

# オンライン手書きデータを用いた定着問題の自動推定 ー幾何数学問題を対象としてー

三浦 将人<sup>†1</sup> 工藤 雅士<sup>†1</sup> 岡田 一洸<sup>†1</sup> 中山 祐貴<sup>‡2</sup> 山名 早人<sup>§3</sup>

<sup>†1</sup> 早稲田大学大学院基幹理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1

<sup>‡2</sup> 早稲田大学グローバルエデュケーションセンター 〒169-0051 東京都新宿区西早稲田 1-6-1

<sup>§3</sup> 早稲田大学理工学術院 〒169-8555 東京都新宿区大久保 3-4-1

E-mail: <sup>†</sup> § {masato, kudoma34,k-okada,yamana}@yama.info.waseda.jp,

<sup>‡</sup> nakayama@aoni.waseda.jp

**あらまし** 近年, タブレット端末を用いた学習用アプリケーションの利用が広がりを見せている. 一般に学習においては学習者一人一人の定着度に合わせた問題を出題することが重要となる. 教室における学習や自習, 及びアプリケーションを用いた学習において, 定着を測る際には解答の正誤情報を用いるのが一般的である. しかし, 学習者が正答・誤答する理由は様々であり, 定着判定を正誤情報から判断することは適当でない. そこで, 本研究では学習の定着を推定する手法として, オンライン手書きデータを活用する手法を提案する. 多数の幾何数学問題における解答データを収集し, オンライン手書きデータから得られる筆記速度, 筆圧等の特徴量を用いて XGBoost(Extreme Gradient Boosting), SVC (Support Vector Classification) により定着を予測する.

**キーワード** 学習分析, Pen-based Computing, 手書き, 幾何問題, XGBoost, SVC

## 1. はじめに

近年, スマートフォンなどのタブレット端末を用いた学習用アプリケーションの利用が広がりを見せている. 令和 3 年 3 月に内閣府から発表された「令和 2 年度 青少年のインターネット利用環境実態調査」<sup>1)</sup>によると, 高校生の 64.6%, 中学生の 48.0%が学習を目的としたアプリケーションやサービスを利用していると回答している.「令和元年度青少年のインターネット利用環境実態調査」<sup>2)</sup>によると令和元年度の学習アプリケーション利用率は高校生 53.1%, 中学生 38.9%であり, 令和 2 年度の利用率は前年度から大きく増えている.

また, 教育の場におけるタブレット端末, ノート PC などの端末利用について, 文部科学省は平成 31 年度から「GIGA スクール」<sup>3)</sup>という学校教育のデジタル化政策を進めており, 1 人 1 台の学習用コンピュータの配布を目指している. 特に近年では新型コロナウイルスの影響でオンライン授業の需要が高まっており, 文部科学省は「1 人 1 台端末」の実現について前倒しを進めている. また, 令和 3 年 10 月に文部科学省が発表した「端末利活用状況等の実態調査 (令和 3 年 7 月末時点) (確定値)」<sup>4)</sup>によると, 全国の公立小学校等の 96.2%, 公立中学校等の 96.5%が端末の利活用をしていると回答しており, 学習時の端末の利用が一般的になっている.

一般に学習においては復習により知識を定着させることが重要である. 既存のアプリケーションにおける復習システムとしては「スタディサプリー」<sup>5)</sup>等のアプリケーションにおいて, 過去に間違えた問題を記録し再出題を行う方法がある. 過去に間違えた問題のみが出題されるため, 学習者は効率的に学習を行うことができる. しかし, 学習者が正答・誤答する理由は様々であるため, 学習の定着判定を正誤情報から判断することは適当でない. 正誤情報から定着を判断した場合, 勘で解いて答えが一致した場合は再出題が行われず, ケアレスミスによる誤答は再出題されるなど, 効率的な学習の妨げになる場合が考えられるからである. 定着の判断を学習者自身が行う手法も考えられるが, 学習者自身が定着判定を行う場合, 誤答であっても解説を読み『十分に定着した』と判断する場合がある. そのため, 主観による定着判定は適切でない.

そこで, 本稿ではオンライン手書きデータを活用し学習者の定着度を推定する手法を提案する. オンライン手書きデータとは, 通常のアフライン手書きにおいては収集することができない, 記入した座標や筆圧などの時系列データを持つ手書きデータである. 幾何数学問題における多数の解答データを収集し, オンライン手書きデータから得られる筆記速度, 筆圧等の特徴量を用いて XGBoost(Extreme Gradient Boosting), SVC (Support Vector Classification) により定着推定を行う.

