

異なる学習形態から得られる出席ログの解析を目指したログの可視化の検証

An integrated analysis of logs obtained from different study styles

讃岐 勝^{*1}, 柳川 信^{*2}Masaru Sanuki^{*1}, Makoto Yanagawa^{*2}^{*1}筑波大学医学医療系^{*1}Institute of Medicine, University of Tsukuba^{*2}筑波大学国際局^{*2}Bureau of Global Initiatives, University of Tsukuba

Email: sanuki@md.tsukuba.ac.jp

あらまし：本稿では、コロナ禍で実施されたオンデマンド授業・ハイブリッド授業・ハイフレックス授業と異なる学習形態から得られたログを一元的に扱い解析した一例について述べる。ログの取得方法から異なるためログの標準化および画像化することでクラスタリングを行い学習形態の推移について分類を行い、学生の行動について検討を行う。

キーワード：ログ解析、ハイブリッド授業、対面授業

1. はじめに

コロナ禍で実施された授業の実践例、動画視聴の解析など多くの報告がなされている。ログを利用した研究として、ページの遷移を観察することによる習熟度の考察などがある⁽¹⁾。種々の授業形態に関する報告も多く報告されているが各々の授業に関して1つの視点で俯瞰する研究はほとんど存在しない。

本稿では、筑波大学医学群医学類（医学部医学科に相当。以下、医学類）の2020年4月～2023年3月までの異なる授業形態から得られた出席ログを用いて解析を行う。オンデマンド授業(2020.4-2020.8)、対面とオンデマンドを選択できるハイブリッド授業(2020.9-2021.3)、対面とライブ配信を選択できるハイフレックス授業(2021.4-2022.3)と異なる授業形態から得たアクセスログを用いて、対面と配信される授業を選択する学生の傾向についてログから考察する。

解析するため、ログをクラスタリングするが、ことなる学習形態を一元的に扱うため、2章ではログの標準化と画像化について述べ、3章で解析結果の検討をし、4章でまとめる。

2. ログの標準化と画像化

2.1 ログの取得

対面授業では各教室に設置されたカードリーダーに学生証をかざすことで出席が記録される。オンデマンドおよびライブの配信は Mediasite 社の配信サーバを利用して実施した。視聴はブラウザを利用してシボレスに認証によって学生を認証する。接続時間と閉じた時間によって接続されていた時間を記録する。接続する URL は講義室によって異なるが、連続して実施される場合もあるため 1-2 限、3-4 限、5-6 限で URL を切り替えるにした。

2.2 標準化

本稿では、異なる授業形態における出席時間を可視化する。出席時間は次のように計算する。対面授

業の出席時間（参加時間）は、“授業終了時間と授業開始時間との差”と“授業終了時間と記録時間との差”の最小値。オンデマンド授業は視聴時間、ライブ授業は対面授業と同様の方法で計算した（筑波大学の授業は1コマ75分）。複数コマ連続の場合、ライブ授業の記録では、最長 75+75+15 の最大 165 分の視聴時間となる(15分は休み時間)。

2.3 画像化

ログを標準化した上で、どの授業形態での参加なのか解析する上では必要である。本稿ではログを画像化することで1つの画像の中に異なる授業形態のログを埋め込むことにした。特徴量を出すために画像して解析することはゲノム解析などでも行われる手法であり、本稿ではこれを採用する。

画像は次の手順でユーザごとに作成する。

1. 3つの授業形態についてフラグをつける。
2. 授業の参加時間（対面）、ないし、視聴時間（オンデマンド）を計算する。
3. 時間軸を横軸、縦軸を参加時間、ないし、視聴時間とした棒グラフを授業形態ごとに色を変えてプロットする。

次の図は、ある1学生のログを上記手順によって可視化したものである。

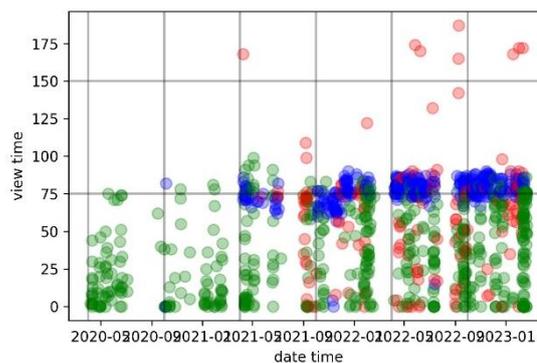


図1 ログを可視化したもの

図1より、次の情報を得ることができる。

- 2020年度はオンデマンドによる視聴
- 2021年は対面授業とオンデマンド(おハイブリッド)による出席
- 2022年は対面授業をメインにライブ配信も利用。オンデマンドによる復習を行っている。

