PalArch's Journal of Archaeology of Egypt / Egyptology

HYBRID BOOK RECOMMENDER SYSTEM USING IMPLICIT FEEDBACK: A MACHINE LEARNING APPROACH

Naveed Jhamat[1], Ghulam Mustafa[2] , Zeeshan Arshad[3], Rizwan Abbas[4]

[1,2] Assistant Professor, Department of Information Technology, University of the Punjab Gujranwala Campus, Gujranwala, 52250, Pakistan.

[3, 4] Lecturer, Department of Information Technology, University of the Punjab Gujranwala Campus, Gujranwala, 52250, Pakistan.

Naveed Jhamat , Ghulam Mustafa , Zeeshan Arshad , Rizwan Abbas , Hybrid Book Recommender System Using Implicit Feedback: A Machine Learning Approach , Palarch's Journal Of Archaeology Of Egypt/Egyptology 18(4). ISSN 1567-214x.

Keywords:Book Recommender System; Implicit Feedback; Machine Learning; Expert System;E-learning.

Abstract:

A recommender system is a common instrument used by businesses to improve customer loyalty and revenues. The two most popular methods when implementing a recommendation system are collaborative filtering and content-based filtering, with the first offering recommendations based on user experience and the second utilizing characteristics of recommendable items.Since the efficiency of each recommendation model is constrained and each has its own strengths and disadvantages in the field of recommender systems, hybrid recommendation models are gaining more interest. The goal of the study was to propose a hybrid book recommendations system with implicit feedback and compares its ability to predict user ratings in an e-book application with basic recommender systems. The models were evaluated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) measures. A hybrid model was built and compared to the pure models by integrating the two basic approaches.In addition, five models, two based on

collaborative filtering, two based on content filtering and one hybrid were created. The results showed that by integrating both methods in a hybrid model with implicit feedback, the lower RMSE and MAE were achieved as compared to the collaborative filter model and content based model on RMSE and MAE measures.

1 Introduction:

A recommendation system (RS) is a method used in the decision-making of a user to recommend items to assist the user. Among companies, RSs are commonly used to increase the number of goods sold or boost customer satisfaction. There are several approaches to RS, the two main ones being Collaborative Filtering (CF) and Content-based Filtering (CN). The advantages and disadvantages of both methods, though the CF approach is seen as the easiest to apply since information on the items is not required. On the other hand, the CN methodology will make recommendations for less customer rating details, although it appears to be over-specialized, so only products that have not yet been rated can be recommended. The downside of the CN method is, however, that products which are not yet rated cannot be recommended. It is the method for RSs most widely used and has its disadvantages because items cannot be recommended until rated[1,2,3].

User feedback in recommendation systems is commonly split into two categories:' explicit' feedback where customers communicate their preferences clearly (e.g. ratings) and 'implicit' feedback where customers implicitly indicate their interest by actions (e.g. clicks). As two independent subjects and various methods have been established to address both of their distinct properties, these two paradigms have long been studied. Further than the limited definitions of explicit and implicit feedback, we note that in many real-world applications, diverse types of user input are plentiful.For example, on e-commerce websites, customer reviews, clicks, orders and ranking scores are readily accessible. All of these signals reflect users' tastes for various types of items. Although there has been a field of study where the relationships between implicit and explicit associations are taken into account[4,5,6], much focus has been put on improving numerical rating predictions by using other signals as auxiliary information.These studies encourage us to cross the line between implicit and explicit indications, but our primary goal is to construct a hybrid system of recommendations for a more general purpose, where multiple types of user input can be viewed simultaneously independently of their particular semantics.

The consumer has associated behavioral data, for instance, ratings of items or number of transactions, while the item has associated metadata as well as material information such as full-text for a book or lyrics for a song[7]. Implicit feedback represents a behavior of a customer, which may be whether a user has consumed an item or finished a book [8]. Implicit feedback could be used independently or used as an extra input to complete a model of sparse explicit feedback [9].

