# Alleviating the cold-start problem by using demographic data and domain-aware similarity measure

**Rahul Crunal Kalaria**
**Supervisor(s): Frans Oliehoek, Aleksander Czechowski, Oussama Azizi, Davide Mambelli**
**EEMCS, Delft University of Technology, The Netherlands**
20-6-2022

**A Dissertation Submitted to EEMCS faculty Delft University of Technology,**
**In Partial Fulfilment of the Requirements**
**For the Bachelor of Computer Science and Engineering**

## Abstract

Recommender systems (RS) are a cornerstone for most online businesses that cater to a large customer base such as e-commerce, social network platforms and many others. RS's enable these platforms to provide tailor-made experiences to each of their customers by strategically utilizing users/items rating data or any other available data. Collaborative filtering (CF) techniques are some of the most popular and successful RS models created. However, CF techniques often suffer from the cold start (CS) problem. In particular, they struggle with complete cold start (CCS) situations in which no user/item rating history is available and incomplete cold start (ICS) situations in which only a limited amount of user/item rating history is available. In this paper, we explore two models which utilize novel ideas to combat the CCS and ICS problems. The first model (DCF) focuses on the intelligent use of user demographic data to combat the CCS problem. The second model (PIPCF) focuses on the use of a novel domain-specific similarity measure called Proximity-Impact-Popularity (PIP) to combat the ICS problem. In addition to this, we also propose our own model (DPIP-CF) which combines these two ideas in conjunction with some of our own modifications to combat the CCS and ICS problems simultaneously. We utilize the MovieLens data set which is a commonly available and popular dataset that is often used to test RS's. Through a series of experiments, we demonstrate the strengths of DCF and PIPCF in dealing with the CCS and ICS problems respectively. Finally, we also show that our DPIP-CF model outperforms all other models discussed in this paper and is a viable solution to dealing with the CCS and ICS problems simultaneously.

# 1 Introduction

With the ever-growing reliance on E-commerce for shopping needs, social networks for facilitating communication and media platforms for entertainment, automated product recommendation from RS's has become a vital part of any online business. Recommender systems have given online businesses the unique ability to create tailor-made experiences for their users allowing them to reap the benefits of increased revenue and user satisfaction [2]. Additionally from the user's perspective, RS's help to reduce informational overload by only recommending products that are most relevant to each user [7; 16]

Over the years various researchers have tried to classify these RS's into several categories. [13] proposed one such categorization in which RS's are split into three types; content-based [9], collaborative-filtering [19; 14], and demographic-based [15; 8]. Content-based RS's cleverly compare the descriptions of items to be recommended to the user with the items that the user has already rated. Collaborative-filtering RS's intelligently infer correlations between various users based on the reviews submitted by these users. Demographic-based RS's exploit user data such as age, sex and location to identify similar types of users. Additionally, some researchers like added to this categorization by introducing Hybrid RS's which are a combination of two or more recommendation techniques [5].

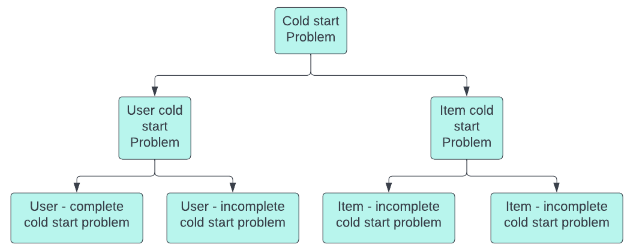


Figure 1: The breakdown of the cold-start problem into user and item cold-start problems which are further broken down into complete cold-start (CCS) and incomplete cold-start (ICS) problems

The cold start problem is one of the most studied problems in the field of recommender systems. The cold start problem can be divided into two types, item cold-start or user cold-start [1]. The item cold-start problem [17] is caused when a new item is added to the pre-existing catalogue of items. Recommender systems struggle to recommend such new items due to insufficient user reviews [4]. On the other hand, the cold-start user problem [21; 18] is caused when a new user enters a system. Recommender systems struggle to make meaningful recommendations to these new users due to the lack of reviews made by them. With such users, it is often the case that the recommender systems can not compute the similarity between users as there is no intersection between the interests of the new user and other users. Due to the user cold-start problem, recommender systems fail to make high-quality recommendations to new users. Additionally, the cold-start problem can further be broken down into the complete cold-start (CCS) and the incomplete cold-start (ICS) problems [22]. The CCS problem refers to situations in which no rating information is available for items or users. The ICS problem refers to situations in which a limited amount of rating information is available for items or users. An overview of the cold-start problem is presented in figure 1. This paper focuses on the user side complete and incomplete cold-start problems.

