

1. はじめに

デジタル教科書の操作を記録した学習行動ログデータから、講義を十分に理解できていない学生を早期発見し、学習に対する改善等を提供する試みが行われている [1]。しかし学習行動ログデータは、学生の講義に対するモチベーションや、内容についての解釈といった抽象的な情報を考慮できない。そのため、具体的な学習支援を提供することが難しい。本研究では、学生の主観が含まれる自由記述形式の学習日誌に注目する。予備調査より、講義に対する日誌の変化と成績の相関が高いことが判明した。そこで、学習日誌の変化量を用いた成績予測モデルを構築し、成績予測時の分析結果を用いた LLM 推論によって、より学生にとってわかりやすく、具体的な根拠と改善策を提供可能な学生向けフィードバックシステムの実現を図る。

2. 研究背景

教育分野において学生の成績の要因を調査する研究が発見である。Stephanie ら [2] は、工学系の大学生に対して自由記述アンケートを実施し、回答内容と GPA に相関があるか調査した。分析の結果、成績の高い学生はエンジニアの実務内容や目的を表現する単語や、数学、科学に対するポジティブな発言が見られ、成績の低い学生が用いる単語と明確な差が確認された。これらから、学生が記録した学習日誌を分析することで、将来の成績を予測するとともに、講義へのモチベーション低下など、成績変動に関与する要因を根拠として具体化できると考えられる。

3. 予備調査

本研究で用いる学習日誌データセットに対して分析を行い、成績との関連性を調査する。本調査では、九州大学で収集した、情報科目の講義における学習日誌と、受講者の成績で構成されるデータセットを用いる。学習日誌は講義直後に、①今回の講義内容の説明、②講義内容や講義についてわかったこと、③講義内容についてわからなかったこと、④講義への質問、⑤講義の感想の 5 項目に回答する形式で行われた。学習日誌は 2020 年から 2022 年までの期間、合計 377 名の学生に対して実施された。各受講者の成績は A, B, C, D, F の 5 段階評価である。

本データセットは講義回ごとに同一の学生から学習日誌を収集しているため、学生 s の講義 i 回目の回答 $x_{s,i}$ は $x_{s,1}, x_{s,2}, \dots, x_{s,15}$ のような時系列データとして表現できる。このとき、各回答は講義順に時系列が進み、講義 i 回目の回答が講義 $i+1$ 回目の回答に影響しうするため、連続性や変化（改善・停滞・悪化）といった時系列的な情報が包含される。この時系列情報を埋め込み間距離を用いて定量化することで、各学生の継続的な講義への取り組み方を調査する。講義間における学習日誌の埋め込み間距離 Dist の算出方法を式 (1) に示す。

$$\text{Dist}(i, g) = \text{median} \left(\sum_{n \in U_g} 1 - \frac{\mathbf{A}_i^\top \mathbf{A}_{i+1}}{\|\mathbf{A}_i\|_2 \|\mathbf{A}_{i+1}\|_2} \right) \quad (1)$$

ここで、 \mathbf{A}_i は Word2Vec[3] によって得られた講義 i 回目の学生回答の埋め込み表現、 U_g は成績 g の学生群である。式 (1) で示すように、隣接する講義回の記事埋め込みに対してコサイン距離を求めることで、各講義間の意味的な変化量を抽出する。また、コサイン距離は学生ごとに変化が大きいため、全学生から算出したコサイン距離の中央値を利用する。

予備実験の結果を図 1 に示す。図 1 より、成績が高い学生ほど講義間コサイン距離が大きく、講義 14 から 15 回目における学習日誌の回答内容が大きく変化していることが確認できる。この結果から、学習日誌の変化量は、成績予測に有用な指標であるといえる。

4. 提案手法

予備実験より、学習日誌における内容の変化量が成績に大きく影響していることが示された。本研究ではこの結果

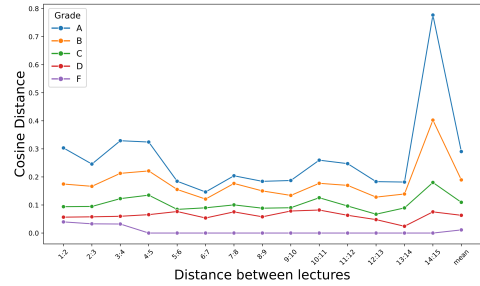


図 1: 講義間のコサイン距離

を踏まえ、以下の 2 つから構成されるアプローチによって、より効果的なフィードバックを提供可能な学習支援を実現する。

1. 学習日誌の取り組み差分情報を用いた成績予測
2. LLM による学生向けフィードバックの生成

4.1. 学習日誌の取り組み差分情報を用いた成績予測

各学生の全 15 回分の学習日誌を文章埋め込みに変換し、予備実験と同様に変化量を算出することで、成績予測モデルの精度向上を図る。成績予測モデルには、予測結果の判断根拠を容易に算出可能である点と系列データへの解釈性の高さから、Light Gradient Boosting Machine (LightGBM) を利用する。予測には、予備実験で用いた講義間コサイン距離だけでなく、講義間ユークリッド距離、学習日誌の各講義回における回答文字数、講義全体の欠席回数の特徴量として利用する。講義間コサイン距離はベクトル間の類似度から文脈や意味的な変化量を算出し、講義間ユークリッド距離はベクトル間の直接的な距離関係から使用単語の変化や文章構成の変化を定量化するものである。