A RS may be utilized for improved performance with implicit ratings. Nichols (1998) explains how explicit ratings are of the disadvantage of not reflecting explicit ratings on the customers' behaviour. Several types of implied ratings can be obtained by the system like Save, Delete, Refer, Read (Time), Consider (Time), etc[10]. The project was conducted by [11] to implement a CF RS news article. They conclude that the time spent on an article was almost as exact as explicit evaluations when predicting ratings were used. The author also states that, rather than simply because their effectiveness and stability is not fully investigated, implied ratings should preferably supplement explicit rating ratings.

A filtering component is used to transform the predicted ratings to recommendations. The filtering component takes the highest ratings and filters out items that are of novel interest to the consumer. This study purposes hybrid recommendation using implicit-feedback approach and compared its performance with the CF and CN approach with and without implicit feedback. The aim of this study is to understand how the addition of implicit feedback into hybrid recommender system affects the performance of the RS. The paper seeks to find how both CF and CN approaches can be mutually complementary and if a hybrid model can exceed both basic models on books dataset.

The paper shall be formulated as follows: In Section 2, we shall introduce the related work. The dataset, evaluation measures and elements of our proposed system will be defined in Section 3. Section 4 discusses the results and discussion and, finally, we summarise and conclude future work in Section 5.

2 Literature Review:

The Collaborative Filtering (CF) approach gives recommendations based on behavioral data patterns. It is based on the idea that if their behavior is similar, two users are similar. Items which were consumed or liked by similar users will be recommended to users. CF models are all based on either explicit or implicit patterns of feedback that may be used to recommend. The approach is seen as the simplest, as it needs user data only as an input and does not require knowledge of the items. The CF approach models are divided primarily into two categories: neighborhood and matrix factorization model [9].The CF method typically relates to a U–I matrix, which tracks each user rating on a single item and compares the total number of entries in this matrix with each item in the matrix. The number of entries observed is the number of actual ratings that the user has recorded, relevant for measurement because this can influence the model's performance and calculation needs.

Neighborhood models based on input from consumers of items are based on the recognition of the associations between users or items. The user-user resemblance and interplay between items are called items–item similarity calculation. The user methodology creates instructions based on ratings from related users. The item-item approach has recently gained popularity, as it scales faster and is more precise than the consumer approach [12]. Cosine similarity and the Pearson correlation are the two most common similarity measures for CF RSs. For item-similarity too, all similarities and neighborhood models are compatible. The [13] has shown that the calculation used usually does not change much. Even if cosine similarities and Pearson correlations were successfully used in several studies, [14]says both can have a problem when the data is very scarce.

Additionally, a neighborhood model is the K-Nearest Neighbor (KNN), a model which calculates the prediction by weighing an average of k values for the nearest neighbourhood. This KNN version considers the average user rating and uses a z-score normalization for each user for user-user similarity. Addressing user mean and standard deviations can aid the model in taking general user behavior. For example, if two users, $u_1$andu_2 , are the same average, but the ratings of each are different, their rating patterns can still be different. The model can also be used to capture other user patterns, for example, if a user is typically high or low or if he or she is a user who rate a lot or not, to get a more accurate forecast of how a user is rate.

In RS, matrix factorization models were increasingly used. They are high in efficiency and good scalability for sparse data. SVD is a matrix factorization model that identifies latent characteristics by mapping users and items within a common latent dimensional space factor. In order to avoid the sparse U–I matrix issues and to be able to scale RS problems effectively, this dimensional reduction approach. For example, with around 109 entries in the U-I matrix, [15] underlines the sparsity and dimension of the Netflix data set, but only 100,480,507 entries observed, implies an estimated sparsity of 1 percent. The author also talks about the data set imbalance with some users less than 5 times and some more than 10,000 times.

