

# 能動的学習データを用いた学生の認知的関与と学習深度の分析

王 笑難<sup>†</sup> 大野 麻子<sup>††</sup> 孫 一<sup>†††</sup> 清光 英成<sup>†</sup>

<sup>†</sup>神戸大学大学院国際文化学研究科 〒657-8501 神戸市灘区鶴甲 1-2-1

<sup>††</sup>神戸大学 DX・情報統括本部情報基盤センター 〒657-8501 神戸市灘区六甲台町 1-1

<sup>†††</sup>神戸情報大学院大学情報技術研究科 〒650-0001 神戸市中央区加納町 2-2-7

E-mail: <sup>†</sup>217c303c@stu.kobe-u.ac.jp, kiyomitu@kobe-u.ac.jp, <sup>††</sup>ohno@kitty.kobe-u.ac.jp, <sup>†††</sup>sun@kic.ac.jp

あらまし 教育機関で実施されるオンラインやオンデマンド授業は、対面授業と同等以上の教育効果が求められている。閲覧履歴や解答履歴など多様な学習データが蓄積・活用されているものの、遠隔環境における学習の深度を正確に把握することは依然として困難である。本研究では、オンデマンドプログラミング授業において、学生の能動的学習データ（教材動画上で学生が行うノート・質問・コメントなどの注釈）を活用し、学生の学びの深さを定量的に測定する手法を提案する。動画注釈によってフレーム単位で取得される学習者の思考過程を、自然言語処理技術とブルーム教育分類法を組み合わせて解析し、拡張した分類テーブルを用いて注釈テキストを自動的に認知レベルへ分類する。結果として、授業の進行に伴い学生の認知スキルが低次から高次へ移行する傾向を定量的に捉えられ、授業設計の影響も把握できることが示唆された。本研究の成果として、学習深度を評価する客観的指標を提示し、指導法や個別支援を最適化するための有用な知見を提供できた点が挙げられる。

キーワード e-Learning, プログラミング教育, 深い学び, 学習評価

## 1. はじめに

近年、プログラミングや工学分野を中心にオンライン教育が急速に普及しており、教育機関で行われるオンラインやオンデマンド授業には、対面授業と同等以上の教育効果が求められている。一方、文部科学省は深い学びを重視し、「習得・活用・探究という学びの過程の中で、各教科等の特質に応じた見方・考え方を働きながら、知識を相互に関連付けてより深く理解したり、情報を精査して考えを形成したり、問題を見いだして解決策を考えたり、思いや考えを基に創造したりすることに向かう」ことと定義している[1]。こうした高次の思考力や問題解決力を伴う学習を遠隔環境で実践するには、対面授業のような即時的フィードバックや学習者間の相互作用が得られにくいため、学習者がどの程度認知面で深く理解しているかを把握する手法は依然として十分に確立されていない。

プログラミング教育においては、文法や構文といった基礎的な知識の記憶・理解から、複雑な問題の解決や応用を要する高次の認知活動まで、多層的なスキルが求められる。同じ出力を得られるプログラムであっても、学習者の思考過程やコードの理解度は大きく異なる可能性があるため、学生には基礎知識の習得に加え、複雑な課題に対して解決策を導き出す高次の思考が求められる。しかし、学習者を評価しようとすると、主に提出物（コードの成果物やレポートなど）を通じた評価に偏りがちである。こうした状況を踏まえ、遠隔プログラミング教育の品質を高めるには、学習者が実際にどのように思考し、どの程度理解を深めている

のかを定量的に把握できる枠組みが必要となる。

我々の先行研究では、オンデマンド形式のプログラミング授業において、プログラミング学習に特有的な動作などを参考にし、講義動画のシーンと学生の質問や学習ノートとを紐づけることにより、状況を特定する記述をほとんど必要としない、動画上に直接注釈を付与する機能をもつ学習支援システムを開発した[2]。導入実験の結果、学生は質問の難易度を下げて取り組めるようになり、ノート・質問・コメントといった注釈を通じて、教員から学生の考えに沿ったフィードバックが得られやすくなった。これらの注釈は、学習者が授業内容をどのように理解しているか、どんな疑問を抱いているかを示すため、学習の深い理解を把握するうえでも有益な情報を含んでいると考えられる[3]。

そこで本研究では、オンデマンドプログラミング学習における学習深度を、学習者が残す能動的学習データを用いて定量的に捉え、授業回数の進行に伴う変化を明らかにすることを目的とする。動画注釈によって取得される学習者の投稿データを、自然言語処理技術と Bloom Taxonomy [4]を組み合わせて解析し、拡張した分類テーブルを用いて注釈テキストを自動的に認知レベルへ分類することで、学生の学習深度を客観的に評価できる枠組みを構築する。

