Project Proposal:
BookBuddy
Book recommendation website for children

CSE 6242 - Spring 2025

Team: Interactive insight (Team 205)
Huyen Trang Ho, Monika Nikam, Ong Jia Wei, Mohammad Parsamehr

Georgia Tech

1. Background and Project Objectives

Problem:

There are a few existing websites that offer appealing book recommendations for children, they often lack the depth of personalization needed to cater to individual preferences and reading abilities (Ng, 2016).

Project Objective:

To create a fun and personalized book recommendation website for children, helping them discover books that match their reading levels and interests.

2. Existing Solutions and Expected Innovations

Limitations of Existing Systems:

- Non-appealing interfaces to gauge user preferences.
- Lacks comprehensive dataset of children's books, including reviews and detailed descriptions.

Challenges in Recommendation Methods:

- **Collaborative Filtering:** sparse data (Ng, 2020) and the cold start problem (Yuan & Hernandez, 2023)
- **Content-Based Filtering:** may lack diversity in recommendations (Xue et al., 2017)

Expected Innovation: BookBuddy will integrate:

- Visually appealing UI
- Hybrid Recommendation Method
- NLP Integration
- Data Utilization

Features	StoryTime (Milton et al., 2019)	BookTrust	Our website (BookBuddy)
Visually appealing User Interface	❌	✅	✅
Personalized Recommendation	❌	❌	✅
Recommendation Method	Content-based Filtering	Content-based Filtering	Hybrid + NLP

Georgia Tech

3. Our Solution: BookBuddy

User-Centered Design: Engaging UI with visual elements for children.	Use React/Flutter for an interactive image-based UI.
NLP Integration: Extract genres, characters, and themes from book descriptions.	Use SpaCy, NLTK, Hugging Face for text processing.
Data Utilization: Build a rich dataset of children's books with metadata.	Utilize Goodread Review scrapped dataset (Wan et al., 2019).
Recommendation Algorithm: Hybrid approach using NLP and collaborative filtering.	Python for building Recommendation Algorithms.

Georgia Tech

4. Impacts and Measurements

Why would BookBuddy be successful?
BookBuddy will succeed by leveraging historical user data, advanced filtering techniques, and a scalable, cost-effective infrastructure to deliver personalized and engaging book recommendations for children.

Who Cares?
- **Parents and Children**: Parents seek tools that encourage their children's reading habits, while children benefit from engaging and personalized book recommendations.
- **Education System**: Educators and institutions aim to foster reading skills and literacy among students, aligning with curriculum goals.

Impacts
- **Provide a Fun UI for Children to Choose Books**: An intuitive and engaging interface can make book selection enjoyable, leading to increased interaction.
- **Encourage Children's Reading Habits**: Personalized recommendations can motivate children to explore new books, enhancing their reading frequency and diversity.

Measurement Methods:
- **User Studies & Feedback:** Conducting user surveys and A/B testing to refine recommendations.
- **Engagement Metrics:** Analyzing session duration, books explored, and user retention.
- **Recommendation Accuracy:** Evaluating recommendation relevance using ground truth data comparison.

Georgia Tech

5. Risks, Payoffs and Expected Costs

Risks

- Cultural Diversity
- Dataset Bias & Cold Start
- Scalability and Cost Constraints
- Privacy Concerns

Payoffs

- Increased Literacy & Engagement
- Scalable Hybrid Recommendation System
- Future Expansion Opportunities

Cost

$0

By using open source and free techstacks (Render, Github)

Duration

10 weeks

Georgia Tech

6. Plan of Activities

Phase	Week 1 - 2 (Feb 17 - Feb 28)	Week 3 - 4 (Mar 3 - Mar 14)	Week 5 - 6 (Mar 17 - Mar 28)	Week 7 - 8 (Mar 31 - Apr 11)	Week 9 - 10 (Apr 14 - Apr 18)
Proposal Research & Data Collection		Proposal Submission Feb 28			
UI/UX design & development					
Algorithm Development & Progress Report				Midterm Checkpoint Progress Report Mar 28	
System Integration					
Testing & Debugging					
Deployment, Feedback & Final Presentation					Final Checkpoint Deploy & Presentation Apr 18

Georgia Tech

Literature Survey

Paper 1: "Book Recommendation System through Content-Based and Collaborative Filtering Method"

- **Main Idea:** Combines content-based filtering, collaborative filtering, and association rule mining to improve book recommendations.
- **Usefulness:** Personalized book suggestions, enhances recommendation accuracy.
- **Shortcomings:** Cold start problem, complexity of combining multiple algorithms.

Paper 2: "StoryTime: Eliciting Preferences from Children for Book Recommendations"

- **Main Idea:** A web-based system for children using visual input and participatory design to address cold start problems.
- **Usefulness:** Child-friendly interface, valuable user-centered design insights.
- **Shortcomings:** Limited data, simplified preferences.

Paper 3: "CBRec: A Book Recommendation System for Children Using Matrix Factorization and Content-Based Filtering Approaches"

- **Main Idea:** Uses Matrix Factorization and Content-Based Filtering to recommend age-appropriate books.
- **Usefulness:** Tailored book suggestions, promotes literacy development.
- **Shortcomings:** Data sparsity, limited by predefined grade levels.

Literature Survey

Paper 4: "User Cold Start Problem in Recommendation Systems: A Systematic Review"

- **Main Idea:** Reviews strategies to overcome the user cold start problem from 2016-2023.
- **Usefulness:** Updated cold start solutions, new methods and metrics.
- **Shortcomings:** Limited focus on item cold start, lacks novel solutions.

Paper 5: "Towards Context Aware and Age-Based Book Recommendation System using Machine Learning for Young Readers"

- **Main Idea:** Context-aware, age-based recommendation system using K-Nearest Neighbors.
- **Usefulness:** Personalized by age, reduces cold start with contextual data.
- **Shortcomings:** Challenges in assessing contextual data, scalability issues.

Paper 6: "Age Recommendation from Texts and Sentences for Children"

- **Main Idea:** Machine learning models to recommend age-appropriate texts for children based on linguistic features.
- **Usefulness:** Tailors reading material to children's comprehension levels.
- **Shortcomings:** Language limitation (French dataset), model explainability issues.