Conventional SVD requires U-I matrix factorization. The conventional SVD performance is an argument of [9]. The conventional approach for missing values is not defined, and it typically leads to overfitting only the few existing values. Instead, further work only suggests that explicit assessments are used to avoid over-replicating the data and that the missing values are not imputed and a regularization component is included. Furthermore, the advantages of using the user prejudices and item choice in the matrix factorization approximation are that a substantial difference of the results will always be clarified by clear reasons of the user or particular items. [15]. In the topic of Netflix, for example, some films are always high regardless of the type of user or some users always score a movie as low, regardless of movie.

Moreover, Content-based Filtering attempts to suggest an item to a user that is close to the previous items the user preferred. The resemblance is extracted from the content of the items, such as the title or the full-text of a book. The pre-processed item information is split into a training and test set, where the training data is used as feedback to the profile learner. The test set is sent to the filtering component directly, so that certain items are not displayed to the learner and hence unknown to them. The filtering portion uses the user accounts, all items processed by the profile learner and any items that possibly have not been processed to generate a list of recommendations.This is achieved by an assessment of the similarities of the related items in the test data and a review of the user preferences provided by the users' profiles to get the best recommendations. The information analyzer conducts the data preprocessing step that takes the item content data and describes items numerically for further analysis.

Content-based filtering is based on the premise that if a person has demonstrated likes to a certain item with certain attributes, it would presumably like other items with the same attributes [16].These programs analyze the characteristics of the items a person has dealt with and introduces candidate items with identical attributes. For example, once a consumer has viewed movies classified of a specific genre, the recommender mechanism will show more movies in the same genre.The two key approaches when designing content-driven filtering systems are to either compare items with items separately, or assign each user a taste profile based on their experience of choice [16], from which suggestions are extracted. Content-based filtering is widely used in systems dealing with text-based items, including news stories and documents.The text is retrieved from the item using text retrieval methods and, by using keyword analysis techniques, the item is assigned those attributes [17].

The greatest benefit of the CF approach is that no content data or techniques are available for the analysis of content. The disadvantages are the primary issue and the tendency to over-specialize. The CN approach recommends items based on the similarity of the content to the items rated by the user. The main advantages with the CN approach are that no new items require users' rating data, and the content does not rely on ratings. The disadvantages are the problem of cold starts, the tendency to recommend items too similar to those already read, and the problem of serendipity.

Nevertheless, hybrid RSs are models based on two or more techniques combined. The aim of hybrid models is to exploit the advantages of each technique. In a hybrid model, there are several strategies for incorporating both approaches. Hybrid RSs, which are weighted, mixed, switched and feature combined provide the same results regardless of the CF or CN approach order[18]. The cascade, the feature increase and meta level are associated with one approach as input to another. Different results could occur depending on the approach used as input. The literature has successfully implemented a hybrid RS[18].

The key idea is to change between what model to use when making a recommendation. The switching method uses a number of rules and determines which models should be employed. By using a weighted hybrid model, [18] managed to achieve a lower MAE and RMSE, which included both predictions. These rules can be set manually, by requesting experts, visual analysis or even classification problems. These rules can be defined manually. We use these methods together to create mixed algorithm recommendations.

3 Hybrid Recommender System with Implicit:

3.1Dataset:

From the book rating website Goodreads, we used a new large-scale dataset. This information contains 229,154,523 documents from 876,145 public book shelves and 2360,655 books (with comprehensive meta-data including authors, series, editions, publishers, numbers of pages, languages of book contents, related books and top user-generated shelf names for these books). Information of multiple user contacts on the item, including date, reading progression, rating score, and summary text, if applicable, are given in each log. The data also revealed that everybody reads the book indicates the user's implicit interest in the book.

Data was anonymized in this analysis, which means no data can be associated with other data. However, it can be difficult to anonymize all data in cases where applied RSs may not include the information, and particularly where multiple implicit information is used. The author wants to make it clear that the goal of the RS is not just to manipulate consumers into any actions, but also to enable them to guide customers into the broadest possible information and not merely to lead customers into the best-selling books. The intention is to improve consumer personalization.