視聴時間0分の記録が多くされているが、サーバへの接続が不安定であったことを表しており、データの不備ではないことを注意しておく。

出席記録を確実にするため授業ごとに再表示をするユーザもあり、図2のような画像として区別できる画像が生成されている(この学生はライブ配信を多く利用したユーザ。復習のためにあまり利用していないことが図から読み取れる。)。300分以上の接続はブラウザを閉じずにそのまま放置されていたことを表す。

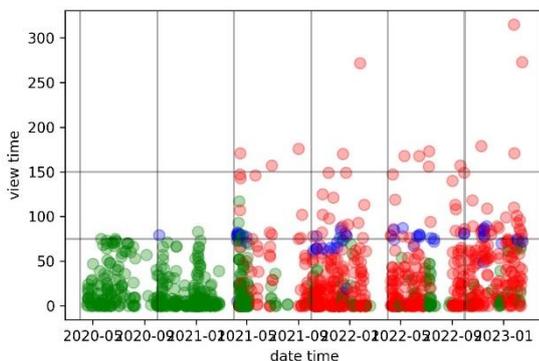


図2 授業ごとにブラウザの再読み込みをした学生

それゆえ、画像自身は情報を適切に表現する方法であるとみなして、次の章にある実験に進んだ。

3. クラスタリングと解析

画像化したログをクラスタリングし、ログを分類する。本稿では、k-means法によってクラスタリングをscikit-learnのモジュールを用いてpython上で実行した。

対象は、2020年度から2022年度に在籍がある筑波大学医学群医学類生の1110名である。休学・留年などの情報は反映せず、登録された学生をそのまま利用した。途中の年度で卒業した学生について、次年度以降は何も視聴していないことが記録されることから、今回は卒業した学生はログから区別できると判断し特別な検討はしない。

3.1 結果と検討

学年を分けずに解析したため、計8学年分の学生の行動をクラスタリングすることになってしまった。そのため、妥当な分割数が100以上と膨大になってしまった。ただ生成させた各クラスタを見ると、次のような傾向が見られた。

- 対面授業の学生は、PCで視聴する学生よりもオンデマンドによる復習は行わない。実施

したとしても限られた箇所(授業中、疑問に思ったことと推察)を視聴している傾向があった。

- 対面授業からライブ配信へ授業を切り替える学生は多かった。一方で、ライブ配信から対面授業に切り替える学生も一定数いた。数名に直接確認したところ、出席ログが保存されていることを確認できたので気切り替えた、接続が不安定で出席ログが残されているか不安に感じたため、切り替えていた、と出席確認は出席のために寄与すると考えられる。
- 出席ログを残すために、対面授業に参加しながら視聴をする学生も存在した。

学年ごとに傾向があるとすると、学年ごとのクラスタ数は10程度になるはずである。図1~図3は同じ学年の画像ログであるが、それぞれ別のクラスに分類されている。

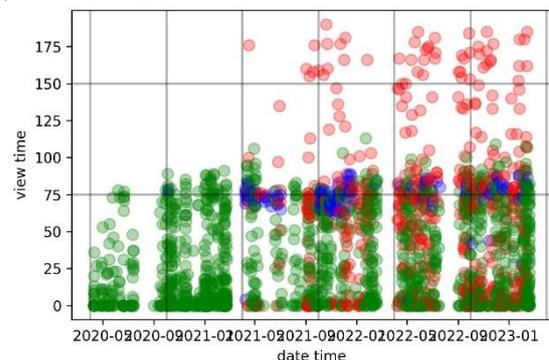


図3 色がたくさんある画像

入学年ごとにログを同一手法でクラスタリングを行うと10程度のクラスに分類される。対面授業を主にした学生、配信を利用した学生、オンデマンドコンテンツを利用した学生、復習に多く利用・多少利用といった傾向を見て取ることが出来た。

4. まとめ

本稿では、授業の参加時間、ないし、視聴時間をもとにログを画像化し、それをクラスタリングすることで学生の傾向を読み解くことに試みた。学習の傾向を観察するだけであれば本稿で述べた方法は有効であると考えられる。ただし、いずれの学習形態においても授業を理解することが最優先事項であり、成績データとの連結が不可欠であることを認識した。また、視聴開始時間のデータを組み合わせることで、夜に学習するユーザが実は多数存在して、朝学校にいけないのでライブ配信を視聴していたのではないかなど深く解析をすることが可能と思われる。

参考文献

- (1) J. A. Larusson, B. White (Eds.), Learning Analytics: From Research to Practice, Springer (2014).