

記述式解答群に対する観点を指定した概要把握 —クラスタリングによる記述の分類—

平澤賢人*1・高瀬治彦*1
Email:425M237 @mic-u.ac.jp

*1: 三重大学大学院工学研究科電気電子工学専攻

◎Key Words

演習支援, 解答要約, クラスタリング

1. はじめに

教育現場では、講師が学習者に対して質の高い講義を行うことが期待される。そのため、講師は講義における学習者の理解状況に応じて、フィードバックを行う必要がある。^(1,2) このため、講師は小テストや課題などの問題演習を活用し、学習者の理解状況の把握を行っている。しかし、講義の形式において多くを占めるのは多人数講義であり、多人数講義では、講義内に学習者の理解状況を把握するのは時間的にも労力的にも困難である。

講義の問題演習には種類があり、穴埋め式や選択式、記述式などの種類がある。これらの演習のうち、学習者の理解状況をより詳細に把握することができるのは記述式の演習であり、講師が学習者全体の理解状況を把握するのに適している。しかし、記述式解答の演習では、学習者の解答が多様であるため、要約等の支援がなされてきた。^(3,4) 文献(3)の手法では、解答を計算機によって採点を行い、支援している。しかし、結果が点数の形で要約されるため、フィードバックに必要な情報が欠落してしまう。また、文献(4)の手法では、解答群から自動抽出された重要語の前後を分類し提供している。しかし、講師が解答に対して着目すべきポイントは複数存在するため、それを考慮した要約が必要とされている。

我々の研究グループでは、クエリという、講師が解答に対して着目したいポイントを提示することで、クエリに沿った学習者の解答の概要を講師が素早く把握できるよう支援する方法について検討してきた。⁽⁵⁾ 本稿では、この手法において、解答群のクラスタリング手法とその結果の可視化の手法について検討を行う。

2. 記述式解答群における解答概要把握支援システム

本節では、従来研究である二村らの解答の概要把握支援システムについて説明を行う。

従来法である二村らのシステムは、解答中に複数存在する着目したいポイントを「クエリ」(質問形式の文章)により指定することで、クエリに関連する記述を学習者の解答から抽出し提供するものである。短文記述式解答の問題の解答に対して講師が自分の知りたい情報に関連するクエリを投げかけることで、クエリに沿った内容を抽出し、クラスタリングして提供する。⁽⁶⁾ 具体的な例を使用して説明する。まず、講義で「コンピュータの5大装置それぞれの働きを100~200文字で説明しなさい」という問題に学習者が解答を行う。そして、講師は5大装置の中から演算装置の働きに対して着目したとする。

その上で、講師はクエリを「演算装置は何を行いますか?」と指定する。システムは、各解答から演算装置の働きに言及している内容を抽出し、その結果をクラスタリングする。最後に、講師はこの結果を参考に演算装置に関する学習者の理解状況を把握する。

文献(5)では、クラスタリング以降の部分について十分検討されておらず、階層型クラスタリングしてその結果であるデンドログラムを提案した例が示されているのみである。そこで、本稿では階層型クラスタリング以外のクラスタリング手法、デンドログラム以外の可視化手法について検討する。

3. クラスタリングと結果の可視化

2節をふまえて、クラスタリング手法とクラスタリング結果の可視化手法について検討を行う。

まず、クラスタリング手法について検討する。文章分類に適しているクラスタリング手法の中から、階層的クラスタリング、K-means、HDBSCANを比較する。階層的クラスタリングは、事前にクラスタ数を決める必要がなく、クラスタの細分化や凝集する可能性に対して柔軟に対応できる。また、得られたクラスタ構造をデンドログラムとして可視化することで、クラスタ間の関係性や階層構造を直感的に把握できる。そのため、学習者の記述解答を網羅的に捉え、個々の解答がどのようなグループに属しているかを一目で確認できる。また、K-meansは計算量が少なく、大規模データに対応することができるだけでなく、クラスタの中心が明確に定義されるため、各クラスタがどのような特徴を持つ解答群で構成されているかを講師が理解しやすいという点も、学習者の理解度を把握する上で有用である。また、HDBSCANは密度ベースのクラスタリング手法であり、データの分布から自動的にクラスタ数を決定するため、事前知識が少なくても柔軟に対応できる。さらに、ノイズとなるデータポイントを明確に識別できるため、講師がクエリと関連性の低い解答をノイズとして排除し、より重要な解答群に焦点を当てることを可能にする。

クラスタリング結果の可視化手法についても検討する。可視化手法では、まとめ表、ワードクラウド、樹形図を比較する。まとめ表は図1のように、学習者の解答からクエリに関する部分を抽出した記述の中から代表的な記述を右側に、代表記述グループの有効解答数での比率を左側にまとめたものである。講師は学習者の主要な解答を必要とするため、その情報をわかりやすく表示しているまとめ表は有用である。ワードクラウドでは、