Our contributions are listed as follows:

1. Study the use of user demographic data to enhance user-profiles and alleviate the CCS problem.
2. Study the use of a domain-aware similarity measure over traditional similarity measures to solve the ICS problem.
3. Propose a new model that is capable of showing improved performance for CCS and ICS problems simulta-

neously by using user demographic data and a domain-aware similarity measure.

4. Experiments are conducted using MovieLens dataset to showcase the strengths and weaknesses of all implemented models. From these experiments, it is concluded that our proposed solution is able to show improved performance for CCS and ICS situations.

This paper aims to study the effectiveness of user demographic data in solving the CCS problem, using domain-aware similarity measures to solve the ICS problems and provide a potential solution to solving the CCS and ICS problems simultaneously by the combined use of demographic data and domain-aware similarity measures. In particular, this paper aims to answer the following research question, *'Can the combined use of user demographic data and a domain-aware similarity measure solve the CCS and ICS problems in collaborative filtering based recommender systems?'*.

The rest of the paper is structured as follows. Section 2 briefly explains some necessary basic concepts used in our research of recommender systems. Section 3 provides an overview of all the models that were studied and implemented within this study. Section 4 first describes the experimental setup and later presents the results of each experiment which are then discussed in section 5. Section 6 discusses the responsible research practices included within this research. Finally section 7 provides some concluding remarks and future extensions.

### 1.1 Related work

To address the CCS problem in CF models, some researchers have found the use of user demographic data to be very effective [10]. The underlying assumption of such demographic-based models is that users that share similar demographics (such as age, sex, and occupation) tend to have more similar preferences than users with different demographic attributes.
[12] proposed the hybrid use of a knowledge-based model and a collaborative filtering model to deal with the CCS problem. In their proposed solution, when the system detects a CCS user, their model switches to using the knowledge-based algorithm. At all other times, their model uses the collaborative filtering algorithm. In [20] the authors argue that similarity measures should not only be confined to user ratings. Their model calculates the similarity between users based on their social Linked Open Data features.

To address the ICS problem, [2] proposed the use of a completely new domain-aware similarity measure. Through a series of experiments they showed that traditional similarity measures fail to properly utilize the domain-specific meaning behind the rating data, hence they have limited use in CF methods. In light of this, they proposed their own similarity measure Proximity-Impact-Popularity (PIP). In a similar fashion, [11] proposed their own modification of PIP called, Proximity-Significance-Singularity which is designed to be an improvement over traditional similarity measures such as Pearson correlation coefficient and cosine similarity.

## 2 Background

This section briefly explains some necessary concepts used in our research of recommender systems. First the general idea and concept of user-based collaborative filtering recommender systems is provided. Then a description of a novel domain-aware similarity measure is provided.

### 2.1 User-based Collaborative filtering

Generally, user-based collaborative filtering systems contain two distinct phases. The first phase focuses on the creation of the neighbourhood set. This set is a group of users who are most similar to our target user. Such a set is computed using a similarity metric. In the second phase, ratings of the target user are predicted for various new items. Finally, the items with the highest predicted ratings are recommended to the target user.

The main objective of the neighbourhood set is to identify users who are most similar or dissimilar to our target user. If there are $m$ users, then these users' similarity or dissimilarity is calculated using a similarity measure and stored in an $m \times m$ user-user similarity matrix. In such a matrix the entry located in cell $i, j$ denotes the similarity score between users $i$ and $j$. Various similarity measures have been used in recommender systems such as cosine similarity and Pearson correlation. In this paper, we use cosine similarity which is defined as follows

$$Sim(u_i, u_j) = \frac{\vec{u_i} \cdot \vec{u_j}}{||\vec{u_i}|| \, ||\vec{u_j}||} \tag{1}$$

Having made the $m \times m$ user-user similarity matrix, the predicted ratings of the target user for unseen items can now be calculated. The idea is that users with a higher similarity score should influence the predicated rating more than users with lower scores. The predicted rating for user $u$ and item $j$ can be calculated using the following formulation:

$$pred(u, j) = \bar{r_u} + \frac{\sum_{v \in \eta}(r_{v,j} - \bar{r_v}) \times sim(u, v)}{\sum_{v \in \eta} |sim(u, v)|} \tag{2}$$

Where $\bar{r_x}$ is the mean rating of user $x$, $r_{x,j}$ is the rating of item $j$ given by user $x$ and $\eta$ is the neighbourhood set

### 2.2 Proximity-impact-popularity (PIP) similarity measure

The PIP similarity measure is domain-aware similarity measure proposed by [2]. PIP can be broken down into three parts, Proximity, Impact and Popularity.

