Majumdar, R. et al. (Eds.) (2025). Proceedings of the 1st International Conference on Learning Evidence and Analytics. Asia-Pacific Society for Computers in Education

Multidimensional Feature-Based Textbook Difficulty Assessment Index

Yu BAI[a], Fuzheng ZHAO[b] & Wenhao WANG[a] & Chengjiu YIN[c*]
[a]*Graduate School and Faculty of Information Science and Electrical Engineering, Kyushu University, Japan*
[b]*Education Technology Center, Jilin University, China*
[c]*Research Institute for Information Technology, Kyushu University, Japan*
*yin.chengjiu.247@m.kyushu-u.ac.jp

Abstract: This study addresses limitations of existing textbook difficulty assessment methods, often relying on single-dimensional metrics. Grounded in cognitive load theory, a multi-dimensional feature index is proposed, integrating five key dimensions: linguistic complexity, formula density, diagram complexity, knowledge abstraction, and structural disorganization. The index is constructed using techniques from natural language processing, image analysis, and knowledge graphs. Linguistic features are derived through tokenization and syntactic analysis; LaTeX formulas are detected via regular expressions; diagram complexity combines structured data and image texture features; knowledge abstraction uses dynamic terminology matching; and structural disorganization is assessed through chapter detection and coherence analysis.

Keywords: Textbook difficulty assessment, multi-modal feature fusion, cognitive load theory, natural language processing

1. Introduction

Textbook difficulty assessment is vital for optimizing educational outcomes by balancing cognitive load (Sweller, 2010). Traditional readability indices, such as Flesch-Kincaid, focus solely on linguistic complexity, neglecting multi-modal elements like formulas and diagrams (Flesch, 1948). Modern textbooks integrate text, visuals, and domain-specific content, requiring a comprehensive evaluation. This study introduces a multi-dimensional framework based on cognitive load theory, addressing linguistic, mathematical, visual, knowledge, and organizational dimensions to improve assessment accuracy and support learning.

2. Textbook Difficulty Assessment Index Design

2.1.1 Linguistic Complexity

It captures linguistic complexity as a proxy for intrinsic cognitive load. It measures intrinsic cognitive load through text complexity. Using NLP tools like NLTK, first apply tokenization and sentence segmentation. Then, average sentence length is determined by dividing total word count by sentence count, with a minimum of one sentence to avoid division by zero. Finally, the proportion of complex words is calculated by counting words after removing stopwords and dividing by total word count, adjusted with a minimum of one word. These metrics assess reading difficulty beyond traditional readability formulas (Graesser et al., 2011).

2.1.2 Formula Density

Formula Density This index measures formula density as an indicator of intrinsic cognitive load in processing mathematical content. LaTeX formulas are extracted using regular expressions, and density is assessed by counting the number of formulas and dividing by the total word count, with a minimum of one word to ensure a valid ratio, yielding a value typically between 0 and 1, though it may exceed 1 for very short texts. This provides a standardized measure of mathematical complexity (Sojka & Líška, 2011).

2.1.3 Diagram Complexity

This index evaluates diagram-related extraneous cognitive load by integrating structured data and image features. It assesses complexity by combining structural features (based on row density, method diversity, and TF-IDF of column names) and visual features (based on contrast and dissimilarity from GLCM texture analysis), with structural components weighted at 60% and visual at 40%, reflecting their relative contributions. This evaluates visual processing demands (Ojala et al., 2002).

2.1.4 Knowledge Abstraction

This study examines knowledge abstraction to assess cognitive load in processing specialized terminology, where high abstraction increases reading difficulty (McNamara et al., 2014). It is measured by counting matched terms from specialized vocabularies, dividing by a threshold of 8 terms per segment, and capping the result at 1.0 to indicate the density of domain-specific terminology relative to word count. It quantifies the abstraction level of domain-specific content.

2.1.5 Structural Disorganization

This index evaluates the extraneous cognitive load caused by structural disorganization in textbooks, focusing on chapter detection, text length, and discontinuity factors. It is computed by considering three factors—chapter segmentation (inversely related to section count), document length (normalized against a maximum word count), and textual discontinuity (based on sentence length variation and newline frequency)—with weights of 20%, 30%, and 50% respectively, resulting in a score from 0.1 to 1.0. Higher scores indicate greater structural complexity (Graesser et al., 2011).

3. Discussion and Conclusion

The proposed multi-dimensional textbook difficulty assessment model offers theoretical innovation by integrating text, formulas, diagrams, knowledge, and structure, enhancing STEM content evaluation. Methodologically, it uses regular expressions, TF-IDF, and image processing for scalable analysis. Limitations include untested predictive validity and limited adaptability to non-STEM fields. Future work should validate with diverse samples, explore dimensional interactions, and develop real-time assessment tools.

Acknowledgements
Part of this work was supported by JSPS KAKENHI (grant number 25K00842)

References

Flesch, R. (1948). A new readability yardstick. Journal of applied psychology, 32(3), 221-233.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. Educational Researcher, 40(5), 223–234.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). Automated evaluation of text and discourse with Coh-Metrix. Cambridge University Press.

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7), 971–987.

Sojka, P., & Líška, M. (2011). Detection of mathematics in LaTeX documents. Proceedings of the 14th International Conference on Text, Speech and Dialogue, 289–296.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. Educational psychology review, 22, 123-138.