Analyzing the Preferences and Personal Needs of Teenage Readers to Make Book Recommendations

Sophie Gao and Yiu-Kai Ng
Computer Science Department, Brigham Young University
Provo, Utah, USA
sophiegao99@gmail.com,ng@compsci.byu.edu

ABSTRACT

Reading is one of the main sources of learning, especially for young readers such as teens. Promoting good reading habits among teens is essential, given the enormous influence of reading on teenagers' development as learners and members of society. For this reason, it is imperative to motivate young readers to read by offering them appealing books to read so that they can enjoy reading and gradually establish a reading habit during their formative years that can aid in enhancing their learning attitude. Books, which provide an indispensable source of reading materials, broaden the horizons of teenagers and allow them to learn from different disciplines, expand their perspective in decision making, and gain knowledge. These days there are a wide variety of books available on the market for teenagers, parents, and librarians to choose from. Due to the diversity of books, choosing a desirable book from a set of unfamiliar books to read is time-consuming and there is no guarantee of the satisfaction of its content. Existing book recommender systems, however, either focus on general audience or very young readers, such as children, and thus might not meet the specific needs of the particular group of users whom we target, i.e., teenagers. To make appropriate recommendations on books that are appealing to teenagers, we propose a book recommender system, called TBRec. TBRec recommends books to teenagers based on their personal preferences and needs that are determined by using various book features. These features, which include book genres, topic relevance, predicted user ratings, and readability levels, have significant impact on the readers' preference and satisfaction on a book. These distinguished parts of a book identify the type, subject area, (un)likeness, and complexity of the book content. Experimental results reveal that TBRec outperforms Amazon, Barnes & Noble, and LibraryThing in making book recommendations for teenagers, and the results are statistically significant.

CCS CONCEPTS

• **Information systems → Recommender systems**.

KEYWORDS

Teenagers, books, recommender systems, metadata

WI-IAT '21, December 14–17, 2021, ESSENDON, VIC, Australia

ACM ISBN 978-1-4503-9115-3/21/12…$15.00
https://doi.org/10.1145/3486622.3493972

ACM Reference Format:
Sophie Gao and Yiu-Kai Ng. 2021. Analyzing the Preferences and Personal Needs of Teenage Readers to Make Book Recommendations. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI-IAT '21), December 14–17, 2021, ESSENDON, VIC, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3486622.3493972

1 INTRODUCTION

Reading is an activity performed on a daily basis: from reading news articles and books to cereal boxes and street signs. According to the National Institute of Child Health and Human Development, "reading is the single most important skill necessary for a happy, productive, and successful life",[1] which is the reason why focusing on *teenager* (or early) *reading* that refers to the knowledge, skills, and dispositions acquired in reading (and writing) in school grades prior to and up till the 12^{th} grade [26], is particularly significant.

Books have long been an important source for their readers to learn, especially in the early ages. With the rapid and advanced development of electronic devices and the popularity of the Internet, teenagers nowadays are distracted by video games, video movies, smartphones, and social interaction on the web and neglect the importance of reading [8]. It was found that most teenagers prefer playing games to reading because games are more entertaining [18]. A large number of surveys, however, have shown that the benefits of reading are irreplaceable. Intensive reading can improve the extended knowledge of teenagers [18]. The importance of reading for teenagers is not limited to knowledge acquisition, but also helps them improve their reading comprehension, which plays an important role in understanding long and complex texts later in their professional studies [17]. To revive the interest of teenagers in reading, it is essential for librarians, teachers, and parents, besides teenagers, to efficiently identify books that the teenagers enjoy. However, choosing appropriate books to read is a time-consuming task and there is no guarantee that the chosen books will be satisfiable to the special demands of teenagers and thus is a major obstacle for people involved in exploring them. Existing book recommender systems such as Amazon, Barnes & Noble, and LibraryThing, however, rely on users' purchase patterns, the same authors, and access approaches, respectively to suggest books that often do not meet the needs of teenagers. In response, we have developed a new book recommender system, called TBRec, that is designed specifically for teenagers. TBRec suggests books to teenagers that they like and helps them develop a healthy reading habit.

