Balancing Thematic Relevance and Cognitive Load in Adaptive Text Recommendation for Learners with SLDs

Francesco Lazzarotto[1], Giulia Coucourde[1,*] and Federica Cena[1]

[1]Department of Computer Science, University of Turin, Turin, Italy

Abstract

Individuals with Specific Learning Disorders (SLDs), and dyslexia in particular, often face barriers to equitable access to reading materials, as texts should be both cognitively accessible and aligned with their thematic interests to support sustained engagement. Existing approaches have addressed these dimensions in isolation, focusing on either readability adaptation or preference-based recommendation, without jointly optimizing both objectives. In this paper, we propose an adaptive text recommendation system that explicitly balances thematic relevance and cognitive load for SLD learners. Each user is modeled through a dual-channel profile comprising a semantic topic vector, capturing evolving thematic interests, and a readability target, tracking the user's current accessibility threshold. A hybrid scoring function combines semantic similarity with an asymmetric cognitive penalty, ensuring that texts exceeding the user's readability threshold are significantly demoted while thematically relevant content is prioritized. The system operates over a preprocessed corpus derived from OneStopEnglish, where each document is represented by an SBERT embedding and a Flesch Reading Ease score, and evolves continuously through explicit and implicit user feedback. We conducted a preliminary offline study with simulated users using standard ranking metrics, including NDCG@K and Precision@K, in which relevance is jointly determined by semantic similarity and readability alignment. Results on the OneStopEnglish dataset demonstrate that the hybrid approach achieves strong performance (HNDCG-Cos ≈ 0.54, HP-Cos ≈ 0.55, and NDCG@10 ≈ 0.90), suggesting its potential to balance thematic relevance with cognitive accessibility.

Keywords

User Modeling, Recommender Systems, Inclusive Education, Social Good, Accessibility

1. Introduction

Reading is a fundamental skill for academic and social engagement, but for individuals with Specific Learning Disorders (SLDs), particularly dyslexia, it is a continual source of cognitive labour and frustration. Dyslexic readers encounter significant challenges not only with decoding and phonological processing but also with the lack of accessible texts that fit with their authentic thematic interests, hence diminishing motivation and prolonged engagement. From the perspective of inclusive education, ensuring equitable access to engaging and cognitively appropriate reading materials is essential to support participation, autonomy, and long-term learning outcomes. Previous research suggests that genuine thematic interest plays a decisive role in reading engagement: when individuals with dyslexia read about topics they care about, they tend to persevere, develop deeper comprehension, and achieve cognitive gains that would otherwise remain inaccessible [1]. Existing AI-based approaches to supporting dyslexic learners have highlighted the potential of AI-driven interventions in education [2]. A substantial body of work focused on text simplification, demonstrating that reducing lexical and syntactic complexity yields measurable benefits for reading speed and comprehension. A parallel line of research explored recommender systems for SLDs, leveraging collaborative filtering and self-assessment data to model user preferences. However, to the best of our knowledge, no prior system jointly optimizes thematic relevance and cognitive load within a unified, adaptive recommendation framework. In this work, we frame the problem as a text recommendation task under cognitive accessibility constraints, aiming to identify content that is not only cognitively accessible but also genuinely aligned with the user's thematic interests. To address this gap, we propose an adaptive recommendation system that represents

Joint Proceedings of the ACM UMAP Workshops 2026, UMAP 2026, June 8–11, 2026, Gothenburg, Sweden

francesco.lazzarotto@edu.unito.it (F. Lazzarotto); giulia.coucourde@unito.it (G. Coucourde); federica.cena@unito.it (F. Cena)

0009-0002-2912-8721 (F. Lazzarotto); 0009-0000-9051-386X (G. Coucourde); 0000-0003-3481-3360 (F. Cena)

each user through a dual-channel profile: a semantic topic vector, capturing evolving thematic interests, and a readability target, tracking the user's current cognitive accessibility threshold. Recommendations are generated by a hybrid scoring function that explicitly balances semantic similarity with a cognitively asymmetric penalty, ensuring that texts exceeding the user's readability threshold are significantly demoted while thematically relevant content is prioritized. The system operates over a preprocessed corpus derived from OneStopEnglish, where each text is represented by an SBERT embedding and a Flesch Reading Ease score, and evolves continuously through explicit and implicit user feedback.