4.2. LLM による学生向けフィードバックの生成

LightGBM モデルによって得られた成績予測結果とモデルの分析を利用することで、より説得力のある、学生向けの学習方法フィードバックを生成する。図 2 に本手法の概要を示す。本手法は LLM による 2 段階推論によって生成文の品質向上を目指す。第 1 段階では、モデルの SHAP 分析結果から寄与度が高い講義回を抽出し、寄与度が高くなった根拠を講義内容と学習日誌の回答から予測する。このとき、寄与度の高い要素が複数ある場合に備え、出力を構造化することで網羅性の欠落を防止する。第 2 段階では、第 1 段階で得られた分析結果と講義資料を入力することで、根拠と改善策を分かりやすく提示した学生向けのフィードバック文を生成する。各 LLM は GPT-4o の生成文を真値として訓練する。

5. 評価実験

本章では、提案手法の有効性を検証するための評価実験を行う。

5.1. 実験条件

学習日誌を用いた際の成績予測精度を比較し、提案手法の有効性を検証する。ベースラインとして、文章埋め込みを直接入力する LightGBM モデルを用いる。各講義回ごとにモデルを訓練し、その出力の平均値を最終的な予測確率とする。評価には分割交差検証法 ($k=5$) を行い、Accuracy, F1-score を用いて提案手法の有無における精度の変化を検証する。

5.2. 実験結果

各手法における講義理解度予測精度の比較を表 1 に示す。表 1 より、ベースラインよりも Accuracy が 6.43pt, F1-score が 8.48pt 向上し、提案手法の有効性を確認した。次に、定性評価として SHapley Additive exPlanations (SHAP) による特徴量寄与度を可視化し、モデルの判断根拠を分析