Two of the three factors use a notion of *agreement* between two ratings $r_1$ and $r_2$ which is defined as follows:
If $if(r_1 > R_{med}\ and\ r_2 < R_{med})\ or\ (r_1 < R_{med}\ and\ r_2 > R_{med})$

$$Agreement(r1, r2) = \textbf{false}$$

Otherwise,

$$Agreement(r1, r2) = \textbf{true}$$

Where $R_{max}$ and $R_{min}$ are the the maximum and minimum possible ratings of the rating scale and $R_{med} = (R_{max} + R_{min})/2$

Proximity calculates the absolute difference between two ratings but also applies a penalty to ratings which are in disagreement. Proximity is calculated as follows:

If *Agreement($r_1$, $r_2$) =* ***true***

$$D(r_1, r_2) = |r_1 - r_2|$$

If *Agreement($r_1$, $r_2$) =* ***false***

$$D(r_1, r_2) = 2 \times |r_1 - r_2|$$

$$Proximity(r_1, r_2) = ((2\times(R_{max} - R_{min}) + 1) - D(r_1, r_2))^2$$

Impact considers how strongly an item is liked or disliked by users. When such a clear preference has been expressed greater credibility can be given to the similarity between the two users. Impact is defined as follows:

If *Agreement($r_1$, $r_2$) =* ***true***

$$Impact(r_1, r_2) = (|r_1 - R_{med}| + 1)(|r_2 - R_{med}| + 1)$$

If *Agreement($r_1$, $r_2$) =* ***false***

$$Impact(r_1, r_2) = \frac{1}{(|r_1 - R_{med}| + 1)(|r_2 - R_{med}| + 1)}$$

The Popularity factor gives greater weight to ratings which are further away from the average rating for an item. If two users have similar ratings for an item but these ratings are far away from the average rating of the item, then this signals that the users must have a high degree of similarity. Hence the popularity component rewards ratings which are close to each other while being distant from the average rating.The *Popularity* can be calculated as follows:

If $(r_1 > \mu_k \; and \; r_2 > \mu_k)$ or $(r_1 < \mu_k \; and \; r_2 < \mu_k)$ then,

$$Popularity(r_1, r_2) = 1 + (\frac{r_1 + r_2}{2} - \mu_k)^2$$

else

$$Popularity(r_1, r_2) = 1$$

where $\mu_k$ is the average rating for item $k$ by all users.

Having defined *Proximity, impact* and *Popularity*, the similarity between two users $u_i$ and $u_j$ can be calculated as:

$$SIM(u_i, u_j) = \sum_{k \in C_{i,j}} PIP(r_{ik}, r_{jk}) \tag{3}$$

where $C_{i,j}$ is the set of co-rated items of the users and $r_{ik}$ and $r_{ij}$ are the ratings given by the users for item $k$ and $PIP(r_{ik}, r_{jk})$ being defined by the three components as follows:

$$PIP(r_1, r_2) = Proximity(r_1, r_2) \times Impact(r_1, r_2) \times$$

$$Popularity(r_1, r_2)$$

# 3 Methodology

This research aims to investigate the effectiveness of enriching user profiles by the use of demographic data to solve the CCS problem, the use of a domain-aware similarity measure to solve the ICS problem and finally the combined use of demographic data and domain-aware similarity measure to solve the CCS and ICS problems simultaneously. This research was conducted in a four-step ablation in which various cold-start (CS) mitigation strategies were iteratively added to a baseline model. Through this type of study, the effectiveness of each CS mitigation strategy can be studied in isolation. As a result, this aids to understand the contribution of each strategy to the overall system. In the first step a baseline CF model (Base-CF) was built. Using this baseline, two more models were built that specifically targeted the CCS and ICS problems respectively. The final model was built by combining the CCS and ICS specific models.

## 3.1 The baseline CF model (B-CF)

The first step of the study was to build the baseline CF system to serve was a point of reference for all other models. Having the other models built upon this baseline helped to empirically observe and understand the effects of each modification on the whole CF model. An overview of baseline RS is shown in figure 2. The baseline recommendation pipeline is split into two phases.