2. Related Work

Research on personalized support for dyslexic learners spans two partially overlapping strands: accessibility-oriented personalization and recommender systems; our work lies at their intersection by addressing text recommendation as a ranking problem, while explicitly personalizing the cognitive accessibility constraints.

2.1. Accessibility-Oriented Personalization for Dyslexic Learners

A substantial body of research has shown that readers with dyslexia benefit from forms of personalization that reduce processing difficulty and better align content delivery with individual cognitive needs. A first line of work concentrates on text adaptation, particularly lexical and syntactic simplification. Previous research indicates that higher-frequency words enhance reading speed, whereas shorter words and simplified constructions can support comprehension, thus reducing the cognitive effort associated with reading [3, 4]. These approaches are closely related to readability research, but they move beyond generic text-level formulas by considering the practical effects of simplification strategies on dyslexic readers. A second line of work shifts from text-centric adaptation to reader-centric personalization. Early work on user-specific readability classification argued that text difficulty should not be treated as an intrinsic property of the document alone, but rather as a relation between document characteristics and reader abilities [5]. This perspective is particularly relevant in dyslexia, where the same text may be manageable for one learner and cognitively overwhelming for another. Building on this idea, subsequent research explored user modeling approaches for dyslexic students in virtual learning environments, linking personalization to cognitive traits, learning preferences, and modality-specific needs [6]. Recent evidence further substantiates the significance of personalized adaptation. A web-based study demonstrated that personalized visual and auditory parameters can considerably enhance the reading experience compared to standard interface configurations [7]. Likewise, adaptive systems have begun to model learning progression over time, for example, by combining dynamic item generation, difficulty management, and student modeling techniques to support individualized intervention [8]. More recently, Large Language Models (LLMs) have begun to be explored as tools for automated text simplification, demonstrating the ability to generate accessible versions of complex content at multiple readability levels, with measurable benefits on comprehension and perceived cognitive load [9, 10]. Collectively, these studies establish a clear consensus: personalization is beneficial for learners with dyslexia, especially when it addresses readability, presentation, and cognitive load. However, most of this literature focuses on making content more accessible, rather than on recommending content that is both accessible and aligned with users' thematic interests.

2.2. Recommender Systems for Reading Support and Dyslexia

The landscape of Recommender systems has progressively evolved to support a wide range of information-seeking and learning processes, moving from generic content retrieval toward increasingly personalized approaches that account for individual preferences, skills, and contextual needs. In educational settings, these systems are commonly used to align learner profiles, preferences, and skill levels with pedagogical resources, with the broader goal of improving engagement, autonomy, and learning effectiveness [11, 12]. Within reading-related applications, early approaches proposed personalized recommendations for digital-libraries, matching books to users based on reading ability and profile information, though without explicitly targeting the dyslexic population [13]. More recent work has moved toward dyslexia-aware recommendation settings. Hybrid recommender systems have been developed in which users' preferences and requirements are collected through self-assessment questionnaires and interviews, then represented using structured formats such as graph-based databases to support personalized recommendations [14]. Additional research provides dual-system recommender architectures that dynamically adapt word recognition tasks according to the learner's evolving capabilities [15]. Similarly, recent work on adaptive sequential recommendation for dyslexia has modeled student trajectories over time in order to optimize intervention and predict future learning needs [8]. These systems are closer to recommendation in the strict sense, because they select or rank candidate items based on estimated utility for the learner. Beyond the dyslexia domain, the idea of incorporating user capability into recommendations has also been explored in behavioral change contexts, where Rasch-based models have been used to match recommended actions to a user's estimated ability level [16]. Yet capability alone is insufficient to drive engagement; the thematic relevance of recommended content plays an equally important role.

Interest modeling has also played an important role in recommendation research. In the broader recommender systems literature, several studies have shown that representing a user through a single preference vector may be insufficient, motivating architectures capable of capturing multiple latent interests from behavioral history [17]. In dyslexia-specific settings, interests have been elicited through self-assessment ratings and incorporated into recommendation models that predict future preferences by comparing users with similar profiles [18]. A more socially grounded perspective has modeled reading interest as a temporally evolving and peer-influenced phenomenon, showing that the preferences of dyslexic learners may also be shaped by group interactions and shared histories [19]. Regardless of these advances, most of the approaches in this domain prioritize preference modeling, without explicitly incorporating a learner-specific model of cognitive accessibility into the ranking function. As a result, recommended items may be thematically relevant but cognitively misaligned with the reader's abilities, thereby limiting sustained engagement.