本稿の構成は以下のとおりである。第2章では関連研究を概観し、第3章で提案手法の詳細について説明する。第4章では実験の詳細について述べ、第5章では実験結果を示す。第6章では考察を行い、第7章で結論を述べる。

## 2. 関連研究

### 2.1 教育における深い学び。

近年、日本の教育機関では、従来の講義型・知識伝達型の授業形態に代わって、学習者が主体的に活動するアクティブ・ラーニングが導入されている。アクティブ・ラーニングでは、単に表面的な知識を習得するだけでなく、学習内容への深い理解が期待される。しかし、これらの学習効果を定量的に把握して評価するには、学生や教員に大きな負担を伴う場合がある。

Honda ら[5]は、書く・話す・発表するといったアウトプット活動への関与や、認知プロセスの外化に注目することで、アクティブ・ラーニングの質を測定するための尺度を開発している。こうした評価基準を活用すれば、教員側は学生の認知状態や学習状況を客観的に捉えられるため、その後の指導や支援を最適化しやすくなると考えられる。

しかしながら、大規模クラスやオンライン授業など学習環境が多様化するなかで、教員が目視や対面で学生の理解度を迅速かつ的確に把握することは容易ではない。そこで本研究では、テキスト解析などの客観的な指標を用いて、学生が高次の思考に至っているかどうかを検証する枠組みの構築を目指す。これにより、学習者同士のやりとりの可視化やフィードバックの効率化を図りつつ、深い学びの実現をより確かなものにできると期待される。

### 2.2 Bloom's Taxonomyについて。

Bloom's Taxonomy[4]は、教育目標を階層的に整理し、認知スキルを評価する枠組みとして広く知られている。従来のタキソノミーでは「知識（Knowledge）」「理解（Comprehension）」「応用（Application）」「分析（Analysis）」「総合（Synthesis）」「評価（Evaluation）」の6段階が設定されていたが、2001年に Anderson と Krathwohl によって改訂され、それぞれが「Remember」「Understand」「Apply」「Analyze」「Evaluate」「Create」という動詞形に置き換えられた[6]。これにより、学習者が実際に行う認知活動がより明確に捉えやすくなり、教育目標との整合性も高められるようになった。

さらに Bloom's Digital Taxonomy [7]では、テクノロジーやデジタルツールを用いた学習行動を考慮し、高次思考力をいかに促進できるかに焦点を当てている。もともとの6段階の認知スキル（記憶から創造まで）がデジタル技術に対応する動詞で再定義され、現代の教育課程や学習設計と密接に結びつけられるようになった。たとえば「Remembering」には「一覧化する」「識別する」「検索する」などが含まれ、「Creating」には「設計する」「構築する」「生成する」といった高次思考力を伴う行為が挙げられる。

本研究では、オンラインプログラミング教育において、Bloom's Digital Taxonomy が示す6つの認知プロセスを基盤とし、学生が行う注釈やコメントなどのテキストを自動分類する。これらの動詞群を整理した例が表1である。各カテゴリーに含まれる代表的な行動動詞を参照することで、どの段階の思考レベルに関する記述が注釈として残されているかを推定できるようになる。学習者の書き込みに現れるキーワードやフレーズを解析することで、彼らがどの段階の思考プロセスを実行しているかを客観的に推定することが可能となる。オンライン授業で得られる大量のテキストデータを有効に活用するためには、このタキソノミーの枠組みが指標として有用であると考えられる。

表1. BLOOM'S DIGITAL TAXONOMY における6カテゴリーの  
例示的動詞

カテゴリー	動詞
記憶する Remembering	定義する、記述する、複製する、記憶する、認識する、紐つける、…
理解する Understanding	関連つける、分類する、説明する、要約する、集約する、…
応用する Applying	構築する、実行する、使用する、実験する、編集する、…
分析する Analyzing	比較する、推論する、連結する、統合する、例証する、…
評価する Evaluating	チェックする、結論する、納得する、整える、省みる、批評する…
創造する Creating	構築する、改める、取り替える、想像する、推測する、仮定する、…

