推定する性能の向上に隠れマルコフモデルの forward アルゴリズムは寄与しないことがわかる。

これらのことから、自動推定モデルは一定数の誤った推定を含むが、3 つの学習状況を推定する手法として有用であるといえる。また、隠れマルコフモデルの forward アルゴリズムは誤った推定を減らし、3 つの学習状況を推定する性能を向上させるといえる。

6.5.2 「筆記具を置いている」の推定に対する考察

「筆記具を置いている」の再現率は 0.982 と非常に良好

で、実際に筆記具が置かれている状況を正しく「筆記具を置いている」と推定できることがわかる。一方で、「筆記具を置いている」の適合率は 0.557 と良好ではなく、推定結果が「筆記具を置いている」であるが、実際には他の学習状況である可能性が高いことがわかる。特に「手が止まっている」が自動推定モデルにより「筆記具を置いている」に誤って推定された回数が多い。これは 6.2 節の実験で問題文を読む際に、被験者の数名が筆記具を水平にして持ち、文字を書かない状態で静止したため、実際の学習状況は「手が止まっている」であるが、自動推定モデルは「筆記具を置いている」と推定したことによるものと考えられる。

また、自動推定モデルの「筆記具を置いている」の再現率はニューラルネットワークのみの推定モデルの再現率とほぼ等しいことから、隠れマルコフモデルの forward アルゴリズムは実際の筆記具を置いている学習状況に対して正しく推定する性能の向上には寄与しないといえる。

一方で、自動推定モデルの「筆記具を置いている」の適合率はニューラルネットワークのみの推定モデルの適合率より 0.032 高いことから、隠れマルコフモデルの forward アルゴリズムは「筆記具を置いている」と異なる学習状況に誤って推定されないようにすることに寄与するといえる。

6.5.3 「手が止まっている」の推定に対する考察

自動推定モデルの「手が止まっている」に対する推定の再現率は 0.831、適合率は 0.924、F1 値は 0.876 と良好で、「手が止まっている」の推定に対して有用であるといえる。

また、自動推定モデルの「手が止まっている」の再現率はニューラルネットワークのみの推定モデルの再現率より 0.021 高いことから、隠れマルコフモデルの forward アルゴリズムは「手が止まっている」に対して正しく推定する性能の向上に寄与するといえる。一方で、自動推定モデルの「手が止まっている」の適合率はニューラルネットワークのみの推定モデルの適合率とほぼ等しいことから、隠れマルコフモデルの forward アルゴリズムは「手が止まっている」と異なる学習状況に誤って推定されないようにすることには寄与しないといえる。

6.5.4 「書いている」の推定に対する考察

自動推定モデルの「書いている」に対する推定の再現率は 0.836、適合率は 0.801、F1 値は 0.817 と良好で、「書いている」の推定に対して有用であるといえる。特に実験では 10863 回「書いている」と推定されたが、実際は「筆記具を置いている」状況であった回数は 2 回であることから、「書いている」状態を「筆記具を置いている」と誤って推定する可能性は非常に低いことがわかる。

また、自動推定モデルの「書いている」の再現率はニューラルネットワークのみの推定モデルの再現率とほぼ等しいことから、隠れマルコフモデルの forward アルゴリズム

は「書いている」を正しく推定する性能の向上に寄与しないといえる。一方で、自動推定モデルの「書いている」の適合率はニューラルネットワークのみの推定モデルの適合率より 0.030 高いことから、隠れマルコフモデルの forward アルゴリズムにより、「書いている」と異なる学習状況に誤って推定されないようにすることに寄与するといえる。

6.5.5 ある学習状況から異なる学習状況に移行する局面における推定の有用性の考察

ある学習状況から異なる学習状況に移行する局面では、自動推定モデルの正解率はニューラルネットワークのみの推定モデルより 0.278 高い正解率を示した。また、「筆記具を置いている」から「手が止まっている」への変化、「手が止まっている」から「書いている」への変化、「書いている」から「手が止まっている」への変化、「手が止まっている」から「筆記具を置いている」への変化の全ての学習状況の変化で、学習状況の変化を正しく推定した個数がニューラルネットワークのみの推定モデルより自動推定モデルの方が多い。これらのことから、ある学習状況から異なる学習状況に移行する局面における推定性能の向上に隠れマルコフモデルの forward アルゴリズムが寄与するといえる。

しかし、正解率は 0.540 と良好ではなく、先行研究の課題でもあった、ある学習状況から異なる学習状況に移行する局面における学習状況の推定については、本研究で提案した自動推定モデルも有用ではなかった。