3. System Design

In this section, we present the proposed architecture (Figure 1), organized into three main components: content preparation, online recommendation, and user profile adaptation. This design aims to accomplish three goals: (i) traceability, (ii) minimizing online latency, and (iii) supporting reproducibility through centralized parameter control.

In the *Offline Processing* stage, text collections are transformed into reusable representations capturing both readability features and semantic information, thereby shifting computationally intensive operations outside the online pipeline and avoiding inconsistencies between the user state and the content repository. The *Online Recommendation* module employs user state to generate texts that meet user interests and readability constraints. A hybrid scoring algorithm that integrates multiple signals ranks these candidates to create a Top-K recommendation list (Section 4.2). This technique avoids lexical or semantic-only requirements. Finally, the *Adaptive Loop* updates the user's thematic representation, readability target, and interaction history based on successive user interactions (including explicit feedback and implicit signals). The system maintains persistent user states across interactions, enabling adaptation over time. Dashed lines in the picture represent penalty rules and hyperparameters for external control, allowing flexible tuning without affecting component operation while keeping modularity, interpretability, and extensibility for future study and experimentation.

4. Methodology

For individuals with specific learning disabilities, text recommendations must balance thematic relevance and cognitive accessibility. Instead of choosing texts based on user preferences, the goal is to find

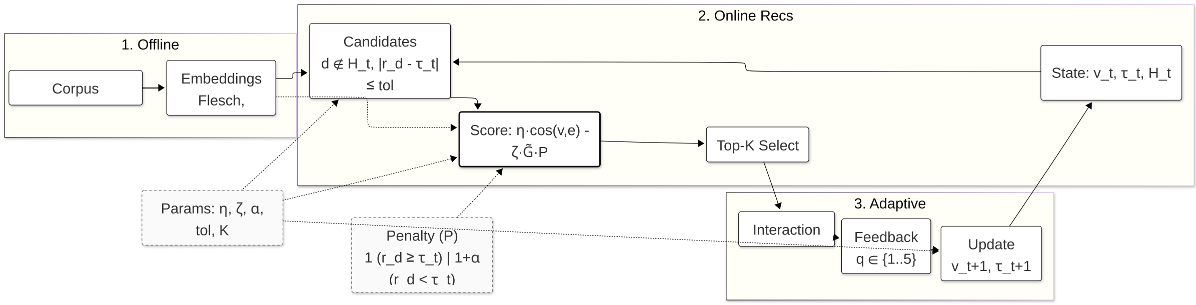

Figure 1: Architectural workflow of the adaptive recommendation system. The diagram highlights the three-phase pipeline: (1) Offline feature extraction and embedding generation; (2) Online multi-objective scoring balancing semantic relevance and readability; (3) The feedback-driven adaptive loop for dynamic profile recalibration.

documents that improve semantic alignment and readability in relation to the user's current state. We model this task as a personalized ranking problem over a document collection $\mathcal{D} = \{d_1, \dots, d_N\}$. Each document d_i is represented by a pair (e_i, r_i), where $e_i \in \mathbb{R}^{384}$ is a semantic embedding and r_i is the Flesch Reading Ease score. The user state at time t is defined as: $u_t = (v_t,\ \tau_t,\ H_t)$, where v_t is the topic vector, τ_t is the readability target, and H_t is the reading history.

4.1. User Modeling

The user profile is structured into three layers: semantic memory, difficulty memory, and episodic memory. Semantic memory represents a continuous mapping that resides in the latent space shared by documents, providing consistency and comparability in the form of similarity geometry. The difficulty memory is used to represent the targeted difficulty range. Finally, episodic memory helps avoid recommendation repetition, maintaining marginal utility. The initialization of the topic vector uses the normalized centroid of the embedded corpus:

$$v_0 = \frac{\dfrac{1}{N}\sum_{i=1}^{N} \dfrac{e_i}{\|e_i\|_2}}{\left\| \dfrac{1}{N}\sum_{i=1}^{N} \dfrac{e_i}{\|e_i\|_2} \right\|_2} \tag{1}$$