TBRec is unique, since it considers and incorporates a wide variety of features of books to suggest appealing books to teenagers to enhance their reading experience. These features include the

[1] http://www.ksl.com/?sid=15431484

genres, topic relevance, predicted ratings, and readability level of books, which target this particular group of readers [22, 23, 30]. By integrating the type, subject area, (un)likeness, and complexity of the text of a book, TBRec can offer recommendations that meet the “needs" and “tastes" of the targeted users.

To verify the performance of the proposed TBRec, we have compared TBRec with Amazon, Barnes & Noble, and LibraryThing using various publicly available book datasets. Experimental results have (i) demonstrated that using all the book features available to our recommender system, TBRec outperforms each individual one in making suggestions, (ii) revealed that the ranked results of TBRec are superior to the other three book recommenders, and (iii) validated that books suggested by TBRec are considered more favorable than its counterparts, i.e., the ones recommendations made by Amazon, Barnes & Noble, and LibraryThing, respectively, based on user relevance feedbacks.

The remaining of this paper is organized as follows. In Section 2, we discuss existing searching tools and recommenders that have been used for identifying books for individual readers, including young readers. In Sections 3, we introduce our teenager book recommender system and its overall design methodology. In Section 4, we present the results of the empirical studies on our recommender system conducted to assess its performance. In Section 5, we give a concluding remark and present directions for future work.

2 RELATED WORK

To the best of our knowledge, there is no existing book recommendation system developed specifically for teenager readers. At present, parents/educators/young readers often rely on existing book websites, including, but not limited to, ARbookfind.com, Kidsread.com, Scholastic.com, and WorldCat.org, which offer different tools to search for books in various domains. These sites, however, either (i) supply (read-alike) non-personalized booklists, (ii) require a particular topic/subject area of interest to be selected from a pre-defined list,[2] which limits the themes of books that can be obtained from the sites, (iii) offer reading choices grouped by age/grade ranges,[3] which is undesirable, since readers in the same grade or age group might not reach the same reading level, or (iv) allow users to create keyword queries to specify their information needs, which often yield an overwhelming volume of items to choose from and impose an additional burden on users to sort though. Unlike the aforementioned websites, the proposed teenager book recommender system eliminates their constraints imposed in locating books, which enhances the process in finding books relevant to the information needs of teenager readers and meeting the preferences of the readers.

A number of existing book recommenders adapt the collaborative filtering algorithm to make suggestions. The personalized recommenders on college books proposed by Bhelawe et al. [3] and Li et al. [15], which apply the user-based collaborative filtering (CF) algorithm, focus on users' personal preferences and recommend the same books to similar users based on the categories to which they belong. Adapting both the user-based and item-based CF algorithms for book recommendations based on book ratings, Anwar and Uma [2] compare the performance of the two CF algorithms.

Besides using the CF approach, many researchers attempted to incorporate machine learning (ML) models into book recommendations. Yuan et al. [31] examine different ML methods and drawbacks of using big data to develop book recommenders. Singh and Jain [28] employ ML to develop a book recommender by using the K-Nearest-Neighbor algorithm, which has been shown to be more optimized than simply applying the CF approach alone. Kwan et al. [14] also develop a book recommender by combining the content-based filtering and a popularity-based strategy.

There exist other book recommenders that unilaterally consider a variety of approaches to recommend books. Sohail et al. [29] use an opinion mining technique based on user reviews to make recommendations. Pera [22] extracts books from the list of social media friends of a user and matches the extracted books with the user's reading preferences to recommend books to the user. Ali et al. [1] combine user's information needs and book content to determine the appropriateness of a book for recommendation. Jomsri [10] adapts a widely-used data mining technique for rule mining, whereas Garrido et al. [6] present a topic map-based approach to make book recommendations. Both Milton et al. [19] and Reuter [25] analyze the reading preferences of children based on their age gaps. Milton et al. design an easy-to-understand interface that categorizes books of interest to children for them to choose, whereas Reuter focuses on identifying the factors that influence children's book preferences, which can lead to book recommendations.

Our teenager book recommender system, TBRec, is different from existing book recommender systems, since the latter either recommend books for targeted audience other than teenagers that cannot be seamlessly adaptable for teenagers, or exclude the usage of various book features that are essential in making appealing/desirable book recommendations for teenagers. Even though TBRec is not a recommender for direct learning, its design goal is to enhance reading selections for teenager readers by locating suitable books among the overwhelming number of choices available these days. (For an in-depth description of existing recommenders in the educational domain, see [16].)