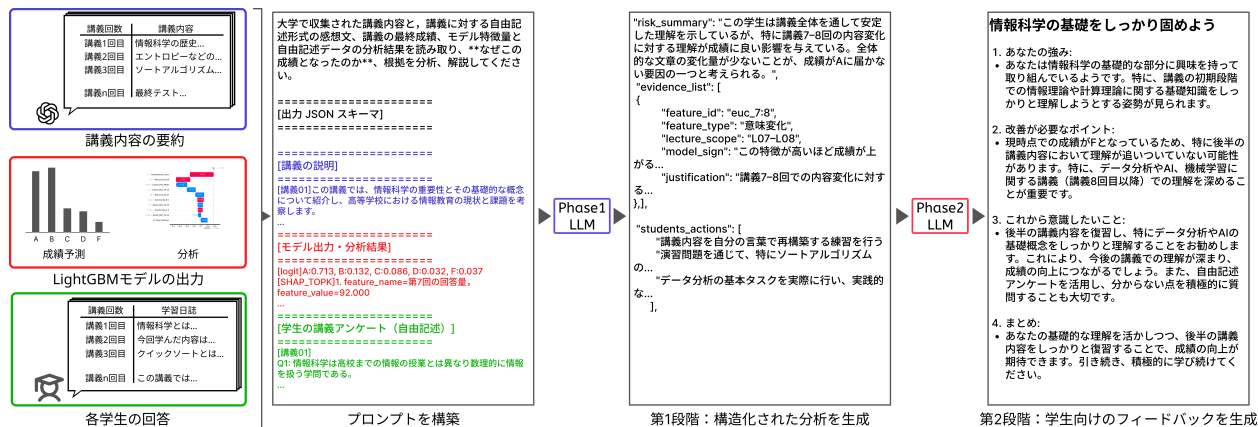


図 2: LLM によるフィードバックシステムの概要

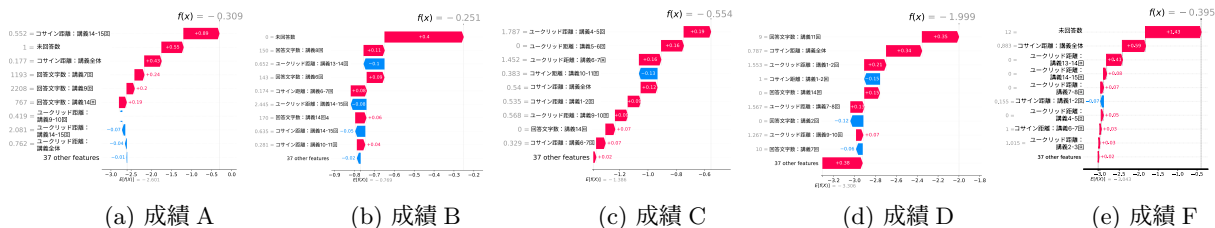


図 3: 各成績の学生における判断根拠の可視化

表 1: 成績予測精度の比較

指標	ベースライン	提案手法
Accuracy	45.70	52.13
F1-score	33.20	41.68

する。図 3 に各成績の学生に対する寄与度分析結果を示す。図 3 より、成績 A, B の学生に対してモデルは講義 14, 15 回間のコサイン距離や未回答数、全体的なコサイン距離の平均に注目していることがわかる。この結果から、成績が高い学生ほど、講義最終時点の回答内容の変化や、講義全体を通して一貫性のある回答がされているかといった要素に注目していることがわかる。一方で、成績 C, D の学生に対して、モデルは講義 4~7 回目のユークリッド距離や、講義 11 回時点の回答など、中盤~後半の学習日誌に注目している。この結果から、成績が低下する傾向にある学生は成績が高い学生と比較し、講義中間以降の取り組みに問題があると解釈できる。成績 F の学生については、学習日誌の未提出回数と全体のコサイン距離に注目していることから、学生の講義不参加といった傾向に強く注目した予測を行っていることが確認できる。以上より、成績予測モデルは各成績の学生が示す学習日誌の変化から、講義全体における取り組みの継続性に注目して予測を行なっていると考えられる。

5.3. LLM による学生向けフィードバックの評価

LLM によって生成されるフィードバック内容を定性的に評価する。図 4 に、成績 C の学生に対するフィードバックの生成結果を示す。図 4 より、予測時の判断根拠をもとに、初期の講義内容である符号化やエントロピーなどの単元に対する取り組みが活発であるといった強みの分析と、後半の講義内容であるデータ分析や AI の単元における理解不足といった改善点を明確に提示している。この傾向は図 3 に示す成績 C の学生の寄与度分析と一致しており、モデルの分析内容をフィードバックに適用できていることを示唆している。さらに、「非構造化データの処理方法について理解を深めるべきである」といった具体的な改善点が言及されていることから、各学生に対して講義理解を促す具体的なフィードバックを提示できていることが示唆される。

情報科学の基礎をしっかりと固めよう

- あなたの強み:
 - あなたは情報科学の基礎的な部分において、特に初期の講義での理解がしっかりしています。情報の符号化やエントロピーの概念に関する基礎知識がしっかりしていることが、講義の進行に役立っています。
- 改善が必要なポイント:
 - 講義の後半、特にアルゴリズムやデータ分析、AI の内容に関する理解がやや不足している可能性があります。これらの分野の具体的な手法や計算量の違いを復習することで、より高い成績を目標せましょう。
 - これから意識したいこと:
 - アルゴリズムやデータ分析、AI の応用について、実践的な演習を通じて知識を定着させましょう。また、後半の講義で扱う非構造化データの処理方法やパターン認識の基本概念についても、理解を深めることをお勧めします。
- まとめ:
 - あなたの基礎的な理解を活かしつつ、特に後半の講義内容を補強することで、さらに高い成績を目指すでしょう。引き続き頑張ってください。

図 4: LLM によるフィードバック結果

6. おわりに

本研究では、学生の理解度や解釈を読み取りやすい自由記述形式の学習日誌に注目し、学生の成績予測を活用するフィードバックシステムを開発した。各講義における学習日誌の変化が成績と強く関連していることが示されたため、これを決定木モデルの入力に利用することで、Accuracy が 6.43pt, F1-score が 8.48pt 向上することを確認した。さらに、成績予測モデルの SHAP 分析において、講義中間時点における学習日誌の内容変化に対して寄与度が高いと示された。この結果を用いた多段階 LLM の構築により、学生に対して具体的なフィードバックを提示可能となった。今後は、フィードバックで示す成績改善案のさらなる具体化と、実環境における有効性の調査を行う予定である。

参考文献

- Yagci, "Educational data mining: prediction of students' academic performance using machine learning algorithms", Smart Learn. Environ., 2022.
- Gratiano *et al.*, "Can a five minute, three question survey foretell first-year engineering student performance and retention?", ASE, 2016.
- Mikolov *et al.*, "Distributed Representations of Words and Phrases and their Compositionality", NeurIPS, 2013.

研究業績

- Koike *et al.*, "An Investigation of the Relationship between Open-Ended Questionnaires and Lectures", LAK, 2025.
(他 学会発表 2 件)