3.2 Evaluation of Recommender Systems:

There are many ways to test a RS. One common method used in many research papers is to calculate the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) of predictions on ratings. MAE computes the mean of the absolute error, for each prediction and actual value, see Equation 1, where $\widehat{y_i}$ are the predictions and y_i are the actual values.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |\widehat{y_i} - y_i|}{n} \quad \text{………………….}(1)$$

RMSE calculates the standard deviation of the residuals, which is a statistic that defines how far the predictions $\widehat{y_i}$ are from the true values y_i in the same manner as MAE, but taxesprediction further away from the true values more than MAE.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

[19] note that both MAE and RMSE are helpful when calculating the performance of a RS. MAE tests how effective the RS is, a lower MAE means better predictions. RMSE instead imply the reliability of the predictions and a low RMSE means low uncertainty of the predictions, since RMSE penalizes greater variations compared to MAE. Furthermore, the models predict the ratings and no recommendations are made, the numerical evaluation of the predictions was delimited to RMSE and MAE.

3.3 Our Propose Hybrid Recommender System using Implicit-Feedback (HI)

There have been two baseline versions. One for the CF, one for the CN. Furthermore, one basic hybrid model was also build. The foundation for basic models was a model that for books in scales 1-5 often estimated the overall mean ranking. The baseline prediction may be useful for analysis, since it captures much of the trends found in the data and the distribution of user rate information focuses into higher ratings. The basic hybrid predicts always the total mean of the predictions of CF and of the predictions made jointly by the chosen CN model. The models were chosen to assess item-item similarity because they are typically more powerful than user-user and have reasonable scalability. All other parameters and any hyper-parameters set to default.

For the K-Nearest Neighbor (KNN) Baseline algorithm additional parameters are needed when a baseline estimator is used. Before the model selection trial, Stochastic Gradient Descent (SGD) and Alternating Least Squares (ALS) were both evaluated using default parameters to determine the appropriate optimization technique. It was noticed that SGD was quicker and more precise, however for the incredibly sparse data ALS could do best. Through searching the grid over a range of parameters, parameters for the chosen models were optimized. The SVD++ matrix factorization model was also chosen for further experimentation. A new, narrower grid based on the lowest RMSE performing parameters was generated from the random scan. In order to further decide whether k = 250 was sufficient or the number of neighbours, k was tested. Without any significant effect on RMSE, neighbors may be lower. The results tend to converge when k = 90, and for final SVD and SVD++ models the rest of the parameters that resulted in the lowest RMSE in grid check are chosen.

These models were used to make validation predictions so that they can be compared with the other models. For the lowest score model, the parameters of the low RMSE have been chosen and the final grid for the data set was created. Each CF and CN model has been trained on the validation data predictions. The goal variable was a binary variable set to 0 when a better prediction was made by the CN model and to 1 otherwise.

Furthermore, for CN component, the model Doc2vec was used to convert all the books in 100-dimensional vectors. User profiles have been learned using vector data in conjunction with usage data. The profiles were learned and analysed using three different regressive techniques: linear regression, regression trees and KNN. The RMSE-based best output model was chosen as the final model for the test results. The CN model was used to render user profile predictions. The text was then tokenized and translated into a lower case. A download of the summary text of each book was done in order to make the retrieval practical due to a huge number of books. The summary text of every book was marked with one ID, with the title and authors' ID, to train the

Doc2vec model. The other tokens have been lemmatized. For 20 epochs, the Doc2 Vec model was trained and a 100D Vector representation for each book was created. In order to achieve a visible representation, the model was developed using the PV-DM version of Doc2vec, as it is sufficient for most tasks according to [20].

To predict ratings, Scikit-learn models have been used to learn the user profile for each user. The methods were linear regression, tree and KNN, all of which were tested. The choice of models was based on the easy interpretation and the fact that large amounts of data are not required. A new model was developed for every single user in the dataset to produce a single model for each user. As the technique is more interpretable than other techniques, and the measuring time between techniques was also lowest, PCA was chosen to serve for input into the final CN model.