This initialization acts as a non-informative semantic prior: by anchoring the topic vector at the centroid of the corpus distribution, Equation 1 prevents premature drift toward irrelevant thematic areas before the adaptive loop captures the first authentic interest signals. After every read, the topic vector is revised using the technique of interpolation as follows: $v_{t+1} = \mathrm{norm}\big((1-\alpha_q)\, v_t + \alpha_q\, \hat{e}_{d_t}\big)$. In this case, the feedback-mapping function $f(q)$ transforms the user-reported difficulty level (q $\in 1, 2, 3, 4, 5$) into an adaptation weight ($\alpha_q \in 0.1, 0.2, 0.3, 0.4, 0.5$), with a monotonic relation on the valid domain: $f(q) = \alpha_q = 0.1q$.

Here, $q = 1$ indicates low engagement or perceived difficulty, while $q = 5$ indicates high engagement and perceived accessibility. From this perspective, higher levels of feedback indicate higher flexibility of the user profile, as the algorithm gives greater importance to interactions carrying stronger informational value. In parallel, the readability target (τ_t) is updated as: $\tau_{t+1} = \mathrm{clip}\,(\tau_t + \Delta\tau_t, ; 20, ; 90)$. This bounded interval prevents both excessive cognitive overload (<20, typically academic/technical prose) and overly simplistic, non-age-appropriate language (>90). Specifically, the update step $\Delta\tau_t$ is formalized as: $\Delta\tau_t = \gamma \cdot f(q) \cdot (r_{d_t} - \tau_t)$ where $\gamma \in (0, 1)$ represents the global damping factor controlling the maximum shift per session, $f(q)$ is a feedback-dependent learning weight mapping the user's explicit/implicit feedback q to an impact magnitude, and $(r_{d_t} - \tau_t)$ is the observed readability distance of the interacted document. The application of minimal shift thresholds in each feedback category ensures that there are no insignificant changes in situations where the user input is significant.

From a theoretical perspective, this design implements a dual-channel dynamic system, where the *semantic channel* determines what to recommend, while the *cognitive channel* regulates how the recommendation should be.

Dataset We use a preprocessed version of the OneStopEnglish dataset [20], which provides multiple versions of each text at different readability levels. This structure enables controlled evaluation of recommendation strategies that jointly consider semantic relevance and readability. Each document is represented by a semantic embedding and a readability score. Specifically, texts are encoded using SBERT (`all-MiniLM-L6-v2`, 384 dimensions), while readability is quantified using the Flesch Reading Ease score. This hybrid representation combines semantic expressiveness with interpretable readability features, supporting the identification of texts that are both topically relevant and cognitively accessible.

4.2. Recommendation Engine

The recommendation engine follows a three-step process: candidate generation, scoring, and ranking. There are two constraints for the set of candidates: (a) no documents already read by the user are considered - H_t (the user's reading history); and (b) the documents must be near the user's level of readability - tol. The candidate set is defined as:

$$C_t(u) = \{d \in \mathcal{D} \setminus H_t \mid |r_d - \tau_t| \leq \text{tol}\}$$

This selection balances search space coverage with cognitive realism, limiting candidates to documents that are both thematically diverse and within the user's readability range. Hybrid scores are then computed for each candidate, sorted in decreasing order, and the top $K = 10$ documents are returned. The computational cost is $\mathcal{O}(n \log n)$, dominated by the sort operation, making the pipeline feasible for interactive inference. In terms of methodology, the proposed architecture conforms to practical settings where: (i) there are relatively small numbers of users or evolving user profiles; (ii) immediate feedback is required from the system; (iii) simplicity in the ranking criterion is preferred over complexity.

4.2.1. Scoring Function

The scoring function represents the theoretical core of the recommendation engine, explicitly encoding the utility function to be maximized. It implements a linear scalarization of two competing objectives: maximizing semantic relevance and minimizing readability mismatch. Formally:

$$S(u, d) = \eta \cdot \cos(v_u, e_d) - \zeta \cdot \tilde{G}(u, d) \cdot P(u, d) \tag{2}$$

The first term $\eta \cdot \cos(v_u, e_d)$ measures semantic alignment between the user profile and the document in the latent space. The second term $\zeta \cdot \tilde{G}(u, d) \cdot P(u, d)$ is the cognitive penalty, which increases as the document deviates from the user's readability target.