### 2.3 自然言語処理と学習分析への応用。

近年、教育分野では学習者が生成するテキストデータ（レポートやコメント、チャットログなど）の分析を自動化する目的で、自然言語処理（NLP）の技術が幅広く利用されている。なかでも注目されているのが、Word2vec [8]や GloVe [9]といった単語埋め込み（Word Embedding）のモデルである。これらは大規模なコーパスから周辺文脈情報を抽出し、単語同士の意味的・統語的類似度をベクトル空間で表現できる点が大きな特長となっている。

Word2vec では、Continuous Bag-of-Words (CBOW) または Continuous Skip-gram というニューラルネットワークモデルを用いて単語ベクトルを学習する。CBOW は周辺単語から中央の単語を予測し、Skip-gram は中央の単語から周辺単語を予測するアーキテクチャであり、大量の文書データから単語間の共起パターンを効果的に捉えることが可能である。これにより、教育現場でしばしば問題となるドメイン特化型の専門用語や略語の多様な使われ方を扱いやすくなり、単語間の潜在的な関連性を可視化できる。

### 2.3 本研究の位置づけ.

近年はビデオ注釈など学習者が残す能動的なデータを活用し、学習の質を可視化・評価しようとする試みが増えている。しかし、プログラミング教育のように認知スキルが多層的で、文法理解やコード記述といった低次スキルから複雑な問題解決を伴う高次スキルまで包含する領域では、学期全体を通じた学習者の認知的变化を動的に追跡する手法がまだ十分に確立されていない。

そこで本研究では、Bloom's Digital Taxonomy を基盤にしつつ、プログラミング教育向けに学習したWord2vec モデルを組み合わせることで、学習者の注釈テキストを自動的に認知レベルに割り当てる枠組みを構築する。これにより、学習者が投稿するノートや質問・コメントの内容から、学期の進行に伴う学習深度の変化を可視化し、教育実践における指導効果や個別支援の最適化に寄与することを目指す。

### 3. 提案手法

本研究では、オンラインプログラミング授業における動画注釈データを活用し、学生の学習深度を定量的に評価する枠組みを提案する。具体的には、Bloom's Digital Taxonomy を基盤とし、形態素解析とWord2vec による単語埋め込みを組み合わせることで、注釈テキスト内の語彙を高次思考力と関連づけて自動分類し、その傾向を可視化するものである。本手法の大きな特徴は、以下の3点にまとめられる。

(1) ビデオ注釈の直接活用。従来、遠隔授業においては、提出物（ソースコードやレポート）の評価を通じて学習状況を推定することが多かった。しかし本研究では、授業動画の特定フレームに学生が残す「ノート」「質問」「コメント」といった注釈テキストそのものをデータ源とし、学習者の思考過程をより細かく

捉えることを可能にする。すなわち、どのタイミングでどのような疑問や着想が生まれたかを直接反映する情報を定量化するアプローチである。

(2) プログラミング分野に特化した自然言語処理。注釈データの多くは、プログラミング特有の専門用語やコード断片、略語が含まれており、汎用的な日本語解析ツールでは十分に処理しきれない場合がある。そこで本研究では、IT専門用語辞書や授業教材のトランスクriptを組み合わせた独自コーパスを構築し、Word2vec の Skip-gram モデルを学習することで、プログラミングの文脈に即した類似度計算を可能にした。これにより、学生の注釈中に現れる多様な言い回しや専門用語を正確に捉え、適切な認知スキルカテゴリへマッピングできるようになる。

(2) Bloom's Digital Taxonomy 動詞表の拡張。本研究では、「Remember」「Understand」「Apply」「Analyze」「Evaluate」「Create」の6分類（Bloom's Digital Taxonomy）をベースとするが、プログラミング教育特有の行動動詞や用語を取りこむため、Word2vec を用いたタクソノミ学習の動詞拡張を行う。すなわち、各タクソノミカテゴリに含まれる動詞から類似度の高い単語を自動抽出し、分類対象となる語彙リストを大幅に拡充する。これにより、学生が異なる表現や略語を用いていても、高次思考力に関連する語彙を網羅的に把握できる。

以上の枠組みによって得られた注釈と分類結果を集約し、授業の進行に伴う学習深度の推移や、学生個人の認知スキルの変化を可視化・評価できる。具体的な手順としては、次のように進める。

1. データ収集：動画から学生が投稿した注釈を収集し、そのテキスト情報を取得する。
2. 前処理とキーワード抽出：形態素解析によって注釈テキストをトークン化し、ストップワード