図2のように代表記述と比率を表示しつつ、解答数の多さによってクラスタの代表記述が大きくなり、大きさによって記述の色を変更できるため、感覚的に解答を把握できるようになる。樹形図では図3のように、縦軸をクラスタ間の距離、横軸を抽出された記述で示す。樹形図は、利用者の裁量でクラスタを分けることができるため、講師が講義内容の特定の側面や、生徒の理解度における微妙なニュアンスに合わせて、クラスタの粒度を調整できるという点で特に有用である。

これらの手法は、学習者の解答概要を把握する上で、それぞれ異なる利点がある。そのため、どの手法が適しているかは定性的な議論のみでは定まらず、具体的なデータを用いて比較検討することが不可欠である。次節では複数のクラスタリング手法と可視化手法を実際に適用し、その結果を比較することで、学習者の解答の概要を講師が素早く把握することを支援するシステムに最も適した手法を調査する。

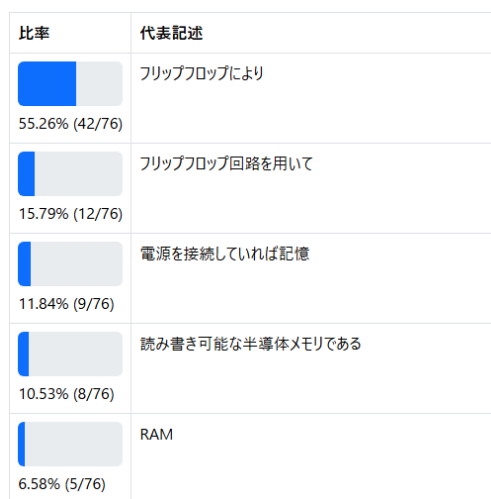


図1 まとめ表

フリップフロップにより (42/76-55%)

フリップフロップ回路を用いて (12/76-16%)

電源を接続していれば記憶 (9/76-12%)

読み書き可能な半導体メモリである (8/76-11%)

RAM (5/76-7%)

図2 ワードクラウド

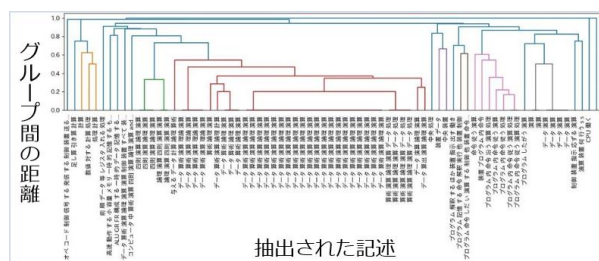


図3 樹形図 (デンドログラム)

4. 実験

本節では、クラスタリング手法とクラスタリング結果の可視化手法を簡単な実験を通じて比較する。

4.1 実験方法と条件

本実験では、クラスタリング手法については、クラス

タの内容の妥当性を評価し、可視化手法については表示された結果を被験者に評価してもらった。評価のために大学生7人を被験者として採用した。いずれの被験者も実験に用いた演習を行った講義を受講した経験がある。

クラスタリング比較では、講義の計算機工学で使用された問題から2つ用意し、それぞれに2つの異なるクエリを投げかけた計4セットの抽出結果に対して最大のクラスタを対象とする。具体的な問題とクエリは以下のとおりである。

問1「マイクロプログラム制御方式における、マクロ命令・マイクロ命令の設計方式の違いを説明しなさい。」

クエリ1「何がマクロ命令を格納する？」

クエリ2「マクロ命令は何によって実行される？」

問2「BIOS, IPL, OSの関係について説明しなさい。」

クエリ1「BIOSは最初に何を読み込む？」

クエリ2「BIOSを実行した後に何を行う？」

いずれの抽出結果から、被験者に最大のクラスタ抽出してもらい、8割以上の被験者によって抽出された内容を最大のクラスタの正解として用意した。これとK-means, 階層的クラスタリング, HDBSCANでクラスタリングして得られた最大のクラスタとの類似率を比較する。また、クラスタリング手法において、K-meansではシルエットスコアを用いたクラスタ数決定法を、階層的クラスタリングでは結合型でウォード法を使用している。いずれの手法も、分類用のデータとして、クエリに従い抽出された記述全体を一つのベクトルで表現する手法(BAAI/bge-m3)⁶⁾によりベクトル化したものを与えた。

可視化比較では、クラスタリング手法を階層的クラスタリングに固定し、クラスタリング結果をまとめ表、ワードクラウド、樹形図でそれぞれ表示したものを見てもらいアンケートを実施した。学習者の解答の概要を把握でき、視認性が十全である手法を10点満点とし、それぞれの点数とその理由を記入してもらった。

4.2 実験結果

まず、クラスタリング手法による最大クラスタを比較する。正解と3つのクラスタリング手法によるクラスタとの類似率をJaccard係数を用いて評価した結果を次に示す。Jaccard係数は、2つの集合の一致度を示す指標の一つであり、和集合全体における共通部分の割合を評価する手法で、1が最高値である。