3 A TEENAGER BOOK RECOMMENDER

TBRec considers essential book features to address the reading preferences and needs of teenagers in making appealing book recommendations to them. These features include *genres, topic relevance, predicted user ratings*, and *readability Levels*, which analyze the category, topic, (un)likeness, and complication level of the content of a book that are especially applicable to teenage readers [22, 23, 30].

To assess and ascertain that recommended books are preferred and interested to a user, TBRec requests the user to provide a book, called *target book*, that the user enjoyed reading. Using the target book TB and a corpus C of published books, TBRec proceeds to analyze the four different book features and obtain four different feature scores for each corpus book CB with respect to TB. The closer the specific features of CB compared with TB are, the higher the combined feature score of CB is. To incorporate the four feature scores into a final ranking score for CB, TBRec uses a simple,

[2]http://www.readingrockets.org/books/booksbytheme
[3]http://www.scholastic.com/teachers/article/ready-go-book-lists-teachers

widely-used linear combination model, called *CombMNZ*. Based on the final ranking scores, TBRec recommends them to the user in a ranked order.

3.1 Book Genres

A genre of a book indicates a particular *category* of the book. The basic assumption of genres is that if two books have the *same* genre, then they are likely *similar* in terms of their plots and contents. *Genres* for books come with various properties: (i) two different genres can be highly similar, such as "dark" and "mystery", (ii) a book is assigned different genres by different experts/users, and (iii) some books have multiple genres assigned to them. For example, the genres of the book, "The Hunger Games" by Suzanne Collins, include *Young Adult, Fiction, Fantasy, Romance, Adventure, Action*, and *Apocalyptic*. Each of these genres is also associated with the number of users who have chosen the label to identify the category of the book, as in the Goodreads dataset.[4] *Similarities* of book genres can be determined using *word-correlation factors* [21] and TBRec computes the score of the genres, denoted GS, of a book in a corpus to determine its similarity to a target book in terms of their genres.

Given a corpus book, denoted CB, and a *target book*, denoted TB, these books should (i) share at least a *common* genre with TB, and (ii) its *genre score*, with respect to TB, denoted $GS(CB, TB)$, must be greater than the *genre score* of TB itself, denoted $GS(TB, TB)$, to ensure that CB and TB are close in terms of their categories. TBRec considers the *top-10* ranked genres (based on their numbers of users who assigned the genres) in computing the two scores.

We realize that solely relying on the genres of a corpus book CB to make recommendations could yield less-than-ideal results, since (i) a *large* amount of data are required to accurately determine a particular user's genre preferences, (ii) a lower (high, respectively) *genre similarity score* of CB does not necessarily indicate that a user dislikes (likes, respectively) CB, and (iii) books can be highly similar with respect to their genres and being dissimilar with respect to other features not related to genres. Therefore, TBRec considers other features of CB in making book recommendations as presented in subsequent sections.

3.2 Topic Relevance

TBRec analyzes the topic, i.e., subject area, of a corpus book by using topic modeling to determine the most dominated topic covered in the book. Latent Dirichlet Allocation (LDA) model is one of the most popular topic modeling methods. LDA generates a collection of topics by using the given text documents. Each topic is a list of keywords arranged by frequency of occurrence. Using the generated list of topics, each text document can be labeled by an LDA model with the possibility for each topic. TBRec adapts the LDA model to determine the topic of a corpus book.

To train an LDA model there are two steps involved: (i) a series of documents to be analyzed for different topics, and (ii) an ideal number of latent topics to be generated. TBRec uses the *descriptions* of corpus books as the set of documents for creating the topics. We extracted 18,000 book descriptions from the Goodreads dataset, which is publicly available, as our training instances. Each one of the book descriptions consists of a sequence of words. During the training process, LDA estimates the probability of a non-stop, stemmed word w given a (latent) topic z, i.e., $P(w|z)$, and the probability of a topic z given a text document D, i.e., a book description in our case, i.e., $P(z|D)$. A number of algorithms have been proposed for estimating $P(w|z)$ and $P(z|D)$, such as variational Bayes [13], expectation propagation [20], and Gibbs sampling [7]. We have chosen Gibbs sampling, since it is easier to implement, more efficient, faster to obtain good approximations, and easily extended than others [24]. As we tried different topic numbers between 5 and 40 for our book recommender system, we have chosen 20 topics as the ideal topic numbers, since 20 yield the highest relevance scores among different keywords within each chosen topic.