In the first phase, given a new user $n$, the similarity between this user and all other existing users is calculated using the cosine similarity metric defined in equation 1.

In the second phase, predicted ratings for all items are calculated as follows:

$$pred(n, i_a) = \frac{\sum_{h=1}^{n'} sim(u_n, u_h)(r_{u_h, i_a})}{\sum_{h=1}^{k'} |sim(u_a, u_h)|} \tag{4}$$

where $n'$ are the users who have also rated item $i_a$

## 3.2 ICS focused model: The PIP CF model (PIP-CF)

The second step of this study focused on the creation of a model which utilized a novel, domain-aware similarity measure. An overview of the main recommendation pipeline for the PIP-CF RS can be found in figure 2. For this model, the PIP similarity measure proposed by [2] was used to calculate the similarity between users.

Just like the B-CF model, the PIP-CF model is split into two phases. In the first phase, given a new user $u_n$, the similarity between this user and all other users is calculated using the PIP similarity measure defined in equation 3.

In the second phase, predicted ratings of user n for all other items are calculated using equation 4

## 3.3 CCS focused model: The demographic filtering model (D-CF)

The third step of this study included the construction of the Demographic recommender system. An overview of the demographic RS pipeline is shown in figure 2. The demographic model was built around the idea that people with a

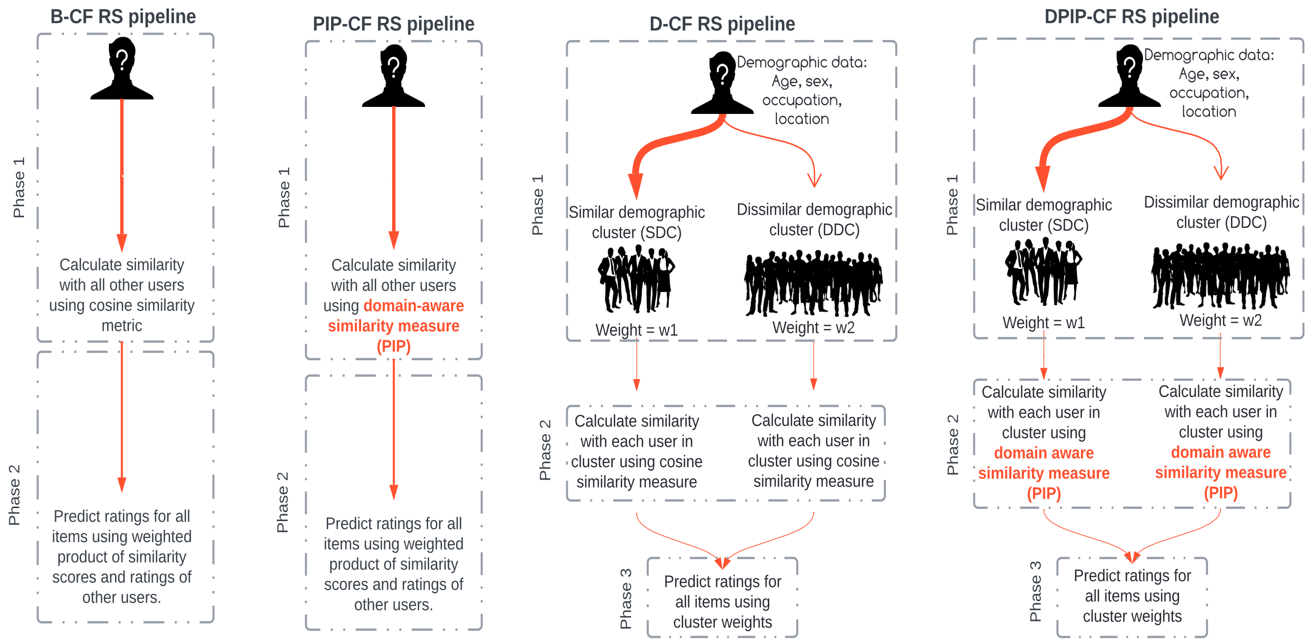


Figure 2: From left to right, the main processes within the baseline (B-CF), PIP-CF, D-CF and DPIP-CF models are presented. The B-CF and PIP-CF models are broken down into two main phases, while the D-CF and DPIP-CF models are broken down into three main phases.

common demographic background are more likely to have similar preferences [10]. To successfully capture this intuition, we split the Demographic recommender system into three phases.