以下, 2 節で関連研究について述べ, 3 節で定着度推定手法を提案する. 4 節で推定に用いるデータ収集

<sup>1</sup> [https://www8.cao.go.jp/youth/kankyouto/internet\\_torikumi/tyousa/r02/net-jittai/pdf/sokuhou.pdf](https://www8.cao.go.jp/youth/kankyouto/internet_torikumi/tyousa/r02/net-jittai/pdf/sokuhou.pdf)

<sup>2</sup> [https://www8.cao.go.jp/youth/kankyouto/internet\\_torikumi/tyousa/r01/net-jittai/pdf/sokuhou.pdf](https://www8.cao.go.jp/youth/kankyouto/internet_torikumi/tyousa/r01/net-jittai/pdf/sokuhou.pdf)

<sup>3</sup> [https://www.mext.go.jp/content/20210608-mxt\\_jogai01-000015850\\_003.pdf](https://www.mext.go.jp/content/20210608-mxt_jogai01-000015850_003.pdf)

<sup>4</sup> [https://www.mext.go.jp/content/20211125-mxt\\_shuukyoku01-000003278\\_001.pdf](https://www.mext.go.jp/content/20211125-mxt_shuukyoku01-000003278_001.pdf)

<sup>5</sup> <https://studysapuri.jp/>

について述べ、5 節で実験結果を示し 6 節で本稿についてまとめる。

## 2. 関連研究

学習履歴を活用した研究、幾何数学問題を対象とした研究、定着度推定を行った研究を紹介する。

### 2.1. 学習履歴を活用した研究

学習履歴を活用した研究として、日本女子大学の藤巻ら[1]は、2021 年に主成分分析及び自己組織化マップを用いて LMS (Learning Management System: 学習管理システム) の学習履歴データを分析し、ドロップアウトの兆候検知について検討を行っている。また、愛知大学の土橋[2]は、2020 年に教材閲覧時のページ遷移データと小テストの得点から学習者のエンゲージメントレベルを可視化し、学習不振に繋がるつまづきを検知できる可能性があることを示した。

### 2.2. 幾何数学問題を対象とした研究

幾何数学問題を対象とした論文として、早稲田大学の森山[3]は 2018 年にオンライン手書きデータを用いて「解法の分類と問題を勘で解いたか否かの分類」を行った。問題を勘で解いたか否かの判定においてオンライン手書きデータを元にした判定により 0.83 の正解率を得ており、オンライン手書きデータの活用が有効であることを示した。

また、早稲田大学の村上[4]は 2019 年にオンライン手書きデータを用いて解法の分類及びヒント参照有無の判定を行っている。村上は図への書き込み内容の特徴量として用いることにより解法の分類における精度が上昇することを示した。また、正答かつサンプル数の多い解法データに限定し分類を行うことによりヒント参照有無の精度が向上することを示した。

村上の論文を元にした著者ら[5]は 2020 年に、ヒント参照有無において問題文・幾何図中への記入における筆圧などの特徴量を用いた手法を提案した。追加した特徴量により、ヒント参照有無の判定において平均 0.636 の F 値を得、一定の有効性があることを示した。また、著者ら[6]は 2021 年に 3 段階の問題難易度毎にヒント参照有無前後における特徴量の変化を調査し、筆記速度の平均値が問題難易度に関わらず、ヒント参照有無の判定に有効に働くことを示した。

### 2.3. 定着度推定を行った研究

定着度推定を行った論文として、中央大学の福井[7]は 2017 年に視線情報を用いて文章の理解度を推定しており、サッカード時間の標準偏差が文章理解度の推定に有効であることを示した。また、NTT コミュニケ

ーション科学基礎研究所の東中ら[8]は 2003 年に視線情報から得られる決定木が文章理解度の推定に有効であることを示した。福井大学の森ら[9]は 2018 年に、英単語の 4 択問題を対象に定着度及び主観的難易度の推定を行っており、ウェアラブルデバイスから得られたデータから定着度合いを推定する場合 FFT 変換を行ったデータを用いることにより精度が向上することを示した。早稲田大学の浅井ら[10]は 2014 年に漢字の暗記定着度の予測を行い、連続値による定着率出力を実現した。

### 2.4. まとめ

幾何数学問題を対象とした研究においては解法の分類[3][4]、ヒント参照有無の判定[4][5][6]がこれまでに行われているが現時点で定着度の推定を行っているものは、著者らが調査した限り存在しなかった。

定着度を自動推定した研究においては視線情報を用いることによる有効性が示されている[7][8]が、視線情報を得るためには顔が端末の正面にある必要があり、学習者とタブレットの位置関係が固定化される問題点がある。オンライン手書きデータを用いた定着度の予測の自動推定を行ったものとしては漢字のような暗記学習における定着度推定[10]において一定の有効性が示されているが、幾何数学問題のような論理的思考力を問われる問題に対する有効性は示されていない。

## 3. 定着度推定手法の提案

### 3.1. 推定手法の概要

本稿では、オンライン学習を行う学習者がより効率的な学習を行うための出題システムを構築するために、オンライン手書きデータを活用した学習者の定着推定を目指す。具体的には、幾何図形問題を対象に、オンライン手書きデータから得られる特徴量から、定着問題を機械学習によって推定する手法の構築を行う。