4 Results and Discussion:

The findings of the experiments suggest that the basic CF model has less RMSE than the basic CN model. The MAE and RME of both basic versions are approximately equal to the CF and CN with implicit feedback. There are lower values for our hybrid model with implicit feedback than the basic hybrid. CF with implicit feedback performed better as compared to all other models except the hybrid model with implicit feedback. Table 1 showed the overall performance of all models using RMSE and MAE measures.

Table 1: RMSE and MAE values for all models including our hybrid model with implicit feedback

Model	RMSE	MAE
Basic CF	1.2749	1.1230
Basic CN	1.3237	1.0423
CF with implicit feedback	1.2733	1.1289
CN with implicit feedback	1.3416	1.0407
Basic Hybrid	1.3241	1.1479
Hybrid with implicit feedback	1.2611	0.9933

For each ranking (1-5) the Basic CF, Basic CN, CF with implicit feedback, CN with implicit feedback, Basic Hybrid and Hybrid with implicit feedback models were assessed. (see Table 2 for the RMSE and MAE results for all tested models)

The lower (1 and 2) scores are now the most difficult to estimate. An incredibly sparse U-I matrix was generated by the user rating dataset used as the input in various models. Both models were retrained and tested with the dataset to assess how data sparsity influences the performance.

Table 2: Hybrid RS with implicit feedback(HI) performance against basic CF, basic CN, CF with implicit feedback (CFI), CN with implicit feedback(CNI) and basic hybrid model (BH) on RMSE and MAE measures.

Subset	RMSE						MAE					
	BCF	BCN	CFI	CNI	BH	HI	BCF	BCN	CFI	CNI	BH	HI
Rating 1	2.4213	2.2815	2.4060	2.2747	2.4819	2.3638	2.2676	1.8310	2.2540	1.8249	2.2892	2.0802
Rating 2	1.5759	1.5184	1.5687	1.5127	1.6356	1.5578	1.4166	1.3532	1.4029	1.3475	1.4467	1.3778
Rating 3	0.8593	0.9892	0.8565	0.9865	0.9013	0.8585	0.7252	0.7957	0.7203	0.7931	0.7550	0.7190
Rating 4	0.5267	0.7567	0.5376	0.7578	0.5678	0.5309	0.4036	0.5523	0.4111	0.5535	0.4323	0.4117
Rating 5	0.9913	1.0726	0.9979	1.0763	1.0336	0.9845	0.8021	0.6792	0.8561	0.6845	0.8164	0.7775

Fewer data points could make it harder for the models to understand any user's behavior. The findings which vary according to the dataset, especially in terms of data sparsity and the amount of ratings a user has rated. For example, if you have a new item, a user, a new data point may increase the sparsity of your U–I matrix. The RMSE and MAE are diminished as a denser data is generated and the results of these studies are equivalent to other datasets. In some cases, the results will also depend on the density of the U-I matrix resulting. In some cases, the results may be different.

The findings of both basic models demonstrate that the CN model average has a lower RMSE than the CF model and better scores than those of other models in other tests.The CF model has a longer time to estimate values for certain data points that are well outside the mean of the dataset. In the other hand, the CN model is able to predict book ratings correctly. Residual analysis conducted on predictions suggests that, when the residues are zero, the CN model tends to estimate accurately more often than CF. A hypothesis is that the larger the ratings a user has, the better the model would be, since the user profiles will be better learnt and the profile can be accurately predicted by the user.

The results show that there are multiple categories of consumers who are different as scores are ingested. It also shows that the variations in the ratings provided by a user, the variation between the ratings of a user and the mean of a rating, and the amount and mean of a person's ratings for books ratings. The average rate is three ratings per client and the majority of customers rate the minimum rating level.

In the prediction distributions, the CN model and the CF model have different variances. The efficiency of various models for users with varying variances is interesting. Where the difference is between 4.8–6.4, the CN model is higher than the CF model. A high variance may mean that the consumer is involved and opinionated. Different user identities may be difficult to recognize and therefore difficult to validate. It is unclear whether the discrepancy or number of ratings influence the model efficiency.