Readability gap. Since the raw readability distance $|\tau_u - r_d|$ numerically dominates the cosine similarity term due to scale differences, we introduce a local normalization relative to the tolerance threshold ($T = \text{tol}$):

$$\tilde{G}(u, d) = \min\left(\frac{|\tau_u - r_d|}{T},\ 1\right)$$

Local normalization relative to T is preferred over global MinMax rescaling, which was empirically found to overly compress the penalty range, causing ranking instability.

Asymmetric penalty. On the Flesch scale, lower scores indicate harder texts, so $r_d < \tau_u$ denotes a document more demanding than the user's readability target τ_u:

$$P(u, d) = \begin{cases} 1 + \alpha, & r_d < \tau_u \\ 1, & r_d \geq \tau_u \end{cases}$$

This asymmetric formulation reflects the clinical-educational priority of avoiding cognitively overwhelming elements. Texts within or below τ_u (i.e., $r_d \geq \tau_u$) incur only a linear penalty based on $\tilde{G}(u, d)$, supporting gradual skill acquisition. Texts exceeding the tolerable difficulty limit ($r_d < \tau_u$) are further demoted by α, prioritizing the reduction of cognitively overwhelming content, irrespective of its semantic attractiveness. This design relies on neuropsychological evidence showing that text difficulty has a nonlinear effect on dyslexic readers, with above-threshold texts triggering a disproportionate cognitive decline and entering the so-called frustration level, where reading sub-processes can no longer be integrated effectively [21, 22]. The parameters η and ζ govern the fundamental trade-off between semantic relevance and cognitive accessibility. Altering the ratio η/ζ produces qualitatively distinct recommendation behaviors. When $\eta \gg \zeta$ (Semantic Supremacy), the system prioritizes thematic relevance at the expense of cognitive accessibility. Users receive texts aligned with their interests but containing linguistically demanding structures that exceed their readability threshold. For individuals with dyslexia, this results in instant cognitive exhaustion, frequent phonological mistakes, and progressive disengagement from the task. Conversely, when $\zeta \gg \eta$ (Cognitive Dominance), the system becomes overly conservative, producing readable but thematically unengaging recommendations that fail to sustain motivation, a known limitation of purely accessibility-driven approaches.

5. Evaluation

We evaluate the offline effectiveness of the recommendation engine on the OneStopEnglish dataset (4.1), using standard Information Retrieval metrics [23], namely NDCG@K and Precision@K. Relevance is defined along two dimensions: (i) semantic similarity between document embeddings and the user interest vector (computed via cosine similarity over SBERT representations), and (ii) alignment between the document's Flesch score and the user's readability target. NDCG@K assesses the quality of the ranking by assigning higher importance to relevant documents appearing at the top of the recommendation list. Precision@K measures the proportion of relevant documents within the top-K recommendations. While NDCG@K and Precision@K quantify ranking quality and relevance concentration, they do not explicitly capture semantic coherence in the top-ranked results. We therefore complement them with MeanCos@5, HNDCG-Cos, and HP-Cos.

$$\text{MeanCos@}K = \frac{1}{K}\sum_{i=1}^{K}\cos(v, e_{d_i}) \qquad \text{DCG@}K = \sum_{i=1}^{K}\frac{2^{\text{rel}_i}-1}{\log_2(i+1)}$$

$$\text{P@}K = \frac{1}{K}\sum_{i=1}^{K}\mathbf{1}[\text{rel}_i \geq 1] \qquad \text{NDCG@}K = \frac{\text{DCG@}K}{\text{IDCG@}K}$$

$$\text{HNDCG-Cos@}K = \frac{2\,\text{NDCG@}K \cdot \text{MeanCos@}K}{\text{NDCG@}K + \text{MeanCos@}K} \qquad \text{HP-Cos@}K = \frac{2\,\text{P@}K \cdot \text{MeanCos@}K}{\text{P@}K + \text{MeanCos@}K}$$

MeanCos@5 measures semantic coherence through a metric that computes the alignment of topics between the user and the documents retrieved for the first five results. The metric directly assesses the ability of the system to retrieve relevant content based on the semantic interests of the user rather than just satisfying readability constraints. With this regard, HNDCG-Cos combines this measure of semantic coherence with the effectiveness of the ranking system in order for higher values to be generated only when both conditions are fulfilled. Likewise, HP-Cos works to complement this notion by linking semantic coherence with a high density of relevant entries at the beginning of the ranked list. Overall, the two metrics work together to provide balanced measures that do not oversimplify the performance of a single model in terms of optimizing one condition without the other.