The classification process of LDA on a given corpus book CB can be described as finding the probabilities of a number of topics covered in the description D of CB and selecting the *topic* with the *highest probability* as the topic covered in D.

3.3 User Rating Prediction

Rating prediction is a classical approach for making recommendations. Attempts have been made in the past by relating users to similar users and an item to similar items on which user- and item-based rating prediction systems have been developed. TBRec adapts the item-based collaborative filtering (CF) approach to predict the rating of a corpus book for a target book so that the higher a predicted rating on an item I for a user u using the ratings of items previously encountered by u is, the more likely I is appealing to u. This recommendation strategy is intuitive and relatively simple to implement. Moreover, it requires no costly training phases that are needed for machine learning models and thus is scalable to millions of users and items. In addition, it is not significantly affected by the constant addition of users, items, and ratings in a large number of commercial applications and does not require retraining.

Based on the item-based CF approach, TBRec predicts the rating of a corpus book i for a target book by analyzing books previously rated by the target user u that are similar to i, denoted $N_u(i)$. Moreover, TBRec extends the item-based CF approach by determining the *degree of similarity* between i and each book j in $N_u(i)$, denoted $w_{i,j}$, to further enhance the item-based CF approach. TBRec applies the pre-defined word vectors, which capture the essential (i.e., non-stop, stemmed) keywords in the book descriptions of i and each book j in $N_u(i)$, to compute the (cosine) *similarity* between i and each book j in $N_u(i)$.

3.4 Readability Level Analysis

The readability level of a book is a useful measure for teenagers to identify reading materials suitable at their reading levels. The majority of published books, however, are assigned a readability level range, such as 8-12, by professionals, instead of a single readability level for their intended readers, which is not useful to the end-users who look for books at a particular readability level. This leads to the development of readability formulas/analysis tools. Unfortunately, these formulas/tools require at least a significant portion of a book content to estimate its readability level, which is a severe constraint due to copyright laws that often prohibit book content

[4] help.goodreads.com/s/article/How-do-I-get-a-copy-of-my-data-from-Goodreads

from being made publicly accessible. To alleviate this constraint imposed on readability analysis on books, we have developed our own readability level analysis tool, called *TRoLL* [5], which relies heavily on metadata of books that is publicly and readily accessible from reputable book-affiliated online sources, besides using previews of books[5], to predict the readability level of books.

3.5 CombMNZ

Based on the respective scores of the features, as discussed in Sections 3.1 through 3.4, computed for each corpus book of a target book, TBRec ranks all the corpus books accordingly. To compute a single score on which the cumulative effect of the four different features of each corpus book is determined for ranking propose, TBRec relies on the CombMNZ model [4]. CombMNZ is a well-established data fusion method for combining multiple ranked lists on an item I, i.e., a book in our case, to determine a *joint* ranking of I, a task known as rank aggregation or data fusion.

$$CombMNZ_I = \sum_{c=1}^{N} I^c \times |I^c > 0|, \text{ where } I^c = \frac{S^I - I^c_{min}}{I^c_{max} - I^c_{min}} \quad (1)$$

where N is the number of ranked lists to be fused, which is the *four* ranked lists in our case, I^c is the normalized score of I in the ranked list c, and $|I^c > 0|$ is the number of non-zero, normalized scores of I in the lists to be fused. Prior to computing the ranking score of a corpus book CB, we transform the original scores in each feature ranked list of CB into a *common range* $[0, 1]$ such that S^I is the score of I in the ranked list c to be normalized, I^c_{max} (I^c_{min}, respectively) is the maximum (minimum, respectively) score available in c, and I^c is the normalized score for I in c.

4 EXPERIMENTAL RESULTS

Given a target book and a set of corpus books, we evaluate the performance of TBRec by (i) determining the *relevance* of its recommendations, (ii) deciding the *ranking accuracy* of its recommendations, and (iii) comparing its suggestions with three well-known book recommenders, Amazon, Barnes & Noble, and LibraryThing. The usefulness of a recommended book provided by either TBRec, Amazon, Barnes & Noble, or LibraryThing is determined by a number of independent appraisers, which serve as the *gold standard* of the conducted empirical study. Based on the ranking on relevant recommended books with respect to a given (i.e., target) book determined by individual appraisers, the *degree of accuracy* of each individual recommendation made by each one of the four recommenders can be computed.