The first phase involved the training of a k-means clustering algorithm on the existing user's demographic data. Let U = [$u_1$, $u_2$... $u_n$] be the set of existing users and n be the new cold-start user. The k-means algorithm is first trained on U, then we assign n to one of the K clusters. The cluster to which user n is assigned is called the similar demographic cluster (SDC) while the other clusters are called dissimilar demographic clusters (DDC).

The second phase focuses on calculating the similarity scores between the new user n and every other user in the SDC and DDC. The cosine similarity measure was used to calculate the similarity scores. Having calculated the similarities between the users, we then predicted the ratings for an item $i_a$ and user $n$ as follows:

$$pred_{SDC} = \frac{\sum_{h=1}^{m^{`}} sim(u_a, u_h)(r_{u_h, i_a})}{\sum_{h=1}^{m^{`}} |sim(u_a, u_h)|} \tag{5}$$

where $m^{`}$ are the users from SDC who have also rated item $i_a$

$$pred_{DDC} = \frac{\sum_{h=1}^{k^{`}} sim(u_a, u_h)(r_{u_h, i_a})}{\sum_{h=1}^{k^{`}} |sim(u_a, u_h)|} \tag{6}$$

where $k^{`}$ are the users from DDC who have also rated item $i_a$

In the third phase, to capture the idea that people with common demographic backgrounds are more similar, we combine the two predicted ratings from clusters SDC and DDC into a single predicted rating as follows

$$pred_{final} = w_1 \times pred_{SDC} + w_2 \times pred_{DDC} \tag{7}$$

Where $w_1$ and $w_2$ are weights given to the SDC and DDC predicted ratings respectively.

### 3.4 CCS and ICS focused model: The final model (DPIP-CF)

The DPIP-CF model is specifically designed to solve the CCS and ICS problems simultaneously. The D-CF and PIP-CF models utilized ideas and concepts which were specifically designed to solve the CCS and ICS problems respectively. The DPIP-CF pipeline can be broken down into three main phases.

The first phase is exactly similar to the first phase of the D-CF model. The existing users are clustered into K clusters and the target user is assigned to one of these K clusters. The cluster to which the target user is assigned is called the SDC and the other clusters are called the DDC.

In the second phase, similarities between the target user and all other users in the SDC and DDC clusters are calculated using the PIP similarity measure defined in equation 3 . Having calculated the similarity scores between the users, the ratings of the test user for an unseen item $i_a$ are then calculated using the two distinct clusters, SDC and DDC. The predicted rating using SDC and DDC are calculated using equation 5 and equation 6 respectively.

Finally, in the third phase, the two ratings are combined using a weighted formulation described in equation 7.

An overview of the recommendation pipeline can be found in figure 2.

# 4 Experiments

This section reports on the performance of all the models that were implemented through the use of various performance metrics. In order to prove the effectiveness of our DPIP-CF model to solve the CCS (zero reviews available situations) and ICS problems (few reviews available situations), Initially experiments were run using the baseline model. Later, experiments on the PIP and demographic models were conducted to showcase the strengths of each model in solving the ICS and CCS problems respectively. Finally, the DPIP-CF model was subjected to the same set of experiments and compared to the previous three models. In section 4.1, various concepts relating to the set-up of experiments are introduced. In section 4.2 the performance results of each model are presented and discussed.

## 4.1 Overview

**Data set**

All experiments were repeated using the MovieLens 1 Million dataset which is provided by the GroupLens research team [6]. The MovieLens dataset contains one million ratings for 3952 movies defined by 6040 users where each user has rated at least 20 movies. Each user profile is associated with some demographic data such as sex, age, occupation and location. The movie profiles contain the genres to which each movie belongs.

**Performance metrics**

This paper makes use of the root-mean-squared-error (RMSE) and mean-absolute-error (MAE) performance metrics to benchmark and compare the various implemented models.