7. おわりに

7.1 本研究の成果

本研究では、ニューラルネットワークと隠れマルコフモデルの forward アルゴリズムから構成される自動推定モデルを用いて学習者の学習状況を推定する手法を設計し、その構築と評価から有用性の検証を行った。その結果、実際の学習を模したタスクに対して高い推定精度を示し、筆記具の動きから学習者の学習状況を推定する手法として提案する自動推定モデルが有用であることが示された。また、隠れマルコフモデルの forward アルゴリズムによるニューラルネットワークの推定結果の補正は、ある学習状況から異なる学習状況に移行する局面における学習状況の推定精度の向上において有用であることが示された。

7.2 今後の課題

自動推定モデルは学習者の筆記具の持ち方や動かし方によって「手が止まっている」状況を「筆記具を置いている」と誤って推定する傾向がある。筆記具の持ち方や動かし方の個人差に対応するため、自動推定モデルのパラメータの再検討が必要である。特にある学習状況から異なる学習状況に移行する局面の正解率は 53.97% と低く、より正解率を高めるための検討が必要である。

筆記データの収集は大学生のみを対象としたが、小学生

と大学生では手の大きさが異なり、筆記具の動きが異なることが考えられる。実用においては対象とする学習者に沿った筆記データの収集を行う必要がある。

学習者の学習状況を推定し教育に役立てる自動推定モデルにするには、「集中している」などの本研究で推定した3つの学習状況より具体的な学習状況の推定が必要である。しかし、このような具体的で教員の指導や学習者の学習の振り返りに有用な推定は、本稿で構築した自動推定モデルだけでは不可能と考えられ、教室内の音声や映像などの筆記具の動き以外のセンシングデータと組み合わせるなどの工夫が必要である。

6 学年国語

参考文献

- [1] 文部科学省：児童生徒の学習評価の在り方について（平成30年中央教育審議会答申）
- [2] 文部科学省：豊かな学びに繋がる学習評価の在り方（令和7年7月8日）
- [3] 文部科学省：令和6年度学校における教育の情報化の実態等に関する調査結果（令和7年3月1日現在）
- [4] 文部科学省：新時代の学びを支える先端技術活用推進方策（最終まとめ）
- [5] 文部科学省：大規模アンケート調査等の実施による学習者用デジタル教科書の効果・影響等の把握・分析等に関する実証研究事業成果報告書（令和6年3月26日）
- [6] 文部科学省：学習用デジタル教科書実践事例集（2022年3月）
- [7] 文部科学省：StuDX Style 事例 「取り組み状況をすばやく把握して個別サポート」
- [8] CAI Yuying, 柴田好章：表情分析の活用による授業分析の試み—教師の発話内容による子どもの感情の変化に注目して—, 日本教育工学会研究報告集, vol.4, pp.216-222 (2021)
- [9] 寺井友太, 田中碧, 藤村真生：生体情報を用いた集中度のフィードバック, 情報処理学会第83回全国大会講演論文集, vol.1, pp.247-248 (2021)
- [10] Shao Bochao, 金尚泰：深層学習を用いたユーザの生体情報認識に基づくリアルタイム集中度検出手法の検討, 情報処理学会第84回全国大会講演論文集, vol.1, pp.451-452 (2022)
- [11] コクヨ：しゅくだいやる気ペン
<https://www.kokuyo-st.co.jp/stationery/yarukipen/>
（最終閲覧：2026年1月）
- [12] 江木啓訓, 尾澤重知：学習者センシングのための筆記行為の検知手法と評価, 日本教育工学会論文誌, Vol.36, Suppl, pp.181-184 (2012)
- [13] 江木啓訓, 尾澤重知：学習者センシングシステムのための筆記行為の検知手法, インタラクション2012 論文集, pp.275-280 (2012)
- [14] Witmotion：WT901BLECL
<https://witmotion-sensor.com/products/bluetooth-5-0-accelerometer-inclinometer-wt901blecl-mpu9250-high-precision-9-axis-gyroscope-anglexy-0-05-accuracy-magnetometer-with-kalman-filter-low-power-3-axis-ahrs-imu-sensor-for-arduino>
（最終閲覧：2025年12月）
- [15] コクヨ：キャンパスノート
<https://www.kokuyo-st.co.jp/stationery/campus/>
（最終閲覧：2025年3月）
- [16] sony：Neural Network Console
<https://dl.sony.com/ja/>
（最終閲覧：2025年3月）
- [17] 国立教育政策研究所：令和6年度全国学力状況調査小学校第