4.1 Data Source and Appraisers

To the best of our knowledge, there is no existing benchmark dataset that can be used for evaluating the performance of a teenager book recommender system. For this reason, we constructed our own dataset, using data made available for the public by BookCrossing[6], Goodreads[7], and Book Cave[8], three of the well-established websites for the book hobbyist community. BookCrossing, which is a free online book club, has established discussion forums, offered blogs, and accumulated more than 13 million books as of November 2019. Goodreads, on the other hand, is a social cataloging website that provides the facility for each user to search its database of books and reviews, whereas Book Cave publishes a database of over 11,000 teenagers books, and the web page for over 98% of books in the Book Cave database also includes a hyperlink to the Amazon Store web page for the Kindle (electronic) version of that book. From the Amazon Kindle Cloud Reader[9] web page we managed to collect the sample texts for the archived books.

As there is no well-established standard nor existing book recommender system that can be adapted or used for comparing the performance of teenager book recommender systems, we turned to students at a local junior high school to conduct an empirical study that allows us to evaluate the performance of TBRec. We needed these students, who were in the 7^{th}- and 8^{th}-grade of their respective class at a local junior high school, to serve as the appraisers of our study, since they are the targeted users of TBRec and have read the target books used in the study so that we could collect their feedback on various recommendation tasks. Altogether, a total of 45 students, who were 13- or 14-year-old, were recruited through their class teachers for our study using 15 target books. The *forty-five* students and 15 target books are ideal numbers for our study according to the calculation based on the *cross-over experiment* [11] so that the experimental results are reliable and objective. We detail the determination of these numbers in subsequent sections.

4.1.1 Determining the Number of Appraisers. In statistics, two types of errors, Types I and II, are defined [11]. Type I errors, also known as α errors or *false positives*, are the *mistakes* of *rejecting* a null hypothesis when it is true, whereas Type II errors, also known as β errors or *false negatives*, are the *mistakes* of *accepting* a null hypothesis when it is false. We apply the formula in [11] below to determine the ideal number of appraisers, n, which is dictated by the probabilities of occurrence of Types I and II errors, to evaluate books recommended by TBRec.

$$n = \frac{(Z_{\frac{\alpha}{2}} + Z_\beta)^2 \times 2\sigma^2}{\triangle^2} + \frac{(Z_{\frac{\alpha}{2}})^2}{2} \quad (2)$$

where $\triangle$ is the *minimal expected difference* to compare our recommendation approach with a librarian who manually chooses a book to be suggested to readers, which is set to 1 in our study as we expect our approach to make high-quality book recommendations as good as the ones made by librarians; σ^2 is the *variance*[10] of the recommended books, which is set to be 2.72 in our study (see the discussion on the computed value in the next paragraph); α (β, respectively) denotes the probability of making a Type I (II, respectively) error, which is set to be 0.05 (0.20, respectively), and 1 - β determines the probability of a false null hypothesis that is correctly rejected, and Z is the value assigned to the standard *normal distribution* of generated summaries. Based on the standard normal

[5]Previews of books can be extracted from the Book Cave(.com) dataset, which consists of more than 20,000 teenager books that are made available by publishers to showcase their books.
[6]https://www.bookcrossing.com/
[7]https://www.goodreads.com
[8]https://mybookcave.com/
[9]https://read.amazon.com/
[10]*Variance* is widely used in statistics, along with standard deviation (which is the square root of the variance), to measure the average dispersion of the scores in a distribution.

distribution, when $\alpha = 0.05$, $Z_{\frac{\alpha}{2}} = 1.96$, and when $\beta = 0.20$, $Z_{\beta} = 0.84$. (See the explanations on setting the α and β values given below.)

We conducted an experiment using a randomly sampled 50 target books to determine the value of σ^2. We chose only 50 books, since the *minimal expected difference* and *variance*, which are computed on a *simple random sample*, do not change with a larger sample set of books. σ^2 is computed by averaging the sum of the square difference between the mean and the actual number of *useful* recommendations[11] created for each one of the 50 target books. We obtained $\sigma^2 = 2.72$ for the recommendations.

