

学習者のレポートテキストデータを対象とした ベイズ推定による理解度推定方式

青木 慎太郎[†] 浅井 悠一朗[†] 中西 崇文[†]

武蔵野大学 データサイエンス学部 データサイエンス学科[†]

1. はじめに

近年，記述式問題やレポートに対する自動採点の研究が盛んに行われている．学習者の取り組みが点数化されることで到達度を知ることができる．一方で現状の自動採点システムからは学習者の習熟理解度を測ることはできないと考えた．記述式問題やレポートでは論理的思考力・判断力・表現力を養うものであるため，教師の添削やフィードバックが重要である．しかしながら，教師側の都合や受講者数の多さから採点やフィードバックが遅れることや，点数のみの返却の場合がある．提出されたレポートに対して習熟理解度をフィードバックできれば今後の学びに活かせると思う．

本稿は，学習者のレポートテキストデータを対象としたベイズ推定による理解度推定方式について示す．本方式は，学習者が提出したレポートの冗長さや語彙の文章校正の観点，正解レポートとの内容類似性を指標として，ベイズ推定による学習者の理解度を推定する．本方式を実現することで，学習者の提出されたレポートに対して指標ごとにどの程度理解しているのかをフィードバックすることができ，今後の学びに生かすことが可能となる．

2. 関連研究

このような研究には，大まかに自動採点のシステムの構築，成績の推定といった位置付けである．それぞれの研究は，特定のタスクを解決するために機械学習アルゴリズムを用いてモデルを構築している．

加藤ら[1]は，記述式テストを対象に BayesianNN を適応し深層学習を用いた予測モデルを作成し自動採点を実現している．

東ら[2]は，アンケート結果，小テストの点数，及び定期試験の点数を対象に Random Forest

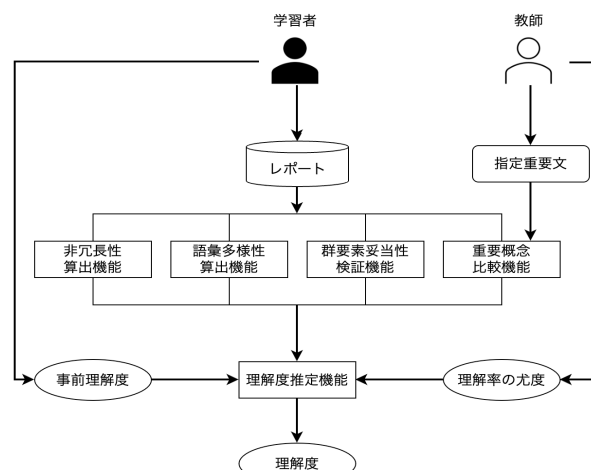


図1：本方式の全体図

を用いて機械学習アルゴリズムを構築し成績の推定を実現している．

本方式は，学習者のレポートテキストデータから文章の冗長さや語彙の豊富さの文章校正の観点と教師が指定した正解レポートから教師が指定した文章との内容類似性を指標として，学習モデルや事前データを必要としないベイズの定理を用いて理解度を推定する．

3. 学習者のレポートテキストデータを対象としたベイズ推定による理解度推定方式

図1に，本方式の全体図を示す．本方式は，指標を算出する非冗長性算出機能，語彙多様性算出機能，群要素妥当性検証機能，重要概念比較機能と最終的な推定を行う理解度推定機能から構成される．データ前処理は，事前に学習者に見立てたテーマ及びレポートデータの作成，非冗長性算出機能，語彙多様性算出機能構築のための形態素解析ライブラリ MeCab を用いた形態素解析処理から構成される．

3.1 非冗長算出機能

非冗長性算出機能は，冗長な要素を検出し，その要素の出現率を算出するものである．テキストデータから文章中における重複部分や不要単語の出現回数を数え上げ，単語の出現確率を求め，単語集合の特性であるエントロピーを求

Bayesian Estimation of Learner's Report Text Data
Comprehension Estimation Method.
Shintaro Aoki[†], Yuichiro Asai[†], Takafumi Nakanishi[†]
[†] Musashino University, Department of Data Science

める．単語エントロピー H は $H = -\sum_k P_i \log_2 P_i$ で定義される．算出した単語エントロピーの最大値 H_{max} を求め冗長度 R を算出する．