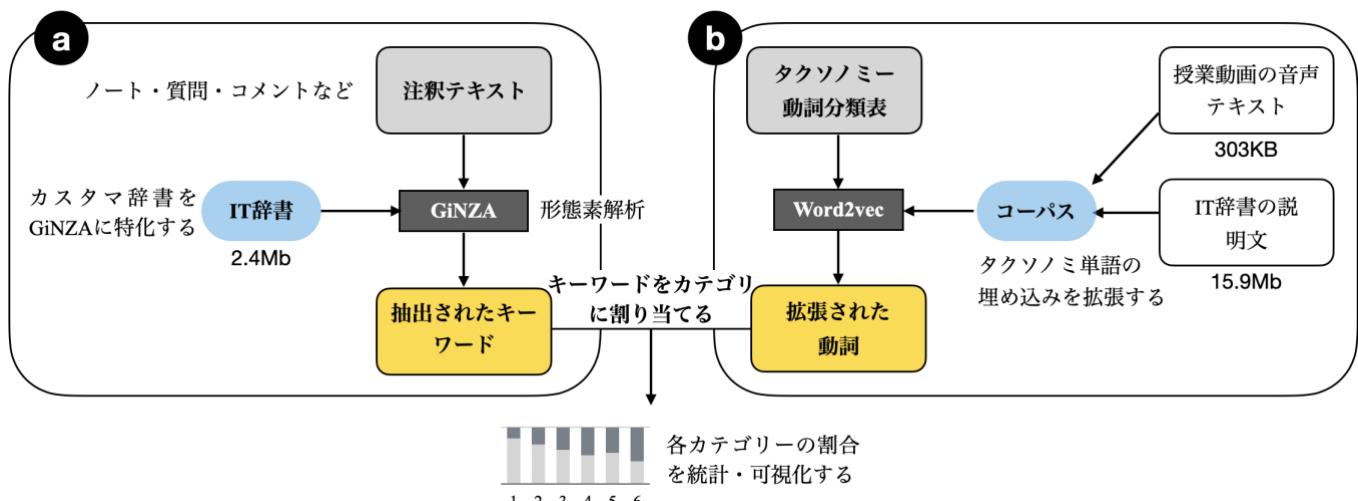


図 1. 動画注釈から学習深度の算出ワークフロー。

- 除去や正規化を行う（図 1a）。
3. コーパス構築と Word2vec 学習：プログラミング教育に特化したテキストコーパスを作成し、Word2vec モデルを学習する（図 1b）。
  4. タキソノミー拡張：学習した Word2vec モデルを用いて、Bloom's Digital Taxonomy に含まれる単語群を類似語で拡張する（図 1b）。
  5. キーワード分類：抽出したキーワードを拡張タキソノミーのいずれかのタキソノミーのカテゴリに自動的に分類する。
  6. 集計と可視化：授業回ごと、あるいは受講者ごとにタキソノミーのカテゴリ分布を集計・可視化し、学習深度の推移を評価する。

## 4. 実験

### 4.1 実験概要とデータ収集

本研究のデータは、日本の大学で開講されている「データマネジメント」科目から得られたものである。プログラミング初学者向けに設計されており、オンデマンド形式の動画教材が 7 本(各約 1 時間)提供され、受講者は一週間に内に任意のタイミングで視聴できる。表 2 に授業スケジュールの概要を示す。

表 2. 授業スケジュール概要

授業回数	授業形式	授業内容
01	対面	ガイダンス
02	オンデマンド	授業動画 1
03	オンデマンド	授業動画 2
04	オンデマンド	課題
05	オンデマンド	授業動画 3
06	オンデマンド	授業動画 4
07	オンデマンド	中間テスト
08	オンデマンド	中間テスト（予備）
09	オンデマンド	復習レポート
10	オンデマンド	授業動画 5
11	オンデマンド	授業動画 6
12	オンデマンド	課題
13	オンデマンド	授業動画 7
14	オンライン	期末試験
15	オンライン	期末試験（予備）

本研究で用いた動画ベースの学習管理システム (LMS) は、授業動画の任意のフレームをクリックするだけでノートや質問、コメントなどを投稿できる機能を備えている。これらの注釈は自動的にタイムラインに紐づけられ、学生が「このフレームで疑問に思った」「この箇所が重要だと思った」といった行動を即時に記録できる。また、演習課題や成績評価の一環として、注釈の投稿量・内容が「授業内の活動評価」に

組み込まれていたため、受講者は積極的に注釈を残す傾向があった。

最終的に、科目に登録していた学生 52 名のうち 42 名から有効な注釈データが得られ、合計 3,014 件の注釈が 7 本の動画に対して付与された。内容は基礎的な用語の説明メモから、高度な概念への質問や実装上の疑問まで多岐にわたる。なお、これらのテキストデータは匿名化を行い、研究目的での利用について事前に学生に伝えた。