ここで、本稿における「定着」を定義する。「定着」とは単なる正解・不正解でなく、ケアレスミスがあった場合であっても解法が合っていれば「定着している」と判断すべきである。また、勘で正解した場合は「定着していない」と判断すべきである。また、また、学習者が定着しているか否かを判断する場合があるが、学習者の「理解できている」判断は信頼できないことが示されている。

本稿で行う被験者実験においては、以下のように実験を行い「定着した問題」を定義する。

- 1) 被験者は 2 週間の間隔を空けて 2 回実験を行う。
- 2) 1 回目の実験では被験者は解法を閲覧しながら問題を解くことができる。

- 3) 2 回目の実験では 1 回目の実験において解いた問題の類題（数値，図の向きを変えたもの）に解答する．
- 4) 2 回目の実験において正解した問題を「定着した問題」と定義する．この時，ケアレスミスによる不正解は正解と同様に扱う．また，アンケートにおいて「理解できなかった/問題の意味が分からない」と回答されている場合は「定着していない問題」とする．

### 3.2.使用する特徴量

今回の推定で使用する特徴量は著者らの論文[5][6]において用いた特徴量に加えて「解法閲覧時間」「解法閲覧時間割合」を用いる．「解法閲覧時間」，「解法閲覧時間割合」は長い程解答者が理解しようとしている，すなわち定着度が高くなるという仮定に基づいて追加した．本手法でを使用した特徴量を以下の表 3.1 に示す．

なお，表 3.1 内で説明に用いている「問題文及び幾何図形部分」とは，左上端(0,0)，右下端(1024,1128)とした矩形形状の解答スペースにおいて，左上端を(0,0)，右下端を(1024,447)とした矩形部分を示す．また，左上端(0,447)，右下端(1024,1128)とした矩形部分を「問題文及び幾何図形部分以外」とする．

表 3.1 で用いている記号の意味は以下の通りである．

$time_s$  : 解答開始時刻

$time_e$  : 解答終了時刻

$P_{i,j}$  :  $i$  番目のストローク（全ストローク数を  $p$  とすると  $1 \leq i \leq p$ ）における  $j$  番目のサンプリング点(サンプリング数を  $q$  とすると  $1 \leq j \leq q$ )

$o\_time_i$  : 「解法表示」ボタン(図 4.1)により  $r$  回中  $i$  回目に解法を表示した時刻（ただし， $r$  は当該被験者が解法表示ボタンを押した回数）

$c\_time_i$  : 「解法表示」ボタン(図 4.1)により  $r$  回中  $i$  回目に表示した解法を非表示にした時刻

各サンプリング点  $P_{i,j}$  は，以下の 4 情報を持つ．

$x_{i,j}^{option}$ :  $x$  座標，  $y_{i,j}^{option}$ :  $y$  座標，

$f_{i,j}^{option}$ : 筆圧，  $t_{i,j}^{option}$ : 筆記時刻

ここで，「問題文及び幾何図形部分」に存在するストロークのサンプリング点は， $option$  に  $in$  を記す．「問題文及び幾何図形部分以外」に存在するストロークのサンプリング点は， $option$  に  $out$  を記す．さらに，「消しゴム機能（4.1.データ収集に用いたアプリケーション

参照）により消去されていないストロークにおけるサンプリング点は， $option$  に  $ex$  を追記する． $option$  が記載されていないサンプリング点は，全てのサンプリング点を示す．

表 3.1.推定に用いた特徴量

特徴量		定義
経過時間		解答開始から終了までにかかった時間[ms] $time_e - time_s$
時間間隔	平均 標準偏差 最大	$i$ 番目のストロークを書き終えてから $i+1$ 番目のストロークを書き始めるまでの時間[ms] $t_{(i+1),1} - t_{i,q}$
ストローク残存率		問題文及び幾何図形部分以外への記入における全ストロークの描画時間の内、残存しているストロークの描画時間の割合 $\frac{\sum_{i=1}^p (t_{i,q}^{ex} - t_{i,1}^{ex})}{\sum_{i=1}^p (t_{i,q} - t_{i,1})}$
速度	平均 標準偏差 最大 最小	問題文及び幾何図形部分以外への記入におけるストロークの筆記速度[pt/ms] $\frac{\sqrt{(x_{i,j+1}^{out} - x_{i,j}^{out})^2 + (y_{i,j+1}^{out} - y_{i,j}^{out})^2}}{t_{i,j+1}^{in} - t_{i,j}^{in}}$
筆圧	平均 標準偏差 最大	問題文及び幾何図形部分以外への記入における，UItouch クラスの <code>force</code> メソッドにより取得した電子ペンのタッチ圧力 $f_{i,j}^{out}$
停滞時間率		ストローク間時間の累計 解答開始から解答終了までの時間 $\frac{\sum_{i=1}^{p-1} (t_{i+1,1}^{in} - t_{i,q}^{in})}{time_e - time_s}$
図中速度	平均 標準偏差 最大 最小	問題文及び幾何図形中への記入におけるストロークの筆記速度[pt/ms] $\frac{\sqrt{(x_{i,j+1}^{in} - x_{i,j}^{in})^2 + (y_{i,j+1}^{in} - y_{i,j}^{in})^2}}{t_{i,j+1}^{in} - t_{i,j}^{in}}$
図中筆圧	平均 標準偏差 最大	問題文及び幾何図形中への記入における UItouch クラスの <code>force</code> メソッドにより取得した電子ペンのタッチ圧力 $f_{i,j}^{in}$
図中ストローク残存率		問題文及び幾何図形中への記入における全ストロークの内、残存するストロークの描画時間の割合 $\frac{\sum_{i=1}^p (t_{i,q}^{in,ex} - t_{i,1}^{in,ex})}{\sum_{i=1}^p (t_{i,q}^{in} - t_{i,1}^{in})}$
解法閲覧時間		「解法閲覧」ボタンを用いて解法を閲覧した時間[ms] $\sum_{i=1}^r (c\_time_i - o\_time_i)$
解法閲覧時間割合		解法閲覧時間 解答開始から解答終了までの時間 $\frac{\sum_{i=1}^r (c\_time_i - o\_time_i)}{time_e - time_s}$