冗長度 R は以下の式で算出される．

$$R = 1 - \frac{H}{H_{max}}$$

3.2 語彙多様性算出機能

語彙多様性算出機能とは，提出されたレポートからユールの K 特性値を用いて語彙の多様さを算出する機能である．ユールの K 特性値とは語彙の豊富さを表現するための公式であり，この機能によりレポートの語彙がどれだけ豊富かを算出することが可能になる．

3.3 群要素妥当性検証機能

群要素妥当性検証機能とは，提出されたレポートに対して群としての所属位置を算出する機能である．この機能では，doc2vec を使い文章を特徴量に変換し，UMAP を用いて得られた特徴量を2次元に次元削減し，その上で特徴量のクラスタリングを行う．その後ユークリッド距離による学習者の所属位置と教師の正解クラスターとの妥当性の比較を行い学習者のレポートが正解クラスターに属するかを判定する．

3.4 重要概念比較機能

重要概念比較機能とは，教師側が指定する正解の文章と，学習者が提出したレポートの各文章の類似度を計算し，その結果から学習者が正しい論述を行っているか否かを比較する機能である．Sentence-BERT を用いることで，意味的な文の埋め込み表現を獲得し，それらのコサイン類似度により算出している．

3.5 理解度推定機能

算出指標を基にベイズの定理を用いて最終的な理解度の算出を行う．始めに学習者が主観的な事前理解度を入力し，各指標に対し教師側が仮定する理解率を尤度として設定する．入力された事前理解度に対して各指標を用いて確率の更新を行い，その結果が学習者の理解度として推定される．

4. 実装・実験

本節では，本方式を実装したシステムについて示す．テーマを『円安と円高の違い』と設定し各機能の特徴付けた疑似レポートを作成した．事前確率と指標算出結果を，理解度推定機能を用いて理解度を推定した結果を表1に示す．

表1の結果から，各機能により疑似レポートの理解度が更新されていることがわかる．また，各機能によって算出された結果がレポートの質に合わせて変化していることがわかる．

評価実験として提案した非冗長性算出機能，

表1：理解度推定結果

	事前確率	重要概念比較機能	群要素妥当性検証機能	語彙多様性算出機能	非冗長算出機能
1	40.0%	80.0%	84.5%	81.3%	53.0%
2	50.0%	36.6%	28.3%	50.6%	32.6%
3	80.0%	96.0%	71.7%	91.9%	88.8%
4	50.0%	69.0%	49.0%	42.4%	31.8%

語彙多様性算出機能，群要素妥当性検証機能，重要概念比較機能が学習者のレポートから理解度を求める上で有効であるかを判断する．機能の有効性を大学生27名を対象に各指標の値が高いレポートと低いレポートを被験者に見せた上で，その指標が採点者にとって回答者の理解度を確認する上で必要か否かをリッカート尺度5段階評価のアンケート方式で検証を行った．その他に大学生24名を対象に講義での課題についてのアンケートを実施した．

各機能及び指標の有効性のアンケート結果は，Q1 非冗長性算出機能の有効性のスコアは($\mu=4.40$, $\sigma=0.56$)となった．Q2 語彙多様性算出機能の有効性のスコアは($\mu=4.03$, $\sigma=0.88$)．Q3 群要素妥当性検証機能の有効性のスコアは($\mu=3.81$, $\sigma=0.72$)，Q4 重要概念比較機能の有効性のスコアは($\mu=4.33$, $\sigma=0.54$)となった．

講義での課題についてのアンケート結果は，課題は87%が記述式のレポート形式の課題であることがわかった．また67%が返却されるのが遅いと回答し，回答者の9割がどの程度理解できているかなどのフィードバックが必要だと回答した．

5. おわりに

本稿では，学習者のレポートテキストデータを対象としたベイズ推定による理解度推定方式について示した．本方式により，学習者の提出されたレポートに対して提案した機能ごとにどの程度理解しているのかをフィードバックすることができ，今後の学びに生かすことが可能となる．

今後の課題として，学習者により明確に何を復習すべきなのかをフィードバックするシステムの実現，語彙の水準を求め，より正確な学習者のレポートテーマにおける語彙力を算出するなどが今後の課題として挙げられる．

参考文献

- [1] 加藤 嘉浩，ベイジアンニューラルネットワークを用いた記述式答案の自動採点，言語処理学会，第26回年次大会発表論文集，pp.969-972，(2020)
- [2] 東 るみ子，機械学習による学習者の理解度推定，第43回教育システム情報処理学会全国大会，pp47-48，(2018)