It is impossible to know the explanation for a high variance. When a user has only scored many times, the deviation is smaller than a user with more ratings. It may also be that, in fact, the user consists of multiple account sharing users and has different views. In contrast to the CF-Model, it appears that the CN model is able to catch variations of ratings far from the mean. For users with smaller variances of the CN model, MAE and RMSE can be found to be lower.

For user classes with various mean scores, RMSE and MAE were measured. For users with an excessive mean ranking, the CN model has a lower score than the CF model on MAE. For users with an average ranking closer to the dataset mean, the CF model is higher, on the contrary. However, it is impossible to understand the precise explanation that some groups are tougher than others to predict. This may be because the book scores are not enough to allow the contrast between the two versions.

An object called the first-rater problem cannot be recommended by a CF RS. A sub-set of test data was evaluated, which only contained new items. The CF model already has the best RMSE and the CN model has the most MAE on the latest data points.

The item bias and factor will be set to zero in the SVD implementation. This means that the predictions are based on the mean of all ratings and the consumer preference. The RMSE rise is the lowest for the book. In comparison to the overall results, the number of tacit ratings available was very limited. The overall RMSE and MAE were affected adversely and raised, as most of the scores were either four or five. The inferred ratings cannot be considered as useful in the exact use of predicting ratings.

The observations taken when the tacit scores were processed may have been incorrect. The reason the findings made the results worse was that the conclusions that were processed were made.

Any user who does not finish a book other that he does not like is for other potential motives. The ranking labeling provided by a consumer may be inaccurate depending on the completion rate. It should always be treated as ranking 1 regardless of how many percentages the consumer reads in the book.

The CN model tends to be doing better when the rating variance between 4.8 and 6.4 is provided by the customer. The gap and the mean of the past ratings of users is necessary to analyze whether CN or CF predictions are to be used. The CN model does well for those consumers with lower score deviation than for those with higher true ratings.

The technique and application of the basic CF model, the CN pattern and the hybrid patterns are discussed in this section. RMSE and MAE have been used to assess the samples, but the accuracy of the recommendations cannot always be mirrored.

Additionally, several tests were replicated to ensure the outcome was not compromised by random elements. When training models containing random components such as SVD model, no seeds were set. The data were re-sampled three times with approximately the same result, suggesting that the findings were accurate.It was best if the tests be repeated more than three times to produce larger performance, but it was not viable due to a lack of time. The findings revealed that designers with a lower RMSE were more prone to the 5-star recommendation being placed in the top 20.100 randomly chosen books and a book of true scores of five were introduced to the CF and CN models. The lowest RMSE model was more likely to be ranked in the top 20 of all 100 books.

For the most fitting models, the hyper-parameter optimization was performed. This segment addresses the application of the CN model. The pre-processing performed in this analysis (see section 2.3.1) was mostly targeted at cleaning up terms that do not provide information. The text representation processing and profile learning will also be addressed. During the experiments, some restrictions were made and studies were limited due to limitations.

On the book data, the selection of the model and the tuning for the hyper-parameters should have been performed. Due to the somewhat different distributions, a new hyper Parametertuning may be claimed for better outcomes.

That would possibly collect more book material than the model used in the report, which uses all word groups and less on the use of vocabulary. By comparing the gap between RMSE and MAE for various Doc2vec versions, users could better understand their ranking based on it. Owing to the time limit, the decision to only use a sub-set was made. A deep dive into stop terms for the entire dataset may have helped to decrease noise in the entire text. If the postholistic stop terms are the alternative in comparison to the whole dataset, it is unclear.

The material was defined by 100-dimensional vectors and PCA was reduced to two dimensions. All books were designed and colored by group PCA and UMAP co-ordinates. Figures demonstrate that there are clusters of groups, especially in Figure 5.17, in which UMAP is used for reducing dimensionality. Books in the genre of "Classic" do not all have the same subject, but they may use comparable vocabulary and they are mostly older than most recent books. This reveals that the Doc2vec model often includes the use of terms to describe various themes and names of the different books.