### 3.3.推定手法

定着の推定には XGBoost, SVC を用いる。解答者の 1 回目の実験におけるオンライン手書きデータから、表 3.1 において示した特徴量の抽出を行い、特徴量が問題に依存しないよう、各特徴量について問題ごとに標準化を行った。すなわち、問題ごとに平均値からの偏差を標準偏差で割った。その後、標準化した解答者のオンライン手書きデータを使用し XGBoost, SVC の構築を行う。

その後、定着度推定に対する寄与率が高い特徴量の順位付けを行う。具体的には、XGBoost, SVC それぞれにおいて表 3.1 に示した特徴量のセットから 1 種類の特徴量を消去した特徴量セットを作成する。収集した解答データの 30% を用いて特徴量セットごとに定着問題の推定を行い、消去した際に F 値の低下が大きいものほど消去した特徴量の寄与が大きいと判断し特徴量の順位付けを行う。順位が上位のものから順に定着率の推定に使用する特徴量を選出することにより、最も精度よく予測できる特徴量の組み合わせを求める。求めた特徴量セットを用いて残りの 70% の解答データを対象として定着問題の推定を行う。

最終の評価においては、取得したデータ数が少ないため、Leave-one-out 交差検証を用いる。推定結果について、正解率により評価する。

また、1 回目の正誤データによる判定、被験者の主観評価による判定を行い、結果を比較する。

## 4. データ収集

### 4.1.データ収集に用いたアプリケーション

被験者実験において用いたアプリケーションについて述べる。

データの収集には村上の論文[4]において用いられたアプリケーション「GeoSol」を利用した(図 4.1)。

「GeoSol」は iPad Pro 上に幾何数学問題を表示し、Apple pencil を用いて画面上に自由に線を書くことができるアプリケーションである。

画面左上のペンアイコン、消しゴムアイコンをタップすることによりペンと消しゴムの機能を切り替えることができる。ペンは図 4.1 に示すように画面下部の解答スペースに加え、問題文や図中への書き込みも可能である。消しゴムは記述済みのストロークをタップすることによりタップしたストロークを削除することができる。

また、画面右上の「解法表示」ボタンをタップすることによりその問題の解法を閲覧することができる。

「できた」ボタンをタップすることで画面のスクリーンショットを撮影し、記入したストロークのオンライン手書きデータ(記入した座標や筆圧などの時系列デ

ータを持つ手書きデータ)がテキストファイルとして iPad 内に保存される。

「GeoSol」を用いた幾何数学問題解答の流れは以下のとおりである。

- 1) アプリケーション側：問題画像を表示する。
- 2) ユーザ：Apple pencil により書き込みを行う。
- 3) ユーザ：(解法を確認したい場合)「解法表示」ボタンをタップする
- 4) アプリケーション側：解法を表示する
- 5) ユーザ：解法確認後「この方法で解く」ボタンをタップする
- 6) ユーザ：「できた」ボタンをタップする
- 7) アプリケーション側：「次の問題」ボタンを表示する
- 8) ユーザ：次の問題へ移行する

