

2ZM-02

# ノートへの記述と web 閲覧記録から生徒の興味に寄り添った情報を 収集・提供する web スクレイピングツールの開発

藤咲喜幸<sup>†</sup> 加藤直樹<sup>†</sup>東京学芸大学<sup>†</sup>

## 1. はじめに

中央教育審議会では「教育課程部会における審議のまとめ」（令和3年1月25日中央教育審議会初等中等教育分科会教育課程部会）が取りまとめられた。そこでは今後の教育課程の在り方について、ICTを新たな学校における土台にし、資質・能力の育成を着実に進める。そして、多様な子供たちを誰一人取り残すことなく育成する「個別最適な学び」の充実が求められる。また、子どもが「個別最適な学び」を進めていくためには、ICTの活用により、学習履歴（スタディ・ログ）を分析・活用することが重要である。

本稿では、学習へのモチベーション低下を防止し自主学習の機会を増加させることと、多様な個性を持つ子供たちの自主学習を支援することを目標とし、ノートへの記述と web 閲覧履歴から生徒の興味関心を分析し、生徒一人一人の興味に寄り添った内容を収集したスクレイピングページを自動で生成するツールを提案し、その開発と評価を通して、ツールの有用性を検証することを目的とする。

既製品に関しては、ユーザの興味に寄り添った情報を選別し収集できるアプリの一つとして SmartNews がある。このアプリケーションでは、閲覧履歴や閲覧頻度の高いチャンネルジャンルからユーザの趣味嗜好に合わせて欲しい情報のみを抽出して閲覧できるという特徴がある。本研究では閲覧履歴に加えノートの記述を用いてユーザの興味関心に合わせた情報を抽出する。

web スクレイピングに関する研究として、松延ら[1]は最新の医薬関連の情報を迅速かつ適切に入手することが求められる薬剤師のために、

web スクレイピングを行い医薬関連の情報を瞬時に得られるかを検討した。その結果として、インターネットにある膨大な情報を適切な解析をすることで、望むデータとして整理することが可能であることが示している。

## 2. 自動スクレイピングツールの設計

### 2.1 基本コンセプト

本稿で提案するツールでは、現在学習している事柄に興味を持ってない、有用性を感じられない生徒、及び教科書の内容だけでなくより発展的な内容を知りたいと考える生徒に対して、学習へのモチベーション低下を防止し自主学習の機会を増加させること、そして、多様な個性を持つ子供たちの自主学習を支援するために、生徒が授業内外でノートに記録した事項と、生徒の web ブラウザでの閲覧履歴から、生徒の学習状況及び興味関心を分析し、生徒一人一人の興味に寄り添った内容を収集したスクレイピングページを自動で生成する。

### 2.2 基本設計

#### 2.2.1 ノート選択機能

本ツールでは、生徒が自宅学習でノートを見返しする際を想定する。そこで、紙のノートをタブレット端末のカメラを用いて読み込む機能を用意する。そして、読み込まれたノートは、次に述べるキーワード指定機能で選択できるキー

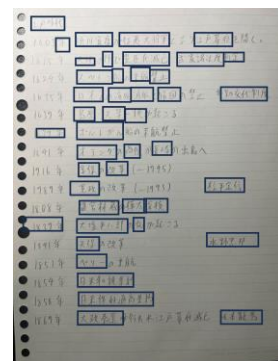


図1 キーワード指定画面

Developing a Web Scraping Tool for Collecting and Providing Information Tailored to Student Interests Predicted from Notebook Entries and Web Browsing Histories

<sup>†</sup> Kiichi Fujisaku, <sup>†</sup> Naoki Kato

<sup>†</sup> Tokyo Gakugei University

ワードであることを示すために、単語ごとに枠をつけて画面上に表示する（図1）。

### 2.2.2 キーワードの指定機能

本ツールでは、ノートに記述された単語の中で、生徒にとってイメージがつかなかった単語及び興味を引き付ける単語についての情報を提供することを目的にしている。そこで選択し表示されたノート内の単語をタップして指定する機能を用意する。

### 2.2.3 スクレイピングページ生成機能

指定されたキーワードに基づくスクレイピングページを生成する機能を用意する。生徒の興味関心に沿った情報の提供のために、生徒の web 閲覧履歴から取り出した頻出単語と選択されたキーワードを元に生成するスクレイピングページと、生徒の新たな知見を増やすための情報提供のために、Google のサジェストキーワードと選択されたキーワードとを元に生成するスクレイピングページの二通り用意する。

### 2.2.4 文章評価による生徒情報の更新

生徒の興味関心にすり合わせていき、各生徒独自の個別最適な学びを実現するために、スクレイピングページの文章を読んだ後、それぞれの文章を評価する機能を用意する。評価は肯定／否定の二択とし、肯定評価された場合は、キーワード指定で次回同じ単語が選択された場合に同内容の文章が表示されるようにする。一方で、否定評価された場合、同内容の文章が今後表示されなくする。

## 2.4 スクレイピングページ生成アルゴリズムの設計

頻出キーワードを用いてスクレイピングページを生成する方法では、web ブラウザの閲覧履歴から web ページを取り出し、形態素解析を用いて頻出単語をランキング化する。検出された上位 2 単語を頻出キーワードとし、「（選択された単語）（頻出キーワード）」を検索キーとした web 検索をして、最上位に表示された web サイトの文章を pysummarization で要約した文章をページの内容とする。

サジェストキーワードを用いてスクレイピングページを生成する方法では、キーワードリサーチツールであるラッコキーワードで選択した単語のサジェストキーワードを検索し、上位 2 単語を用いて「（選択された単語）（サジェストキーワード）」を検索キーとした web 検索をして、

最上位に表示された web サイトの文章を同様に要約したものをページの内容とする。

ここで、取得した頻出キーワードとサジェストキーワードの中に、生徒が過去に否定評価をつけた単語があった場合は、そのキーワードは選択せず、下位のキーワードを用いる。このことによって、生徒の興味にそぐわなかった文章の表示を制限し、生成する文章を更新していくことで、生徒の興味にマッチした内容を表示する。

なお、選択された単語が過去に肯定評価した単語である場合は、上記のアルゴリズムは動かさず、その際に生成したページを表示する。

## 3. 実装・試作

本ツールは、Visual Studio Code を用いて開発を行い、Windows11 上の Google Chrome を使用して動作確認を行った。

## 4. 予備的評価実験

提案した自動スクレイピングページ生成機能が、生徒にとって一人一人の興味に寄り添った自己学習を支援できているかどうか検証することを目的に評価実験を行った。被験者である大学生 10 人に本ツールを使用してもらいアンケートに答えてもらった。結果として、8 割の被験者がこのツールを利用することで個々の興味に寄り添った自主学習の時間を増加できると思うと回答した。

## 5. おわりに

本稿では、生徒が授業内外で記録したノートの記述、生徒のブラウザでの閲覧履歴から、生徒の学習状況及び興味関心を分析し、生徒一人一人の興味に寄り添った内容を収集したスクレイピングページを自動で生成するツールを提案し、開発した。評価実験からは、本ツールが個々の興味に寄り添った自主学習を支援できる結果が得られた。今後の課題として、書き込み機能の開発、評価機能を用いた web サイト選択精度の向上が挙げられる。

## 参考文献

- [1]松延他：Web スクレイピングを利用した医薬関連情報の収集と入手データの活用，第一薬科大学研究年報，第 38 号，pp. 51-64 (2022)