表1 問1クエリ1のクラスタリング結果

手法/係数	Jaccard
K-means	0.98
Hierarchical	0.98
HDBSCAN	0.075

表2 問1クエリ2のクラスタリング結果

手法/係数	Jaccard
K-means	0.81
Hierarchical	0.061
HDBSCAN	0.84

表3 問2 クエリ1のクラスタリング結果

手法/係数	Jaccard
K-means	0.95
Hierarchical	0.97
HDBSCAN	0.81

表4 問2 クエリ2のクラスタリング結果

手法/係数	Jaccard
K-means	0.89
Hierarchical	0.89
HDBSCAN	0.53

表5 クラスタリング結果の平均値

手法/係数	Jaccard
K-means	0.91
Hierarchical	0.72
HDBSCAN	0.56

また、可視化手法の見やすさの平均点数を次に示す。

表6 可視化手法の平均点数

可視化手法	点数
まとめ表	9
ワードクラウド	5
樹形図	4

アンケート結果では、まとめ表の点数が高く、それ以外が同列に低い点数となった。それぞれの評価理由として、まとめ表では、「記述が頻度の多い順に並んでいて、比率のメーターも分かりやすい」、「解答の見やすさ、比率が直感的に判断しやすい」などの意見が挙げられた。ワードクラウドの評価理由としては、「文字の大きさによって頻度の多い順番が分かるが、頻度の少ない解答が見にくかった」、「似たような記述が違う大きさ、違う色で並べられていても分かりにくい」などの意見があった。また、樹形図の評価理由としては、「どういう図か理解するまでが難しい」、「グループ分けされているのはわかるが、見にくい。どの記述が多いか少ないのかもわかりづらい」などの意見となった。

5. 考察

表1～5より、人手クラスタとそれぞれのクラスタリング手法のクラスタの類似率は、全体的に K-means が高い類似率を示しており、K-means が記述クラスタリングにおいて有効であることがわかる。また、K-means は階層的クラスタリングや HDBSCAN よりも計算量が少ない。そのため、講義中に素早く解答概要を把握できると考えられる。この点からも、K-means が有効であるといえる。

階層的クラスタリングは、おおむね K-means と同程度の類似率を示しているが、類似率が極めて低い場合があった。これは、クラスタリング結果で示された最大のクラスタと、正解の最大のクラスタとの間に食い違いがあったためである。階層的クラスタリングの結果の解釈

は、デンドログラムの解釈のしかたに影響を受けるため、自動で解釈すると人の解釈と食い違ってしまったといえる。

また、HDBSCAN は、安定して他の手法より低い類似率となった。これは HDBSCAN が密度の高い記述同士でクラスタを形成するため、記述式解答のようなゆらぎの大きい文書群では、ゆらぎがノイズみなされてクラスタとしてまとめることに失敗したといえる。

以上の定性的な分析をふまえると、記述式解答から抽出した記述を分類し、その主要な内容を素早く把握するという用途では、K-means が適しているといえる。

可視化手法の結果について、表6より、まとめ表が適していると考えられる。これは、まとめ表が「記述が頻度の多い順に並んでいて、比率のメーターも分かりやすい」、「解答の見やすさ、比率が直感的に判断しやすい」などの評価理由から、解答概要がシンプルにまとまっており、視認性が良かったと考えられる。

以上の結果から、クラスタリング手法では K-means、可視化手法ではまとめ表という組み合わせが現状では最適だと考えられる。

6. おわりに

本稿では、クエリに沿った学習者の解答の概要を講師が素早く把握することを目指とした。そのために、現在構築を目指している支援システムのクラスタリング手法と結果の可視化手法を比較するため実験を行った。実験から、クラスタリング手法では K-means、可視化手法ではまとめ表という組み合わせが最適であるという結果が得られた。しかし、クラスタリング手法や可視化手法は今回紹介した方法以外にも複数ある。また、当システムは記述のベクトル化についても十分な検討がなされていないため、今後はベクトル化手法についても検討を重ねていきたい。

参考文献

- (1) 中島英博:“多人数講義で学生の深い学習を促す教員の特質”, 名古屋高等教育研究, Vol.15, pp.161-177 (2015).
- (2) 西森敏之:“大学生の授業における態度と数学教師の対策—日本数学会のある調査より—”, 高等教育ジャーナル—高等教育と生涯学習—, vol.6, pp.1-31 (1999).
- (3) 田中博, 斎藤健:“ニューラルネットワークを用いた記述式問題の自動採点”, 言語処理学会 第22回年次大会 発表論文集, pp.123-126 (2016).
- (4) 伊藤慎治, 大庭知也, 高瀬治彦, 川中普晴, 鶴岡信治:“記述式解答群の主要な内容の把握 支援—三段階表示の効果の検証—”, 2016年度 PCカンファレンス論文集, pp.41-44 (2016).
- (5) 二村駿輔:“記述式演習の解答群の概要把握の支援—大規模言語モデルを利用したインタラクティブな要約—”, 三重大学大学院工学研究科 修士論文 (2023).
- (6) Jianlv Chen, Shitao Xiao, Peitian Zhang, Kun Luo, Defu Lian, Zheng Liu. “M3-Embedding: Multi-Linguality, Multi-Functionality, Multi-Granularity Text Embeddings Through Self-Knowledge Distillation.” arXiv preprint arXiv: 2402.03216 (2024)