The values of α and β are set to be 0.05 and 0.20, respectively, which imply that we have 95% *confidence* on the correctness of our analysis and that the *power* (i.e., probability of avoiding false negatives/positives) of our statistical study is 80%. According to [12], 0.05 is the commonly-used value for α, whereas 0.80 is a conventional value for 1 - β, and a test with $\beta = 0.20$ is considered to be statistically powerful. Based on the values assigned to the variables in Equation 2, the ideal number of appraisers used for our study is

$$n = \frac{(1.96 + 0.84)^2 \times 2 \times 2.72}{1^2} + \frac{1.96^2}{2} \cong 45 \qquad (3)$$

4.1.2 Determining the Number of Target Books. To determine the ideal number of target books as test cases to be included in the controlled experiments, we rely on two different variables: (i) the *average attention span* of a teenager and (ii) the *average number of test cases* that a person can handle at a time. As mentioned in [27], the average attention span of a teenager is between twenty to twenty-five minutes to spend on a web search engine in one session. Based on this study, each appraiser was asked to evaluate our book recommendations for *three* test cases, i.e., target books, since evaluating all the *eight* recommendations made for each one of the three test cases takes approximately 25 minutes, which falls into a teenager time span. Since each appraiser was committed to spend appropriately 120 minutes on the empirical study, we set up 5 sessions with 25 minutes each for each appraiser to conduct the empirical study. Altogether, a total of *15* (= 5 × 3) target books were used for the study.

4.1.3 Our Verification approach. The students, who participated in the performance evaluation of TBRec, were asked to determine which ones of the eight recommendations[12], if there were any, were relevant books with respect to the corresponding target book. The two books marked as *relevant most often* by the appraisers were treated as the *gold standard* for the target book.

4.2 Performance Evaluation

To evaluate the effectiveness of TBRec in recommending relevant and highly-ranked books to teenagers, we applied several performance measures commonly used in information retrieval and recommender systems: *precision@1* (P@1), *precision@2* (P@2, since each recommender suggests two books), and *Mean Reciprocal Rank* (MRR), which measure the *top-ranked* and *first* useful recommendation among all the ranked recommendations, respectively.

[11] A recommendation is considered *useful* if it is regarded as relevant to the corresponding target book determined by librarians recruited at a local school.

[12] Two each from TBRec, Amazon, Barnes & Nobles, and LibraryThing which were the top-2 recommendations made by the four recommender systems on a given target book, respectively. The appraisers had no idea which recommendation was made by which book recommender.

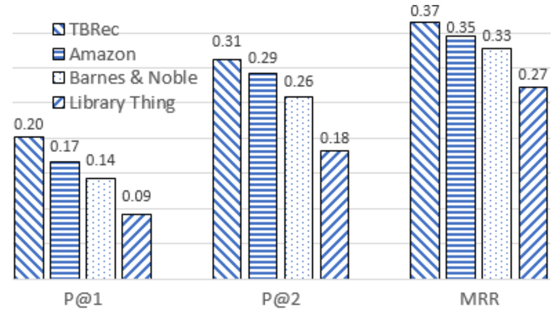

Figure 1: The P@1, P@2, and MRR values for various book recommenders, including TBRec

4.2.1 Comparing the Performance Evaluation of TBRec and Other Book Recommender Systems. Users of a recommender system tend to look at only the top few ranked results to find relevant recommendations. Some search tasks have only the top-ranked recommendation, i.e., P@1, in mind, whereas others might consider the top-5 ranked recommendations. After the gold standard for each one of the 15 test cases provided by the 45 appraisers were recorded, we computed the P@1, P@2, and MRR values for various book recommenders involved in our empirical study. Figure 1 shows the performance metrics for the *average precision* at rank 1 (i.e., average P@1) and *average Precision* at rank 2 (i.e., average P@2), in addition to the average of the reciprocal ranks at which the *first* useful recommendation (among all the ranked recommendations) for each target book is made, i.e., MRR. The P@1, P@2, and MRR scores of TBRec are higher than the corresponding scores of Amazon, Barnes & Nobles, and LibraryThing, respectively, and the results are statistically significant based on the Wilcoxon Signed-Ranks Test ($p < 0.05$).