The material experts may explore the resulting clusters with reduced embedding measurements to decide whether derived similarities seem to be sufficient.

A difference between fiction and non-fiction appears to be made with the Doc2vec model. It is necessary to make the same experiment with other methods of document representation by vector representation of the text. Term frequency-inverse document frequency (TF-IDF) or Latent Dirichlet allocation (LDA) should be checked to decide if it could lead to higher embedding efficiency.

The sentence BERT (S-BERT) edition (Reimers et al. 2019) of BERT can also be helpful to try out how it can be used in Sentence Sentence-BERT. It may have created some noise in the model because some books may have a text written in a different language than English. The model does not work in several languages. The two architectures Doc-2vec: PV-DBOW and PV-DM can be mixed to best represent text.

This could not have been the case with regression trees or linear regression because they could only gain from a single dimension. It can also be useful to look at different proportions, but also to explore different reducing methods with different sizes. In this area, a detailed test may have led to better CN model predictions. The findings of this analysis may have contributed to a clearer forecast of the process projections for the CN model. The results could be used to boost the estimation accuracy of CN models for regression and linear regression.

The experiment may also benefit from the analysis of various dimensional technologies to minimize regression trees efficiency. As KNN can accommodate greater data dimensions than linear regression and regression of the arboreal. The data with implied ratings included only a particular category of rating. The root of these ratings was just about a year old, resulting in a few tacit ratings. The aim was to make the data collection more equal, with knowledge while training a model still can contribute to performance (4 and 5) while it makes the dataset more unevenly

For instance, details on whether a consumer books a book (saved it later) or has seen a book will boost the findings and give more detail. If these acts are to be converted into evaluations, they are likely to be viewed as strongly optimistic (3-5)

Instead of grid search the matrix factorization was used by Lin et al. (2014), but the same outcome is predicted. The optimum weighting factor was chosen, taking into consideration only the optimum RMSE. In order to improve the efficiency of the switching model more information should have been obtained into book RSs behavior. It will possibly only be possible to test out more functionality such as an average book length of a user's read book and do a more detailed review. Improving the way model rules are calculated can contribute to a significant performance improvement.

In the literature, there are several hybrids for feature increase that incorporate several approaches. The most effective CF model, SVD, was expected to be the right model for the hybrid for this experiment. It is of paramount importance for the program to be straightforward and transparent to the consumer using the applied RS. The issue of openness in an RS is important, not only to gain consumer interest, but also for the consistency of the RS for ethical reasons. It does not influence a user's privacy to store and use an overt rating while the implied rating may be in problem when addressing it.

5. CONCLUSION AND FUTURE WORK:

In this study, we introduce a new hybrid recommendation model with implicit feedback for books, consisting of two high performing CF model and CN model. We use the book reading to gather auxiliary information instead of only modelling user-item experiences as previous works do. A book reading implicit feedback is used to capture the trends of the consumer

The analysis showed that RMSE = 1.2749and MAE = 1.1230were accomplished by using basic CF methodology, opposed to pure CN approaches, which achieved RMSE = 1.3237and MAE = 1.0423. It may be argued that CF is higher than the MAE model of the CN, while CF model have the betterRMSE.In general, the CF was more reliable because it had less significant errors than the CN model. However, the CN method was more effective, since a more reliable assessment was more widely expected. The methods for various types of ratings were carried out differently. In the prediction of ratings 3 and 4, CF outperformed CN, while CN became the best model for ranking 1 and 2. It was also found that when estimating ratings for new items, the output of CN was less affected than CF.Finally, the book-RS performance was improved with the hybrid approach. The implicitly feed-backed hybrid model was the most popular RMSE (1.2611) and MAE (0.9933). The findings show that the combination of CF and CN approach in a hybrid model, in which the strengths of one approach balance each other's drawbacks, enhances performance.