The RMSE is a frequently used measure of the differences between values predicted by a model or an estimator and the values observed. The RMSE is defined as follows:

$$RMSE = \sqrt{(\frac{1}{n}) \times \sum_{i=1}^{n} (y_i - x_i)^2}$$

where:

$n$ = total number of predictions
$y_i$ = the predicted rating for item $i$
$x_i$ = the actual rating for item $i$

The MAE calculates the magnitude of difference between the prediction of an observation and the true value of that observation. The MAE is defined as follows:

$$MAE = (\frac{1}{n}) \times \sum_{i=1}^{n} |(y_i - x_i)|$$

where:

$n$ = total number of predictions
$y_i$ = the predicted rating for item $i$
$x_i$ = the actual rating for item $i$

**Model parameters**

For the D-CF and DPIP-CF models, various combinations of weights were first inspected. Through these investigations, a weight of 0.8 for $w_1$ and 0.2 for $W_1$ were chosen as they showed the best performance. Additionally, the user's sex and age were used as demographic data.

**Simulating artificial cold-start**

The models were designed to overcome various cold-start scenarios. Hence artificial cold-start scenarios were simulated by only allowing the RS models to utilize a small number of ratings per new test user. This was manually done by hiding the required number of reviews from the RS systems. Twenty-five per cent of the data was reserved for testing purposes and the ratings per user were artificially increased (by reintroducing the hidden reviews) from 0 to 60 to simulate CCS and ICS scenarios.

## 4.2 Results

In this subsection, the four models are evaluated using the following performance metrics - RMSE and MAE- and visualized in the following figures.

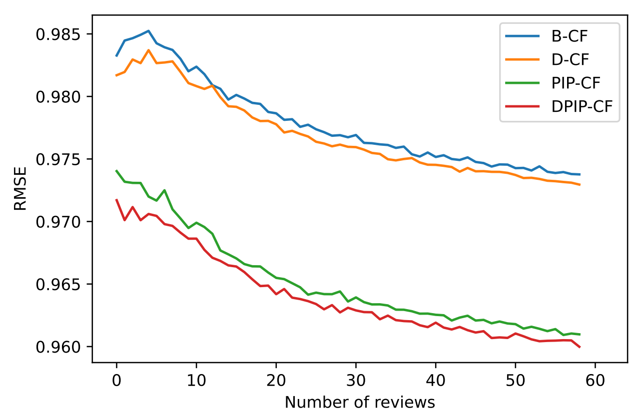


Figure 7: Performance graph of all models using the RMSE metric

First, figure 7 shows the performance of each model, as measured by the *RMSE* metric. It can be seen that the final DPIP-CF model which is a combination of the D-CF and PIP-CF model consistently outperforms every other model in CCS and ICS conditions. Compared to the baseline, the D-CF model seems to show a notable improvement in performance only in CCS conditions. The PIP-CF model shows good performance in ICS conditions.

Next, figures 3, 4, 5 and 6 show the variance in the RMSE for each model. The variance in the RMSE for the B-CF model appears to decrease at a very slow rate. In contrast, all other models show a more rapid decrease in the variance of