5.1. Empirical Study

We evaluate 10 independent seeds with stratified simulated users, using $\eta = 0.6$, $\zeta = 0.4$, $\alpha = 0.4$, tol $= 15$, $\gamma = 0.6$, and $K = 10$. Table 1 reports mean ± standard deviation across seeds. Parameters were set following a principled constraint: since the asymmetric penalty P(u,d) already amplifies the cognitive component for texts exceeding the user's readability threshold, assigning $\zeta > \eta$ would result in a double penalization of the readability dimension, producing overly conservative recommendations (as discussed in Section 4.2.1). Accordingly, the semantic weight was constrained to $\eta \geq 0.5$. Within this admissible range, the specific values $\eta = 0.6$ and $\zeta = 0.4$ were selected via a small-scale grid search over the simulated user protocol, optimizing for the harmonic balance between HNDCG-Cos and Precision@K. Simulated users were generated through a stratified multi-seed protocol to balance diversity and experimental control. For each of 10 independent seeds, we created 128 profiles (8 topic clusters x 4 readability anchors x 4 users per cell), for a total of 1280 profiles. Topic strata were obtained via K-Means over normalized SBERT embeddings. Each user's readability target was sampled around anchors (30, 45, 60, 75) with Gaussian noise and clipped to [20, 90]. Initial history included 5 documents sampled without replacement from the assigned topic stratum, prioritizing documents close to the sampled readability target. The initial user vector was computed as the normalized centroid of history embeddings, with small stochastic perturbation to increase within-cell heterogeneity. This protocol yields broad yet structured synthetic variability, ensures systematic coverage of the topic-readability space, and supports robust reporting across seeds (mean +- standard deviation).

Method	NDCG@10	Precision@10	Precision@5	MeanCos@5	$H_{NDCG-Cos}$	H_{P-Cos}
Readability only	**0.9996 ± 0.0000**	**0.9891 ± 0.0056**	**0.9975 ± 0.0032**	0.1674 ± 0.0044	0.2828 ± 0.0062	0.2826 ± 0.0061
Similarity only	0.5168 ± 0.0086	0.6261 ± 0.0170	0.6464 ± 0.0286	**0.4906 ± 0.0049**	0.4948 ± 0.0053	0.5374 ± 0.0133
Hybrid	0.9036 ± 0.0030	0.9706 ± 0.0068	0.9822 ± 0.0065	0.3929 ± 0.0053	**0.5367 ± 0.0063**	**0.5520 ± 0.0065**

Table 1
Ablation with local gap normalization and asymmetric penalty on harder texts. The hybrid objective provides the best joint readability-semantic harmonic balance.

Leakage across versions. On the same protocol, leakage at Top-10 (same article family in history and recommendation) is: Readability only 0.0250 ± 0.0040, Hybrid 0.2268 ± 0.0105, Similarity only 0.4461 ± 0.0141. The Hybrid leakage rate is partially expected, as recommending the same article at a different readability level may support scaffolded skill acquisition. Nevertheless, to limit redundancy, future work will extend H_t to track entire article families rather than individual documents, excluding same-family candidates from the candidate set.