4.2.2 Feature Evaluation of TBRec. To evaluate the performance of each individual feature used by TBRec and its accuracy in making ranked recommendations on its own, we calculated their P@1, P@2, and nDCG values using one feature at a time to make recommendations. (nDCG, the normalized discounted cumulative gain, penalizes useful suggestions that are ranked *lower* in the list of suggestions.) Hereafter, we calculated the same performance values using all the features combined, which is adapted by TBRec. All of these evaluations are based on the top-2 suggestions of each target book, and there are 13 test cases in this study. We have recruited 35 freshman students from our university, who are 18- or 19-year-old, to conduct the study (dictated by the cross-over experiment [11]) who come from various academic disciplines. Given a target book, these students were asked to rank the relevance of each recommended book (based on a brief description of its content), and their evaluations yield the ground truths of our study.

According to the performance values shown in Figure 2, which uses different test cases compared with the ones in Figure 1, we realize that the *Book Genres* feature is the most promising one, which is followed by the *Rating Prediction* feature, since they perform significantly better than other features in terms of recommending and ranking relevant books to the users. Collectively, using

Figure 2: The nDCG, P@1, and P@2 values for each feature and the combined features of TBRec

all of the features, the algorithm of TBRec outperforms each individual feature, and the result is statistically significant based on the Wilcoxon Signed-Ranks Test ($p < 0.02$). The average nDCG, P@1, and P@2 values indicate that suggestions made by TBRec rank higher than the corresponding ranked books recommended by each individual feature as determined by the college students.

5 CONCLUSIONS AND FUTURE WORK

The ability to read and comprehend what has been read is significant for many reasons. Through reading and comprehending, humans acquire knowledge and understanding of the world around them. This allows the reader to cooperate and accomplish tasks that (s)he would not be able to do without instructions and/or examples. To the youths, the contents of books read by them can stimulate their imagination and expand their understanding of the world. It helps them develop language and listening skills and prepares them to understand the written words. However, these days the numerous number of books in the market make it difficult for teenagers, parents, and librarians to find appealing and interested books for teenagers to read.

To help teenagers find suitable books more efficiently and easily, we have developed a unique book recommender system, called *TBRec*. TBRec combines multiple features specific to books to provide book recommendations for teenagers. The conducted empirical study shows that by combining multiple book features to make book recommendations, TBRec performs better than using only a single book feature. Moreover, books recommended by TBRec for teenagers are considered more favorable than the ones suggested by Amazon, Barnes & Nobles, and LibraryThing, respectively, which are three of the well-known, existing book recommender systems.

For future work, we would like to extend the performance evaluation on TBRec to determine the impact TBRec has on the reading and learning habits of teenager readers. Furthermore, we would like to enhance the functionality of TBRec by examining existing image-matching models and, if necessary, develop one that would allow us to perform a more in-depth examination of book illustrations, especially tailored towards younger teenagers. In doing so, we anticipate that more relevant book suggestions could be generated, which will improve the effectiveness of our proposed book recommender system for teenagers.

REFERENCES

[1] Z. Ali, S. Khusro, and I. Ullah. 2016. A Hybrid Book Recommender System Based on Table of Contents (TOC) and Association Rule Mining. In *Intl. Conf. on Informatics & Systems*. ACM, New York, 68–74.

[2] T. Anwar and V. Uma. 2020. Book Recommendation for eLearning Using Collaborative Filtering and Sequential Pattern Mining. In *ICDABI*. IEEE, 1–6.

[3] D. Bhelawe, S. Dhurat, J. Gharat, and S. Shinde. 2021. *Enhanced Book Recommendation System using Collaborative Filtering*. SSRN.

[4] G. Cormack, C. Clarke, and S. Buettcher. 2009. Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods. In *ACM SIGIR*. ACM, New York, 758–759.

[5] J. Denning, M. Pera, and Y.-K. Ng. 2016. A Readability Level Prediction Tool for K-12 Books. *Association for Information Science and Technology (JASIST)* 67, 3 (2016), 550–565.

[6] A. Garrido, M. Pera, and S. Ilarri. 2014. SOLE-R: A semantic and linguistic Approach for Book Recommendations. In *ICALT*. IEEE, New York, 524–528.

[7] A. Gelfand. 2000. Gibbs Sampling. *American Statistical Association* 95, 452 (2000), 1300–1304.