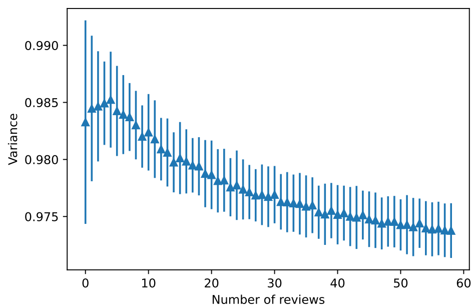


Figure 3: RMSE variance of the Baseline (B-CF) model

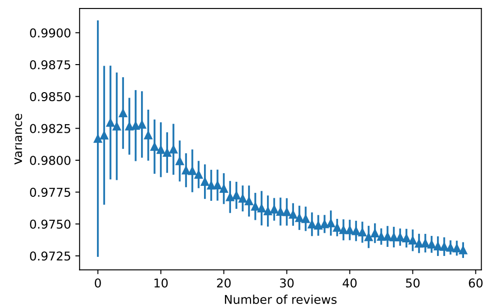


Figure 4: RMSE variance of the D-CF model

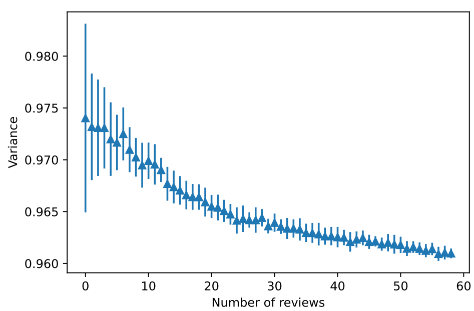


Figure 5: RMSE variance of the PIP-CF model

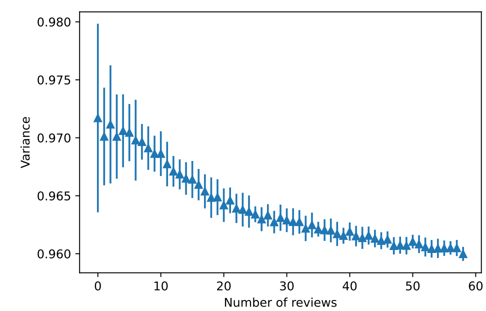


Figure 6: RMSE variance of the DPIP-CF model

their RMSE scores. Among the remaining models, the DPIP-CF and PIP-CF models show the fastest decrease in variance. The slow decrease in variance of the B-CF model indicates that it takes the model substantial number of reviews to be able to make decently accurate predictions. While the three other models are able to quickly converge to a tighter confidence interval. This in conjunction to the lower RMSE values of the three models shows that with just a few reviews, these models can quickly converge to a reasonable predicted rating.

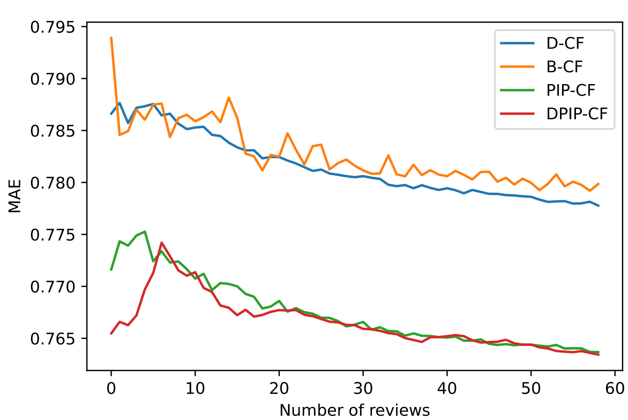


Figure 8: Performance graph of all models using the MAE metric

Finally, figure 8 shows the performance of all the models using the *MAE* metric. If we focus on CCS conditions (where the number of reviews is 0), it is immediately clear that all three models show a substantial improvement compared to the B-CF model, with the DPIP-CF model showing the greatest improvement. Focusing on ICS conditions (where a number of reviews are greater than 0), it can be seen that the B-CF and D-CF models show similar performance. Likewise, the PIP-CF and DPIP-CF models show similar performance in this area.

## 5 Discussion

The results from the previous section indicate that the enrichment of the user profiles by their demographic data seems to show an improvement in CCS recommendation conditions (where the number of reviews available is 0). The D-CF model utilized user demographic data to weight predictions from similar demographic users higher than dissimilar demographic users. When compared to the B-CF model, D-CF showed the greatest improvement in CCS conditions while in ICS conditions the improvement was negligible. This is perfectly in line with the previous research in which demographic data was used to mostly solve the CCS problem.

Although the PIP-CF model was primarily designed to excel in ICS conditions, it also displayed a notable performance increase in CCS conditions compared to the B-CF model. It is worth noting that although this model showed great performance in CCS and ICS conditions, there is a steep increase in performance when moving from CCS to

ICS conditions with the PIP-CF model. This observation indicated that there is potential room for improvement in CCS conditions for this model.

The DPIP-CF model was built on the observation that the strengths of the D-CF model perfectly counteracted the weaknesses of the PIP-CF model and vice versa. This is exactly what we observe in the performance results of the DPIP-CF model that utilized demographic data and a domain-aware similarity measure as mitigation strategies against the CCS and ICS problems respectively. Amongst all models, the DPIP-CF model displayed the best performance in all metrics through the entire testing range. When specifically compared to the PIP-CF model, it can be seen that the transition from CCS to ICS conditions shows a substantial increase in performance. This increase in performance at CCS conditions can be attributed to the use of user demographic data in the recommendation pipeline.

## 6 Responsible Research

**Reproducibility**

Reproducibility is a challenge in the field of computer science. In light of this, In this paper, we present two alternative methods to help any future reader reproduce our research and results. First, the methodology section provides a clear and concise description of the implementational details for each model. Furthermore, the experiments make use of a standard and an openly available dataset which can be downloaded by any future reader. Details for the experimental setup such as parameters and algorithms used are also provided in section 4. Second, we also understand that some readers might not have much experience implementing recommendation systems. Hence we also provide all source code through a public repository [1].