The different models provided predictions on ratings, and no recommendations were derived from these predictions. Future work can possibly explore these different approaches, by deriving recommendations and putting them into production. There are several ways of improving the RS that was implemented in this study. A/B testing to see which of the models are performing best would add more validity to the study and enable new discoveries in userbehavior. A mixed hybrid approach can also be explored if one were to put production. Improvements and fine-tuning of the RS implemented could improve the performance of the ratings. The research was restricted to e-books only. It is helpful to also research the audio book consumers to see how the behaviour is different and to collect more details.Any time-aware models may have been discussed, for example. One may contend that the book describing the best of an author is the book written by that author most recently.

Acknowledgement:

We are thankful to Department of Information Technology, University of the Punjab Gujranwala Campus, Pakistan, for providing technical support and computational resources.

Conflicts of Interest:

The authors declare that they have no conflicts of interest to report regarding the present study.

References:

[1] Gazdar, A., & Hidri, L. (2020). A new similarity measure for collaborative filtering based recommender systems. Knowledge-Based Systems, 188, 105058
[2] Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., & He, Q. (2020). A survey on knowledge graph-based recommender systems. IEEE Transactions on Knowledge and Data Engineering.
[3] Sun, Z., Zhang, X., Li, H., Xiao, Y., & Guo, H. (2021). Recommender Systems Based on Tensor Decomposition. CMC-COMPUTERS MATERIALS & CONTINUA, 66(1), 621-630.
[4] Jawaheer, G., Szomszor, M., & Kostkova, P. (2010, September). Comparison of implicit and explicit feedback from an online music recommendation service. In proceedings of the 1st international workshop on information heterogeneity and fusion in recommender systems (pp. 47-51).
[5] Jawaheer, G., Weller, P., & Kostkova, P. (2014). Modeling user preferences in recommender systems: A classification framework for explicit and implicit user feedback. ACM Transactions on Interactive Intelligent Systems (TiiS), 4(2), 1-26.
[6] Koren, Y. (2010). Factor in the neighbors: Scalable and accurate collaborative filtering. ACM Transactions on Knowledge Discovery from Data (TKDD), 4(1), 1-24.
[7] Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In Recommender systems handbook (pp. 1-35). Springer, Boston, MA.
[8] Schafer, J. Ben, Dan Frankowski, Jon Herlocker, and Shilad Sen (2007). “Collaborative Filtering
[9] Recommender Systems”. In: LNCS 4321 - The Adaptive Web .
[10] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30-37.
[11] Nichols, D. (1998). Implicit rating and filtering. ERCIM.
[12] Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). Grouplens: Applying collaborative filtering to usenet news. Communications of the ACM, 40(3), 77-87.
[13] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001, April). Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th international conference on World Wide Web (pp. 285-295).

[14] Lathia, N., Hailes, S., & Capra, L. (2008, March). The effect of correlation coefficients on communities of recommenders. In Proceedings of the 2008 ACM symposium on Applied computing (pp. 2000-2005).
[15] Ahn, H. J. (2008). A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. Information sciences, 178(1), 37-51.
[16] Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press.
[17] Iaquinta, L., De Gemmis, M., Lops, P., Semeraro, G., Filannino, M., & Molino, P. (2008, September). Introducing serendipity in a content-based recommender system. In 2008 Eighth International Conference on Hybrid Intelligent Systems (pp. 168-173). IEEE.
[18] Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12(4), 331-370.
[19] Lin, W., Li, Y., Feng, S., & Wang, Y. (2014, June). The optimization of weights in weighted hybrid recommendation algorithm. In 2014 IEEE/ACIS 13th International Conference on Computer and Information Science (ICIS) (pp. 415-418). IEEE.
[20] Balasingam, U., & Palaniswamy, G. (2019). Feature Augmentation based Hybrid Collaborative Filtering using Tree Boosted Ensemble. Informatica, 43(4).
[21] Le, Q., & Mikolov, T. (2014, June). Distributed representations of sentences and documents. In International conference on machine learning (pp. 1188-1196). PMLR.