### 4.2 前処理とキーワード抽出

本研究では、日本語形態素解析ツール GiNZA [10] を用い、注釈テキストに含まれるキーワードを抽出する。GiNZA は、Recruit の研究機関 Megagon Labs と国立国語研究所の共同研究で開発された学習モデルをベースとしており、spaCy のフレームワークと SudachiPy のトークナイザを組み合わせて日本語を解析する[20]。

しかし、標準的な GiNZA モデルでは IT 固有の専門用語やコード断片が数多く含まれる本研究のデータを十分に扱えない場合がある。そこで、著者らが独自に作成した IT 専門用語辞書 (27,381 語) を GiNZA に統合し、「情報科学総合教材用語リスト」[21]も併用することで IT 用語の認識精度を高めた。これにより形態素解析時に IT 用語が適切に 1 トークンとして扱われ、意図しない分割や読み誤りが大幅に削減できた。

前処理の具体的な手順は図 1a に示すように以下のとおりである。

1. トークン化：カスタマイズ済みの GiNZA で文を単語単位に分割する。
2. ストップワード除去：助詞・接続詞など意味を持たない単語を取り除く。
3. 標準化：同義表現を一つの基本形に統合し、「テーブル」「テーブル化」「テーブルにする」などのゆらぎを軽減する。

以上を経て得られたキーワードは、後続の分類処理に供する。

### 4.3 タキソノミ学習の動詞の拡張生成

単語埋め込みモデルの学習には、大規模でドメイン特化したコーパスを使用することが望ましい。本研究では、(1) カスタム IT 用語リストの各語に対応する Wikipedia 日本語版の解説文（計 15.9MB）と、(2) 授業動画のトランスクリプト（計 303KB）を組み合わせた独自コーパスを構築した。後者のテキストは ffmpeg で抽出した音声データを OpenAI Whisper で文字起こしたものである。

図 1b に示すように、これらのコーパスを用いて Word2vec モデルを学習する。Word2vec には主に

Continuous Bag-of-Words (CBOW) と Continuous Skip-gram の 2 種類があるが、本研究では周辺文脈が比較的広範に分散していると想定し、Skip-gram を採用した。学習時には単語ベクトルの次元数やウィンドウサイズなどのハイパーパラメータを調整し、プログラミング教育領域において有用な語彙類似度を得られるよう最適化を図った。

学習済みの単語ベクトルは、コサイン類似度によって語彙間の意味的近さを定量的に比較できる。たとえば「function」「return」のようなプログラミング特有の単語同士が高い類似度を示すだけでなく、「データベース」「正規化」「クエリ」などの専門用語間の文脈的な近さも把握可能となる。こうした特性により、Bloom's Digital Taxonomy に含まれない表現であっても、意味的に対応するカテゴリーを推定しやすくなる。

#### 4.4 キーワード分類

本研究では、Bloom's Digital Taxonomy に含まれる動詞（例：Remember カテゴリーであれば「列挙する」「識別する」など）を出発点とし、Word2vec モデルを組み合わせることで語彙を拡張している。タキソノミーの各カテゴリーには複数の行動動詞が定義されているが、プログラミング教育の実際の場面では学生がそのまま動詞を使用するとは限らない。

具体的には以下の手順でタキソノミーを拡張する。

1. 単語ベクトルの取得：Bloom's Digital Taxonomy にある動詞（表 1. 各カテゴリーには約 30 単語程度）について、学習済み Word2vec モデルから対応するベクトルを取得する。
2. 類似度計算：コーパス中のすべての単語ベクトルとのコサイン類似度を算出し、あらかじめ設定した閾値（本研究では 0.3）以上を抽出する。
3. カテゴリーへの追加：閾値を超えた単語を、対応動詞のカテゴリーに加える。たとえば「理解する」と類似度が高い単語であれば“Understand”に追加する。

このプロセスにより、最初は 180 語程度だったタキソノミーの語彙が、プログラミング文脈で学生が用いる表現や専門用語を含む大規模な分類テーブルへと拡張される。拡張後のタキソノミーを用いることで、学生が独特な表現や略語を使った場合でも、高い精度で認知レベルを推定できるようになる。