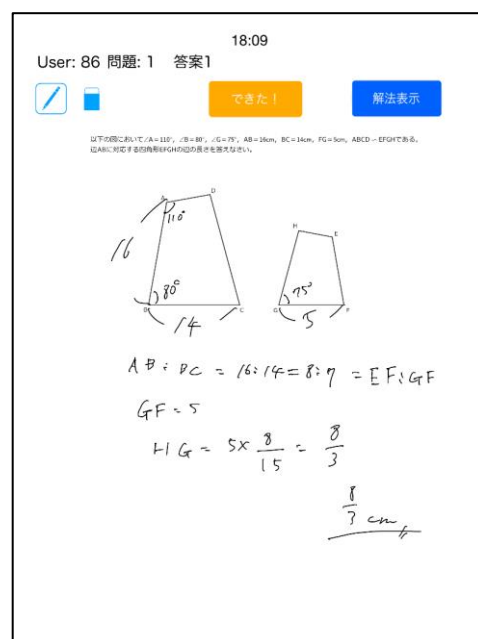


図 4.1 使用したアプリケーション

### 4.2.データ収集に用いた問題

今回推定に用いた問題は「チャート式中学数学 3 年」<sup>6</sup>から相似、円、三平方の定理の問題を計 60 問著者が改題し出題した。具体的には、図中の長さの変更、図の反転・回転により作成した。

Q1~Q10 は相似の問題、Q11~Q20 は円の問題、Q21~Q30 は三平方の定理の問題であり、Q31~Q60 はそれぞれ Q1~Q30 の類題である。

<sup>6</sup> チャート研究所『チャート式中学数学 3 年』数研出版、2021 年

### 4.3.被験者実験

「GeoSol」を用いて早稲田大学の学生 52 名（男性 32 名，女性 20 名，18～26 歳）を対象にオンライン手書きデータ収取実験を行った．データ収集は 2 回に分けて行い，1 回目の実験では Q1～30，または Q31～60 を 1 問あたり 5 分の制限時間で解いてもらった．1 回目の実験の翌週に 2 回目の実験を行い，2 回目の実験では 1 回目の実験で Q1～30 を解いた被験者(25 名)には Q31～60，1 回目の実験で Q31～60 を解いた被験者(27 名)には Q1～30 を解いてもらった．また，10 問毎に 5 分程度の休憩を取ってもらった．

1 回目の実験において解答者は「解法表示」ボタンをタップすることにより問題の解法を見ながら問題を解くことができる．2 回目の実験においては「解法表示」ボタンをタップしても解法は表示されない．

また，1 問終了するごとにアンケートの回答を依頼した．アンケートの内容を表 4.1 に示す．

表 4.1.実施したアンケート内容

質問内容	解答方法
あなたが感じた問題難易度	1.非常に簡単 2.簡単 3.比較的簡単， 4.比較的難しい 5.難しい 6.非常に難しい
制限時間をどう感じたか	1.非常に短い 2.短い 3.比較的短い 4.比較的長い 5.長い 6.非常に長い
同じ問題が出題された際に正解する自信の有無	1.かなり自信がある 2.自信がある， 3.どちらかと言えば自信がある 4.どちらかと言えば自信がない 5.自信がない 6.かなり自信がない
各プロセスを理解できたか	1.解法を見る前に理解できた 2.解法を見て理解できた 3.理解できなかった/問題の意味が分からない 4.別の手法で解いた

表 4.1 の「各プロセスを理解できたか」の質問につ

いては各問題の回答に必要なプロセスを列举し一つ一つのプロセスについて理解できたか否かを確認した．

## 5. 実験結果と考察

### 5.1.データの前処理

52 名の被験者を対象に 1 人あたり 30 問の問題を解いてもらい，オンライン手書きデータの収集を行った結果 1 回目の実験において，1,560 件中 1,444 件（Q1～30：818 件，Q31～60：626 件）のオンライン手書きデータを正常に収集した．116 件のオンライン手書きデータは被験者が「できた！」ボタンを押さずにアプリケーションを終了した，解答途中で問題が発生した，体調不良により解き切らずに帰宅した等の理由で収集できなかった．

2 回目の実験でアンケートの収集を行った．表 4.1 に示した 2 回目の実験のアンケートにおいて，「各プロセスを理解できたか」の設問に対し「別の手法で解いた」と回答した答案については，定着・未定着の判断が困難であるため今回の推定の対象から除いた．また，2 回目の日程の都合が合わなくなりアンケートが収集できなかった答案について今回の推定の対象から除いた．取り除いた結果，1,268 件（Q1～30：695 件，Q31～60：573 件）のオンライン手書きデータとアンケート結果が残った．

1,268 件のオンライン手書きデータについて，表 3.1 に示した特徴量の抽出を行った．

99 件の解答は「問題文及び幾何図形部分」への記入のみが行われていた．そのため，「問題文及び幾何図形部分」以外の書き込みを対象とする「ストローク残存率」「速度」「筆圧」の項目が欠損していた．欠損部分については，同オンライン手書きデータの「図中ストローク残存率」「図中速度」「図中筆圧」の値により穴埋めを行った．

同様に，3 件の解答は「問題文及び幾何図形部分」への記入が行われておらず「図中ストローク残存率」「図中速度」「図中筆圧」の項目が欠損していた．欠損部分については「ストローク残存率」「速度」「筆圧」の値による穴埋めを行った．