[8] R. Hopper. 2005. What Are Teenagers Reading? Adolescent Fiction Reading Habits and Reading Choices. *Literacy* 39, 3 (2005), 113–120.

[9] B. Jansen, A. Spink, and T. Saracevic. 2000. Real Life, Real Users, and Real Needs: a Study and Analysis of User Queries on the Web. *IPM* 36, 2 (2000), 207–227.

[10] P. Jomsri. 2014. Book Recommendation System for Digital Library Based on User Profiles by Using Association Rule. In *INTECH)*. IEEE, New York, 130–134.

[11] B. Jones and M. Kenward. 2003. *Design and Analysis of Cross-Over Trials, 2nd Edition*. Chapman and Hall.

[12] L. Kazmier. 2003. *Schaum's Outline of Business Statistics*. McGraw-Hill.

[13] D. Kingma and M. Welling. 2013. Auto-Encoding Variational Bayes. arXiv preprint arXiv:1312.6114.

[14] C. Kwan, M. Koh, and M. Jasser. 2020. Comparison Study Between Content-Based and Popularity-Based Filtering Via Implementing a Book Recommendation System. *IJARET* 11, 12 (December 2020), 1121–1135.

[15] J. Li, T. Xu, and J. Zhou. 2020. A Personalized Recommendation Algorithm for College Books based on User Interest. In *IEEE CSEI*. IEEE, New York, 144–149.

[16] N. Manouselis, H. Drachsler, K. Verbert, and E. Duval. 2013. *Recommender Systems for Learning*. Springer Briefs in Electrical and Computer Engineering.

[17] S. McGeown, L. Duncan, Y. Griffiths, and Sue. Stothard. 2015. Exploring the Relationship Between Adolescent's Reading Skills, Reading Motivation and Reading Habits. *Reading and Writing* 28 (2015), 514–569.

[18] M. Merga. 2019. *Reading Engagement for Tweens and Teens: What Would Make Them Read More?* ABC-CLIO, Santa Barbara, California.

[19] A. Milton, M. Green, A. Keener, J. Ames, M. Ekstrand, and M. Pera. 2019. StoryTime: Eliciting Preferences from Children for Book Recommendations. In *RecSys)*. ACM, New York, 544–545.

[20] T. Minka. 2013. Expectation Propagation for Approximate Bayesian Inference. arXiv preprint arXiv:1301.2294.

[21] M. Pera. 2009. *Improving Library Searches Using Word-Correlation Factors and Folksonomies*. Master's thesis. BYU.

[22] M. Pera. 2014. *Using Online Data Sources to Make Recommendations on Reading Material for K-12 and Advanced Readers*. Ph.D. Dissertation. BYU.

[23] R. Plutchik. 1997. *Circumplex Models of Personality and Emotions*. American Psychological Association, Washington, D.C., Chapter The Circumplex as a General Model of the Structure of Emotions and Personality, 17–45.

[24] I. Porteous, D. Newman, A. Ihler, A. Asuncion, P. Smyth, and M. Welling. 2008. Fast Collapsed Gibbs Sampling for Latent Dirichlet Allocation. In *SIGKDD*. ACM, New York, 569–577.

[25] K. Reuter. 2007. Assessing Aesthetic Relevance: Children's book Selection in a Digital Library. *JASIST* 58, 12 (2007), 1745–1763.

[26] K. Roskos, J. Christie, and D. Richgels. 2003. The Essentials of Early Literacy Instruction. *Young Children* 58, 2 (2003), 52–60.

[27] L. Rozakis. 2002. *Test Taking Strategies and Study Skills for the Utterly Confused*. McGraw Hill.

[28] H. Singh and A. Jain. 2020. Recommendation of Books in Online Book Store Using KNN. *JAC* XIV, VI (December 2020), 4.

[29] S. Sohail, J. Siddiqui, and R. Ali. 2013. Book Recommendation System Using Opinion Mining Technique. In *ICACCI*. IEEE, 1609–1614.

[30] G. Spache. 1953. A New Readability Formula for Primary-Grade Reading Materials. *Elementary School* 53, 7 (1953), 410–413.

[31] M. Yuan, S. Yang, M. Gu, and H. Gu. 2020. Analysis and Research on Book Recommendation Model Based on Big Data. In *EITCE*. ACM, New York, 21–25.