最後に、前処理で得られた注釈テキストのキーワードを拡張タキソノミーに照合し、以下のステップで分類を行う。

1. 直接マッチング：キーワードが拡張タキソノミーの単語と完全一致すれば該当カテゴリーに割り当てる。

2. レママッチング：一致しない場合、語形変化やゆらぎを吸収するために基本形や派生形を探索して再照合する。
3. 集計：注釈内に含まれるキーワード数をカテゴリー別にカウントする。注釈全体としては、キーワード出現頻度の合計を用いて、どの認知スキルに関連する言語表現が多いかを測定する。

こうして得られた分類結果をもとに、授業回単位や学生個人単位で認知スキルの分布を集計し、学習深度の可視化・評価を行う。すなわち授業が進むにつれて深い学びへ移行しているか、あるいは各学生がどのような認知パターンを示しているかを時系列で把握できるようになる。

## 5. 結果

### 5.1 前処理の結果

IT 専門用語辞書を GiNZA 形態素解析ツールに統合した結果、本研究のデータセットに含まれる IT 特有の用語の認識精度が大幅に向上した。実験では、最適化前の標準 GiNZA モデルと比較して IT 用語として正しく認識された単語数が 5,641 から 6,516 に増加し、約 15.57% の増加率が得られた。さらに IT 用語が全キーワードに占める割合も約 13.5% から 16.4% へ上昇した。一方で、最終的に抽出されたキーワードの総数は最適化前の 41,848 トークンから 39,819 トークンへと 4.85% 減少しており、不要な分割や誤判定が減少したこと示唆している。

これらの結果は、IT 用語辞書の組み込みによって解析精度が向上し、プログラミング教育のテキストデータ特有の専門用語を的確に捉えられるようになったことを意味する。とりわけコード断片やアルゴリズム名などが多用されるプログラミング学習の注釈データでは、形態素解析の正確性がその後のキーワード抽出や分類に直接影響を与えるため、この精度向上は本研究の分析手法において重要な役割を果たす。

### 5.2 タキソノミ学習の動詞の拡張結果

Word2vec モデルを用いて Bloom's Digital Taxonomy に含まれる動詞群を拡張したところ、カバー範囲と精度が大きく向上した。拡張前は各認知スキルに関連する動詞のみを含む合計 180 語ほどだったが、Word2vec による類似語追加により 4,383 語へと大幅に拡張され、プログラミング教育で頻用される専門用語や表現もカバーできるようになった。

実際の分類では、前処理段階で抽出された 39,819 のキーワードのうち 8,404 語（約 21.1%）が拡張タキソノミーテーブルのいずれかの認知スキルカテゴリーに割り当てられ、残る約 78.9% は未分類となった。これ

は日本語表現の多様さや、固有名詞・コメント的発話など、本分析の対象外語彙が含まれていることが理由と考えられる。なお、この分類率（20～30%程度）は著者らの別研究の結果とも概ね一致している。

タキソノミーを Word2vec でドメイン特化拡張したことで、学生が異なる言い回しや文体を使った場合でも、適切に認知スキルを推定しやすくなつた点は、本手法の有効性を示す根拠となる。

### 5.3 学習深度の定量的評価

図 2 に、各動画（レッスンごと）における認知スキル分類結果を示す。あわせて、図 3 には動画ごとの注釈数・抽出キーワード数・分類したキーワード数の推移を示した。

**低次認知スキル（Remember, Understand, Apply）の減少傾向。**学期が進むにつれて、いわゆる浅い学習レベル（Remember, Understand, Apply）の割合が徐々に低下している。

**高次認知スキル（Analyze, Evaluate, Create）の増加傾向。**一方で、Analyze, Evaluate, Create といった高次の認知活動を示すキーワードの割合は全体として増加傾向を示し、特に Create は最終回に向かって顕著に上昇した。ただし、第 5 回レッスンのみ Analyze が大きく落ち込み、Remember と Understand が増加する例外的パ