### 5.2XGBoost のパラメータ調整

2 値分類であるため

‘objective’: ‘binary:logistic’

‘eval\_metric’: ‘error’

として 2 クラス向けのロジスティック回帰による学習により 2 クラス分類のエラー率が小さくなるように学習を行った．

また，収集したデータからランダムに 30%を用いて学習ラウンド数を変えながら表 3.1 に示す特徴量をも

ちいて Leave-one-out 交差検証により推定を行った。ラウンド数と正解率の関係を以下の図 5.1 に示す。

ラウンド数 4 で正解率の上昇が緩やかになり、20 ラウンドまで実施した結果 12 ラウンドで正解率は最大 0.690 となった。よって、以降はラウンド数を 12 として実験を行った。

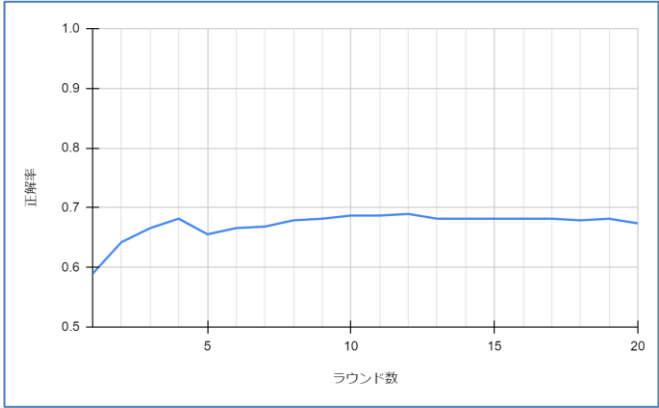


図 5.1 正解率とラウンド数の関係

順位付け結果をもとに、上位から特徴量を追加することにより最適な特徴量の組み合わせを見つける。上位から特徴量を追加を行い、順位付けに用いた 30% のデータを対象に Leave-one-out 交差検証により正解率を求めた結果を以下の図 5.2、図 5.3 に示す

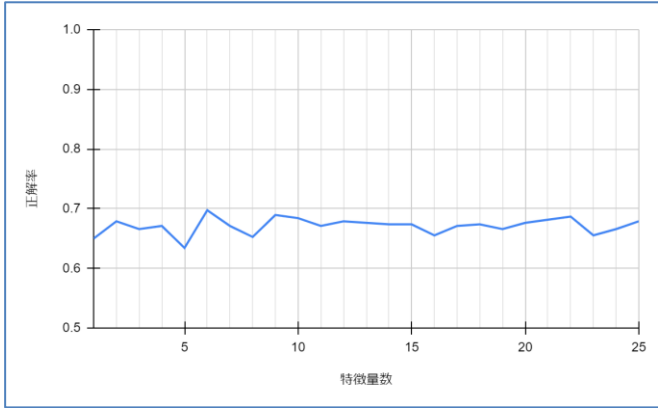


図 5.2. XGBoost における正解率

5.3.特徴量順位付け

3.3「推定手法」で示した手法により、5.2「XGBoost のパラメータ調整」で用いた 30% の解答データにより特徴量の順位付けを行った。

XGBoost, SVC それぞれにおける順位付け結果を以下の表 5.1 に示す。

表 5.1 特徴量順位付け結果  
(黄色マーカ部分が最終的に採用した特徴量)

順位	XGBoost	SVC
1	解法閲覧時間	解法閲覧時間
2	速度 標準偏差	時間間隔 平均
3	速度 平均	時間間隔 標準偏差
4	時間間隔 最大	停滞時間率
5	筆圧 第三四分位数	速度 第一四分位数
6	経過時間	國中速度 標準偏差
7	速度 第一四分位数	筆圧 第一四分位数
8	國中ストローク残存率	経過時間
9	國中筆圧 第一四分位数	國中筆圧 第一四分位数
10	筆圧 平均	筆圧 平均
11	時間間隔 平均	國中筆圧 標準偏差
12	國中速度 標準偏差	解法閲覧時間割合
13	國中筆圧 第三四分位数	國中筆圧 第三四分位数
14	ストローク残存率	國中筆圧 平均
15	國中速度 第三四分位数	速度 標準偏差
16	解法閲覧時間割合	筆圧 第三四分位数
17	國中筆圧 標準偏差	國中速度 第一四分位数
18	停滞時間率	筆圧 標準偏差
19	國中速度 第一四分位数	速度 平均
20	筆圧 第一四分位数	國中速度 第三四分位数
21	時間間隔 標準偏差	國中速度 平均
22	國中速度 平均	時間間隔 最大
23	速度 第三四分位数	速度 第三四分位数
24	國中筆圧 平均	ストローク残存率
25	筆圧 標準偏差	國中ストローク残存率