6. Limitations and Future Work

To address the limitations that come with offline evaluation, we plan to incorporate a user-based evaluation, which relies on two measures: the Completion Rate and the System Usability Scale (SUS) [24]. Completion Rate, defined as the ratio of completed to attempted texts, serves as a behavioral proxy for cognitive adequacy, under the assumption that cognitively overloaded texts are more likely to be abandoned. It enables comparison between experimental and control conditions, providing insight into the effectiveness of the proposed recommendations relative to traditional selection criteria. As opposed to the behavioral metrics, the latter, in particular, the SUS, is designed to capture subjective measures of the system's usability and usefulness. At the same time, the adopted metrics primarily assess the system's internal consistency rather than its clinically relevant effectiveness, which introduces a potential risk of evaluative circularity. Moreover, the Flesch score, originally validated on neurotypical populations [25, 26], may not fully reflect the cognitive load experienced by dyslexic readers, for whom factors such as the density of complex graphemes and lexical familiarity are critical [27]. As a result, high ranking performance (e.g., NDCG) does not necessarily indicate the actual appropriateness of the recommended texts. User-based evaluation introduces further challenges: the target population is inherently heterogeneous [28], and metrics such as the Completion Rate do not imply actual comprehension. For instance, Likert-scale feedback can be systematically distorted by emotional factors typical of individuals with SLDs. Detecting an improvement causally attributable to the system would require a controlled longitudinal design using clinically validated instruments (such as CTOPP and GORT-5 [29, 30]) that fall outside the scope of this study. Despite these limitations, the adopted evaluation strategy provides a first step for assessing the consistency of the proposed approach.

7. Conclusion

This work introduces an adaptive text recommendation system aimed at supporting individuals with SLDs, particularly dyslexia, by jointly modeling thematic relevance and cognitive accessibility. The proposed approach relies on a dual-channel user representation, combining a semantic topic vector with a dynamic readability target, and a hybrid scoring function that integrates semantic similarity with an asymmetric cognitive penalty. Preliminary offline evaluation on the OneStopEnglish dataset suggests that the proposed method can produce rankings that balance thematic relevance with readability constraints, yielding promising results in terms of both NDCG and Precision. While these findings provide an initial indication of the potential of the approach, they remain limited to simulated user scenarios and offline metrics.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT and Grammarly in order to: Paraphrase and reword. The authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] R. Fink, Why Jane and John couldn't read–and how they learned: A new look at striving readers., International Reading Association, 2006.

[2] A. Paglialunga, S. Melogno, The effectiveness of artificial intelligence-based interventions for students with learning disabilities: a systematic review, Brain Sciences 15 (2025) 806.

[3] S. S. Bautista, P. Gervás, R. I. Madrid, Feasibility analysis for semiautomatic conversion of text to improve readability, in: Information and Communication Technologies and Accessibility, 2009. URL: https://api.semanticscholar.org/CorpusID:15286886.

[4] L. Rello, R. Baeza-Yates, S. Bott, H. Saggion, Simplify or help? text simplification strategies for people with dyslexia, in: Proceedings of the 10th international cross-disciplinary conference on web accessibility, 2013, pp. 1–10.

[5] C. Litsas, M. Mastropavlou, A. Symvonis, Text classification for children with dyslexia employing user modelling techniques, in: IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications, IEEE, 2014, pp. 1–6.

[6] F. E. Benmarrakchi, J. El Kafi, A. Elhore, User modeling approach for dyslexic students in virtual learning environments, International Journal of Cloud Applications and Computing (IJCAC) 7 (2017) 1–9.

[7] M. L. Lorusso, F. Borasio, P. Panetto, M. Curioni, G. Brotto, G. Pons, A. Carsetti, M. Molteni, Validation of a web app enabling children with dyslexia to identify personalized visual and auditory parameters facilitating online text reading, Multimodal Technologies and Interaction 8 (2024) 5.

[8] J. I. M. Trujillo, I. Rodríguez-Rodríguez, D. Castillo-Barnes, A. Ortiz, A. Sánchez, J. L. Luque, Optimizing dyslexia intervention through an adaptive sequential recommender system, Knowledge-Based Systems (2025) 114309.

[9] J. Qiang, M. Huang, Y. Zhu, Y. Yuan, C. Zhang, K. Yu, Redefining simplicity: Benchmarking large language models from lexical to document simplification, arXiv preprint arXiv:2502.08281 (2025).

[10] T. Guidroz, D. Ardila, J. Li, A. Mansour, P. Jhun, N. Gonzalez, X. Ji, M. Sanchez, S. Kakarmath, M. M. Bellaiche, et al., Llm-based text simplification and its effect on user comprehension and cognitive load, arXiv preprint arXiv:2505.01980 (2025).

[11] M. C. Urdaneta-Ponte, A. Mendez-Zorrilla, I. Oleagordia-Ruiz, Recommendation systems for education: Systematic review, Electronics 10 (2021) 1611.

[12] G. Lampropoulos, Recommender systems in education: A literature review and bibliometric analysis, Advances in Mobile Learning Educational Research 3 (2023) 829–850.