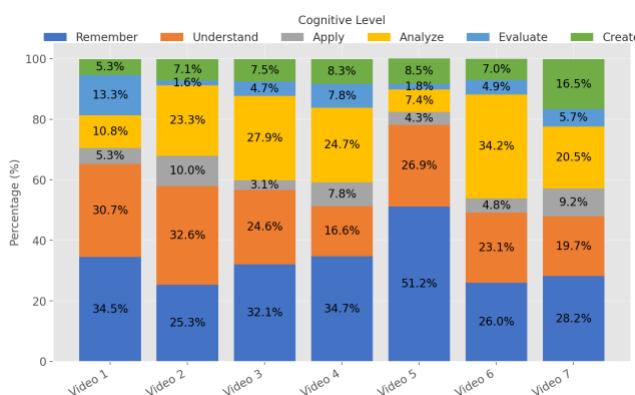


図 2. 各動画による認知スキル割合分布。

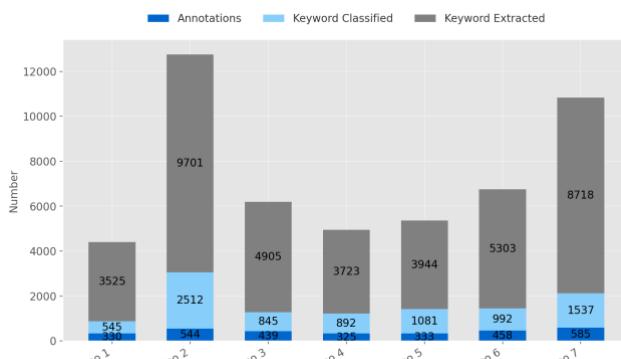


図 3. 各動画による注釈数・抽出キーワード数・分類したキーワード数。

ターンが見られる。これは授業内容が基礎的な復習を要する領域にシフトしたためと推察される。

図 3 からは、動画ごとの注釈数（300～600 件）に明確な増減トレンドは見られないものの、抽出キーワード数は大きく変動し、第 2 回で最も多い 9,701 語、第 1 回レッスンで最も少ない 3,723 語となっている。しかし、注釈 1 件あたりのキーワード数は概ね 10～18 前後で安定しており、注釈 1 件あたりの分類成功キーワード数も 1.65～4.62 の範囲で推移している。すなわち、レッスン内容や学生の書き込み方によって注釈数やキーワード数に差はあるものの、各注釈に含まれる情報量は一定レベルで保たれていると考えられる。

## 6. 考察

### 5.1 学習深度の変化

本研究の結果によれば、授業が進行するにつれて「Remember」「Understand」「Apply」といった低次認知スキルが徐々に減少し、「Analyze」「Evaluate」「Create」のような高次認知スキルが増加する傾向が確認された。これは学期を通じて、学習者が基礎知識の習得から、より高度な批判的思考力や創造的思考力へ移行していく可能性を示唆する。特に「Create」の割合が最終回で顕著に増加していた点は、授業後半でより複雑な課題に取り組むだけの下地が形成されたことを示すといえる。

一方、第 5 回レッスンのみ「Analyze」の割合が急減し、「Remember」や「Understand」の割合が再度高まる現象が観察された。これは授業内容が復習や基礎再確認に重きを置く構成にシフトしたこと、あるいは同時に他科目の課題や試験が集中したことによる認知負荷の増加が原因として考えられる。学習者が一時的に低次スキルへ回帰することは、高度な学習を継続するうえで必要な基礎知識を再度固めるプロセスでもあるため、必ずしも学習効率の低下を意味しない点に留意が必要である。

加えて、本研究ではクラス全体の注釈に対し、注釈 1 件あたりの抽出キーワード数や注釈 1 件あたりの分類成功キーワード数といった指標を導入し、学生の書き込み量や情報量を把握した。注釈 1 件あたりの抽出キーワード数が約 10～18 の範囲で比較的安定していることからは、ビデオ注釈という簡易な入力手段であっても、学生の思考や疑問が一定以上の情報量をもつて表現されていると推測できる。一方、Bloom's Digital Taxonomy と照合できたキーワードは 21.1% 程度にとどまり、残る 8 割近くは分類されない単語がある。

分類されない単語については、頻繁に登場する授業中の名詞をメモ的に書き留めるものが多いと考えられる。これらは学習者にとって重要なキーワードである

が、動詞を中心に認知スキルを捉える仕組みであることを踏まえると、こうした名詞の増加や単純な用語メモの記録が、必ずしも高次思考力や認知技能の変化を直接的に反映しているわけではないだろう。専門用語が多く注釈されていても、それだけで「Analyze」や「Evaluate」といった認知プロセスが十分に機能しているとは断定できないと考えられる。

## 5.2 教育への応用

こうした動的変化を注釈データからタイムリーに把握できる点は、本研究で提案した手法の大きな意義といえる。従来のテスト結果や提出物だけでは捉えきれない学習過程の揺らぎを可視化できるため、教員は学生がどの段階で理解に行き詰まっているか、あるいはどの領域でより高度な思考に進んでいるかを早期に見極めやすくなると考えられる。また、授業進行に伴う学習深度の推移を定量的に評価できるようになることで、授業設計の狙いや学習到達度との整合性を検証しやすくなる。たとえば、高次スキルの育成を掲げているにもかかわらず、注釈の大半が「Understand」にとどまっている場合、教授法や学習活動を見直す契機と捉えられる。さらに、個別学習者の注釈を時系列で追えば、どの時点で深い学びに進めているか、あるいは理解が停滞しているかを早期に捉えられるため、より効果的なフィードバックや支援が実現しやすくなるだろう。