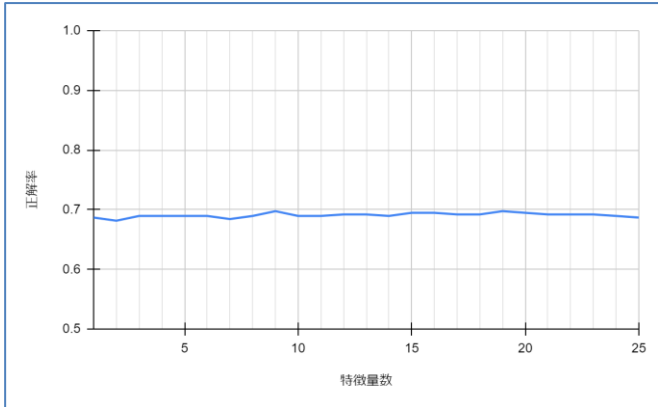


図 5.3. SVC における正解率

XGBoost においては上位 6 個の特徴量を用いた際に正解率は最大 0.697 となった。そのため、推定においては表 5.1 の上位 6 個の特徴量を採用した。

SVC においては上位 9 個の特徴量を用いた際に正解率は最大 0.697 となった。そのため、推定においては表 5.1 の上位 9 個の特徴量を採用した。

5.4.推定結果

5.3「特徴量順位付け」において示した特徴量セットを用いて、特徴量順位付けに用いなかった 70% のデータを用い、問題ごと及び解答者ごとに分割した leave-one-group-out 交差検証により定着推定を行った。

また、「定着した問題」を「適合」として学習を行った。すなわち、「適合率」とは「定着していると推定された問題の内、実際に定着している割合」を指し、「再現率」とは「定着している問題の内、定着してい

ると推定された割合」を指す。

また、比較のため 1 回目の正誤データを用いた推定、すなわち「1 回目の解答が正答であれば 2 回目も正解、1 回目の解答が誤答であれば 2 回目も誤答である」とした結果について検証した。

また、学習者の判断による推定、すなわち 1 回目の実験において、表 4.2 に示すアンケートの「同じ問題が出題された際に正解する自信の有無」の項目で「かなり自信がある」「自信がある」「どちらかと言えば自信がある」と回答した場合は 2 回目は正答する、「かなり自信がない」「自信がない」「どちらかと言えば自信がない」と回答した場合は 2 回目は誤答するとした結果について検証した。

leave-one-group-out 交差検証において問題ごとに分割した場合、すなわち当該問題の解答を含まず学習器を構築した場合の結果を表 5.2 に示す。また、解答者ごとに分割した場合、すなわち第三者のデータのみで学習器を構築した場合の結果を表 5.3 に示す。

表 5.2 当該問題の解答を含まず学習する場合

	正解率	適合率	再現率	F 値
XGBoost	0.666	0.710	0.845	0.771
SVC	0.687	0.694	<b>0.949</b>	0.802
正誤情報	<b>0.738</b>	<b>0.808</b>	0.796	0.802
学習者判断	0.726	0.770	0.841	<b>0.804</b>

表 5.3 第三者のデータのみで学習する場合

	正解率	適合率	再現率	F 値
XGBoost	0.632	0.690	0.816	0.748
SVC	0.667	0.677	0.956	0.793
正誤情報	<b>0.738</b>	<b>0.808</b>	0.796	0.802
学習者判断	0.726	0.770	<b>0.841</b>	<b>0.804</b>

結果として XGBoost、SVC の両方において当該問題の解答を含まず学習器を構築した場合の推定結果の方が、第三者のデータのみで学習器を構築した場合に比較し、正解率が高いことが分かる。つまり、本稿で用いた特徴量は、解答者に依存する特徴を必ずしも排除できていないことが分かる。

一方で何れの場合においても、正誤情報や学習者判断に比較して高い再現率を得ることができており、特に SVC を利用した提案手法において顕著である。すな

わち、提案手法は、従来の正誤情報や学習者判断による定着判断に比較して、より漏れを無くすことができることが分かる。このことから、定着している問題の再出題を行いたくない場合、提案手法が有効である可能性がある。

ここで、正解率の向上が見られなかった理由として、「易問・難問における推定に今回用いた特徴量が有効でない」ことが 1 つの理由であると考えた。易問においては記述量が少ない問題が多く、特徴量をうまく抽出することが難しい。また、難問においては解くことができる解答者であっても悩む時間が長くなるため「経過時間」「停滞時間率」などの特徴量において差が生じにくくなると考えた。さらに、易問・難問を対象とした場合、正誤情報や学習者判断での定着・未定着の判断が容易になると考えた。

そこで、「易問・難問以外の解答データ」に対する推定結果を確認した。ここで、「易問・難問」の判断には表 4.2 に示すアンケートの「あなたが感じた問題難易度」の平均値を用いた。

難易度の平均が上位 25% 以内であった 7 問、及び下位 25% 以内であった 7 問を除いた、30 問中 16 問の問題に対する推定結果を以下の表 5.4、表 5.5 に示す。

