YouUnderstood.me? Academic material retrieval for students and educators

Ion Madrazo
Computer Science Department
Boise State University
Boise, Idaho, USA
ionmadrazo@boisestate.edu

Sole Pera
Computer Science Department
Boise State University
Boise, Idaho, USA
solepera@boisestate.edu

ABSTRACT

K-12 students make use of online resources to fulfill their academic information needs on a daily basis. However, they can often get discouraged because the contents they retrieve are outside their comprehension level, whether being too easy or too difficult for them to read. On the other side, educators find several challenges in finding materials for their classes that suit the students reading skills. Those reasons, make both students and educators spend a reasonably large amount of time seeking for adequate materials.

In this paper, we present a web application, that makes use of multidisciplinary techniques, such as, natural language processing, machine learning and information retrieval techniques, to help both students and educators in the process of finding materials that suit the reading skills of the students in a faster and more efficient way. For this purpose, the web application we present, puts together: (1) a material searcher that by combining a search engine and a readability formula, permits the fast retrieval of documents from different sources, (2) a readability tracking system that enables both types of users to see how the reading skills of the student are evolving with time and (3) an analysis tool that enables educators to analyse materials from outside the application for determining their complexity level.

CCS Concepts

•**Information systems** → **Personalization; Search interfaces;** *Clustering and classification;* •**Computing methodologies** → **Natural language processing;**

Keywords

Search engines; Filtering; Readability assessment; Student tracking

1. INTRODUCTION

K-12 students make use of the internet in a daily basis to fulfill their information needs for completing their academic

tasks, such as, finding information for a class presentation or discovering the meaning of a new word. For this purpose, they mostly use search engines and online available catalogues for retrieving contents such as, news articles, books or term definitions. However, they can often get discouraged because the contents they retrieve are outside their comprehension level, whether being too easy or too difficult for them to read.

In the academic environment students are not the only ones facing the problem of retrieving adequate contents in terms of reading abilities and information needs. For example, even in a same grade class, students' reading skills can differ significantly, so not all students in the same class can be provided with same texts. This supposes a personalization need that the instructor needs to handle on a daily basis. However with the high number of students in class this task becomes impossible to tackle. All of these reasons, make instructors spend a significant amount of their time seeking adequate materials for their students, a time that could be reduced with the help of an specialized application.

We have developed an application oriented to both instructors and students in the process of finding materials that suit the students reading level. The system is centered on a readability formula that together with a search engine makes looking for leveled reading material easier and more efficient. The system lets students log in in the web application, which keeps track of the materials read by each student and his feed-back for each or the read materials(too easy/OK/too complex). This enables the application to make suggestions about the readability level for each student. Both the student and the teacher are able too see the readability score in order to see how it evolves. Furthermore, the application users can use the included material searcher which retrieves results tailored to the readability score. The material searcher can act as an interface for various types of material sources depending on the capabilities of the education center that makes use of the application. Internal sources such as, school libraries or external resources such as online book catalogues like AR[1] or Lexile[2] or encyclopaedias such as Wikipedia.Moreover, even the internet can act as a source for the application, enabling it to be used as a conventional search engine, which filters results depending on the readability score. Cannot find a way of structuring this last part, beginning at "The material searcher..."

---

[1] http://www.acceleratelearning.com
[2] http://lexile.com

Apart from the search engine, the instructors have access to an analysis page, where they can submit texts they found outside the application for determining their readability score. This tool, together with the track of readability scores of each students, helps them make sure the materials they found are adequate or not for the class.
,

2. READABILITY ASSESSMENT

Most of the features of the application, from material retrieval filtering to external material analysis, are based on the complexity analysis of the reading materials. This analysis is performed by the XXXXX [Madrazo and Pera, ress] readability assessment tool, which, is capable of detecting the input language of a text on the fly and providing a readability score for it.

XXXXX is based on a supervised learning paradigm, at the moment, makes use of more than a hundred features for learning and prediction. The features are extracted using Natural Language Processing tools, such as a tokenizer, a part of speech tagger and different semantic analyzers. Those features, are used to train a model that will later be used for predicting the readability score for new materials.

3. MATERIAL SEARCHER

4. TRACKING STUDENTS

5. USER EXPERIENCE

- **Analyzing materials.** An analysis page is provided to instructors, so that they can make use of the readability assessment algorithm [Madrazo and Pera, ress] with materials from outside the application. This page provides an input form where the instructor can submit the material, and receive a readability level prediction for the it.

- **Looking for new materials.** A material search page is provided so that the user can insert text queries for looking for materials. The user can select a certain readability level he wants to look for, of leave it blank letting the application choose the best level for the logger user. Furthermore, the material source can be selected, depending on the user needs. In case of students, they can select a material for reading and provide feed-back on it. This feed-back will be used for tracking purposes.

- **Tracking students.** The educator can see a table with data about all his students, where the number of materials read and the readability score for each student are provided. The instructor can go deeper if needed, and see each of the materials each students has read and the individual readability scores for each material, as well as, the feed-back given by the student. The data is also presented in a summarized way, so that the educator can see the average, maximum and minimum reading skills of the students in class.

6. FUTURE WORK

7. MISSING IDEAS

- The application is fully multilingual.

8. REFERENCES

[Madrazo and Pera, ress] Madrazo, I. and Pera, S. (In press). Steps towards multilingual readability assessment tool.

[Robinson et al., 2000] Robinson, R. D., McKenna, M. C., and Wedman, J. M. (2000). Issues and trends in literacy education.