オンライン教育や大規模クラスの文脈でも、ビデオ注釈のような簡易な操作で得られるテキストデータを活用すれば、多様な学習者の高次思考の状況を可視化・評価できる可能性がある。本研究の手法を他分野の科目や異なる学習環境に展開していくことで、さまざまな授業形態における学習深度の把握や教員の指導設計への還元をさらに促進できると期待される。

## 7. おわりに

本研究では、Bloom's Digital Taxonomy と自然言語処理技術を組み合わせ、オンラインプログラミングコースにおける学生の注釈データから学習深度を定量的に測定する枠組みを示した。具体的には、日本語形態素解析ツールの最適化や専門用語辞書の導入、ドメイン特化型 Word2vec モデルの学習を通じて、学生が残すノート・質問・コメントのキーワードを認知スキルに分類する手法を構築している。その結果、授業の進行に伴い学生の認知活動が下位レベルから上位レベルへ移行する傾向が明確に捉えられ、指導内容や教授方略が学生の認知スキルに与える影響を示唆した。

従来のテスト成績や提出物中心の評価では把握しづらい学生の思考過程を、より詳細かつ動的に追跡で

きる点は本手法の大きな特徴である。学習深度を可視化することで、教育者は授業設計の有効性を検証し、個々の学生への指導やサポートを的確に行うことが期待できる。特に、学期中のどの時点で学習深度が停滞・後退しているかを把握すれば、適時のフィードバックや支援策を講じることが可能となる。

今後の課題としては、タキソノミー拡張のさらなる精緻化や、リアルタイム解析を学習管理システムに組み込む実装面の検証が挙げられる。また、プログラミング以外の分野や大規模クラスへの適用可能性についても検討が必要である。多様な教育現場で本手法を応用し、学習者の認知的発達をデータに基づいて把握・支援する枠組みとして発展させることが期待される。

## 参考文献

- [1] 文部科学省「個別最適な学び」と「協働的な学び」の一體的な充実」.  
[https://www.mext.go.jp/a\\_menu/shotou/new-cs/senseiouen/mext\\_01317.html](https://www.mext.go.jp/a_menu/shotou/new-cs/senseiouen/mext_01317.html). (2025年3月18日確認)
- [2] X. Wang, H. Kiyomitsu, Y. Su, K. Ohtsuki, Y. Sun, and K. Shigyo, "Visualization of potential differences in comprehension by distribution of notes and questions in online programming courses," in *2023 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*, IEEE, pp. 1-8, 2023.
- [3] X. Wang, Y. Su, Y. Sun, T. Nishida, K. Ohtsuki, and H. Kiyomitsu, "Voices of Asynchronous Learning Students: Revealing Learning Characteristics Through Vocabulary Analysis of Notes Tagged in Videos," in *Database Systems for Advanced Applications (DASFAA 2024)*, pp. 277-292, 2024.
- [4] B.S. Bloom, "Taxonomy of Educational Objectives," Pearson Education. 1956.
- [5] S. Honda, H. Konda, N. Miho, Y. Yamada, T. Mori, and S. Mizokami, "The Influence of Perception about Relationship with Others in Class and Externalization in Active Learning," *Journal of Japan Association for College and University Education*, vol. 41, no. 1, pp. 88-96, 2019.
- [6] L.W. Anderson, D.R. Krathwohl, et al.: "A Taxonomy for Learning, Teaching, and Assessing: a Revision of Bloom's Taxonomy of Educational Objectives", Longman, 2000.
- [7] A. Churches, "Bloom's digital taxonomy," ed: Australian School Library Association NSW Incorporated, 2010.
- [8] T. Mikolov, S. Ilya, K. Chen, G. Corrado, and J. Dean: "Distributed Representations of Words and Phrases and their Compositionality", *Advances in Neural Information Processing Systems* 26, pp.3111-3119, Curran Associates, Inc, 2013.
- [9] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532-1543, 2014.
- [10] GiNZA, Japanese NLP Library.  
<https://megagonlabs.github.io/ginza/> (2025年3月18日確認)