表 5.4 当該問題の解答を含まず学習する場合  
(易問・難問を除くデータセット)

	正解率	適合率	再現率	F 値
XGBoost	0.647	0.708	0.831	0.764
SVC	<b>0.704</b>	0.718	<b>0.939</b>	<b>0.814</b>
正誤情報	0.687	<b>0.771</b>	0.776	0.774
学習者判断	0.696	0.749	0.840	0.792

表 5.5 第三者のデータのみで学習する場合  
(易問・難問を除くデータセット)

	正解率	適合率	再現率	F 値
XGBoost	0.611	0.687	0.801	0.739
SVC	0.677	0.695	<b>0.945</b>	<b>0.801</b>
正誤情報	0.687	<b>0.771</b>	0.776	0.774
学習者判断	<b>0.696</b>	0.749	0.840	0.792

易問・難問を除いた問題を対象とした推定においては、当該問題の解答を含まず学習した場合、すなわち

解答者自身のデータを学習させている場合において、SVCによる提案手法の正解率が最も高く0.704となった。

以上をまとめると、ユーザは易問・難問に対する定着判断は高い正解率で行うことができるが、学習器では難しい。これは、「易問においては記述量が少なくなる」「難問においては解くことができる解答者であっても悩む時間が長くなるため「経過時間」「停滞時間率」などの特徴量において差が生じにくくなる」ためだと考えられる。一方で、易問・難問を除いた問題を対象とした場合、SVCを用いた学習による定着推定は、問題の正誤から定着を判断する方法やユーザ自身が定着を判断する手法に比較して高い正解率を得ることができると分かった。

なお、今回の実験から、現在用いている特徴量は解答者間での特徴量差をうまく吸収できておらず、解答者自身のデータを用いて学習器を構築しなければ定着推定が高い正解率で行えないことも分かった。

## 6. おわりに

本稿では、複数の過程を含む問題における定着推定を目的としてオンライン手書きデータを用いた手法を提案した。提案手法は、解答者間の違いをうまく吸収できていないものの、解答者のデータを用いて学習器を構築することで、定着推定を行うことが可能であることを示した。

特に、易問・難問を除いた問題に対しては高い推定正解率を得ることができ、正解率0.704を達成した。一方、易問・難問に対しては、問題の正誤からのみ定着を判断する手法やユーザ自身の定着・未定着判断の方が有効であった。

今後の展望としては、易問・難問の定着率推定の正解率向上手法の研究、解答者に依存しない推定手法の研究、さらには、実際の学習用アプリケーションに組み込むことにより効率的な学習の助けになるか調査を行うことが挙げられる。

## 謝辞

本研究の一部は科学研究費補助金(20H04292)によるものである。

## 参考文献

- [1] 藤巻星香, 小川賀代, “オンライン授業における学習履歴を用いた学習支援の検討”, 研究報告教育学習支援情報システム Vol.2021-CLE-33, No.14, pp.1-5, 2021
- [2] 土橋喜, “エンゲージメントヒートマップーMoodle ログのデータマイニングによる学習傾向の可視化ー”, 情報教育シンポジウム論文集, pp.23-30, 2020
- [3] 森山優姫菜, 山名早人, “シーケンス情報を用いて幾何学の手書き解答データを解答パターンに分類する手法”, 早稲田大学基幹理工学研究科修士論文, 2018
- [4] 村上統馬, 山名早人, “オンライン手書きデータを用いた幾何学図形問題に対する答案の自動分類と解答者の理解度推定”, 早稲田大学基幹理工学部卒業論文, 2019
- [5] 三浦将人, 村上統馬, 中山祐貴, 山名早人, “オンライン手書き解答データを用いた解答者のヒント参照有無の判定ー幾何学問題を対象としてー”, 電子情報通信学会信学技法, vol.119, no.393, pp.27-32, 2020
- [6] 三浦将人, 工藤雅士, 中山祐貴, 山名早人, “オンライン手書きデータを用いたヒント参照前後における変化の調査-幾何問題を対象として-”, 第13回データ工学と情報マネジメントに関するフォーラム(DEIM2021), C33-4, 2021
- [7] 福井優太, “視線情報に基づくユーザの理解度推定システムの開発”, 大学院研究年報 理工学研究科編 47, 中央大学, 2017
- [8] 東中竜一郎, 大野健彦, “視線に基づく文書理解度測定法とその応用”, 情報処理学会研究報告. HI, ヒューマンインタフェース研究会報告 102, pp.31-38, 2003
- [9] 森朝春, 長谷川達人, “ウェアラブルデバイスを用いた英単語4択課題における定着度合いおよび主観的難易度の推定”, 分散協調とモバイルシンポジウム 2018 論文集 (2018), pp.589-595, 2018
- [10] 浅井洋樹, 山名早人, “オンライン手書き情報を用いた未定着記憶推定システム”, 研究報告コンピュータと教育(CE), vol.2014-CE-127, no.1, pp.1-6, 2014