[13] M. Omisore, O. Samuel, Personalized recommender system for digital libraries, International Journal of Web-Based Learning and Teaching Technologies (IJWLTT) 9 (2014) 18–32.

[14] K. El Hage, A. Remadi, Y. Hobeika, R. Ma, V. Hong, F. Bugiotti, A multi-source graph database to showcase a recommender system for dyslexic students, in: 2023 IEEE International Conference on Big Data (BigData), IEEE, 2023, pp. 3134–3138.

[15] J. I. Mateo-Trujillo, D. Castillo-Barnés, I. Rodríguez-Rodríguez, A. Ortiz, A. Peinado, J. L. Luque, A. Sánchez-Gómez, Dual-system recommendation architecture for adaptive reading intervention platform for dyslexic learners, in: International Work-Conference on the Interplay Between Natural and Artificial Computation, Springer, 2024, pp. 205–214.

[16] A. Starke, M. Willemsen, C. Snijders, Effective user interface designs to increase energy-efficient behavior in a rasch-based energy recommender system, in: Proceedings of the eleventh ACM conference on recommender systems, 2017, pp. 65–73.

[17] C. Li, Z. Liu, M. Wu, Y. Xu, P. Huang, H. Zhao, G. Kang, Q. Chen, W. Li, D. L. Lee, Multi-interest network with dynamic routing for recommendation at tmall, 2019. URL: https://arxiv.org/abs/1904.08030. arXiv:1904.08030.

[18] G. Morciano, J. M. Alcalde Llergo, A. Zingoni, E. Yeguas Bolívar, J. Taborri, G. Calabrò, Use of recommendation models to provide support to dyslexic students, Expert Systems with Applications 249 (2024) 123738. URL: http://dx.doi.org/10.1016/j.eswa.2024.123738. doi:10.1016/j.eswa.2024.123738.

[19] E. Ahmed, D. Mincu, L. Harrell, K. Heller, S. Roy, Study: Socially aware temporally causal decoder recommender systems, 2023. URL: https://arxiv.org/abs/2306.07946. arXiv:2306.07946.

[20] Sowmya, nishkalavallabhi/onestopenglishcorpus: Onestopenglish corpus release, 2018. doi:10.5281/zenodo.1219041.

[21] K. Stanovich, Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy, Reading Research Quarterly 21 (1986) 360–407. doi:10.1598/RRQ.21.4.1.

[22] M. Treptow, M. Burns, J. McCOmas, Reading at the frustration, instructional, and independent levels: The effects on students' reading comprehension and time on task, School psychology review 36 (2007) 159–166. doi:10.1080/02796015.2007.12087958.

[23] C. D. Manning, P. Raghavan, H. Schütze, Introduction to Information Retrieval, Cambridge University Press, Cambridge, UK, 2008. URL: http://nlp.stanford.edu/IR-book/information-retrieval-book.html.

[24] J. Brooke, Sus: A quick and dirty usability scale, Usability Eval. Ind. 189 (1995).

[25] R. Flesch, A new readability yardstick, Journal of Applied Psychology 32 (1948) 221–233. doi:10.1037/h0057532.

[26] P. Kincaid, R. P. Fishburne, R. L. Rogers, B. S. Chissom, Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel, 1975. URL: https://api.semanticscholar.org/CorpusID:61131325.

[27] L. Rello, R. Baeza-Yates, L. Dempere-Marco, H. Saggion, Frequent words improve readability and short words improve understandability for people with dyslexia, in: P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, M. Winckler (Eds.), Human-Computer Interaction – INTERACT 2013, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 203–219.

[28] M. J. Snowling, Dyslexia, 2nd ed., Blackwell Publishers, Oxford, UK, 2000.

[29] J. M. Hintze, A. L. Ryan, G. Stoner, Concurrent validity and diagnostic accuracy of the dynamic indicators of basic early literacy skills and the comprehensive test of phonological processing, School Psychology Review 32 (2003) 541–556. doi:10.1080/02796015.2003.12086220.

[30] A. H. Hall, Test review: J. l. wiederholt & b. r. bryant. (2012). gray oral reading tests–fifth edition (gort-5), Journal of Psychoeducational Assessment 31 (2013) 516–520. doi:10.1177/0734282912468578.