**Ethical considerations**

Some of the recommender systems used in this paper utilise machine learning algorithms, in particular, the K-means clustering algorithm. Based on various user demographics, the algorithm segregates the users into different clusters. A key point to remember for any future researchers who may expand on the work of this paper is that these machine learning algorithms are sensitive to biases in data. If there are any biases within the training data, they will affect the final results of the algorithms [3]. Hence any readers of this paper must keep this in mind and also make sure that the data they use is free from bias, especially if they intend to expand on the user demographic data by using features such as location and ethnicity.

## 7 Conclusions and Future Work

This paper investigated the use of user demographic data to solve the complete cold-start problem (CCS), domain-aware similarity measure to solve the incomplete cold-start (ICS) problem and their combined use to solve the complete and incomplete cold-start problems simultaneously. Based on the various experiments conducted, it can be concluded that the combined use demographic data and a domain-aware similarity measure yields the best results.

A few starting points for future research were identified during creation of this paper.

1. In the current implementation of the DPIP-CF model, the weights w1 and w2 were manually set. A potential extension of these models could be to dynamically scale these weights based on the number of reviews that are available for the test user.
2. Equations 4, 5 and 6 only consider the set of users who have also rated the same items as the test user. In some instances these set of users are very small. As a result the quality of the predicted ratings are negatively affected. This issue could be solved by the use of a content-based RS as a subroutine. The content-based RS could include in this set, new users who have rated items which are *similar* to the test users rated items instead of just users who have rated the exact same items.

[1] https://cse3000-research-project.github.io/2021/Q4

## References


[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.

[2] Hyung Jun Ahn. A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. *Inf. Sci.*, 178(1):37–51, jan 2008.

[3] Ricardo Baeza-Yates. Data and algorithmic bias in the web. In *Proceedings of the 8th ACM Conference on Web Science*, WebSci '16, page 1, New York, NY, USA, 2016. Association for Computing Machinery.

[4] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Jesús Bernal. A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-Based Systems*, 26:225–238, 2012.

[5] Robin Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4):331–370, nov 2002.

[6] F. Maxwell Harper and Joseph A. Konstan. The movielens datasets: History and context. *ACM Trans. Interact. Intell. Syst.*, 5(4), dec 2015.

[7] Axel Hirschel. Alleviating the cold-start problem of collaborative filtering by hybridising it with a demographic recommender system, 2017.

[8] Bruce Krulwich. Lifestyle finder: Intelligent user profiling using large-scale demographic data. *AI Magazine*, 18:37–45, 01 1997.

[9] Ken Lang. Newsweeder: Learning to filter netnews. In *in Proceedings of the 12th International Machine Learning Conference (ML95*, 1995.

[10] Blerina Lika, Kostas Kolomvatsos, and Stathes Hadjiefthymiades. Facing the cold start problem in recommender systems. *Expert Systems with Applications*, 41(4, Part 2):2065–2073, 2014.

[11] Haifeng Liu, Zheng Hu, Ahmad Mian, Hui Tian, and Xuzhen Zhu. A new user similarity model to improve the accuracy of collaborative filtering. *Knowledge-Based Systems*, 56:156–166, 2014.

[12] Luis Martinez, Luis G. Perez, and Manuel J. Barranco. Incomplete preference relations to smooth out the cold-start in collaborative recommender systems. In *NAFIPS 2009 - 2009 Annual Meeting of the North American Fuzzy Information Processing Society*, pages 1–6, 2009.

[13] Michael J. Pazzani. A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review*, 13:393–408, 1999.

[14] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, CSCW '94, page 175–186, New York, NY, USA, 1994. Association for Computing Machinery.

[15] Elaine Rich. User modeling via stereotypes. *Cognitive Science*, 3(4):329–354, 1979.

[16] J. Ben Schafer, Joseph Konstan, and John Riedl. Recommender systems in e-commerce. pages 158–166, 1999.

[17] Andrew I. Schein, Alexandrin Popescul, Lyle H. Ungar, and David M. Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '02, page 253–260, New York, NY, USA, 2002. Association for Computing Machinery.

[18] Andrew I. Schein, Alexandrin Popescul, Lyle H. Ungar, and David M. Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '02, page 253–260, New York, NY, USA, 2002. Association for Computing Machinery.

[19] Upendra Shardanand and Pattie Maes. Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, page 210–217, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[20] Yadav Usha, Duhan Neelam, and Bhatia Komal Kumar. Dealing with pure new user cold-start problem in recommendation system based on linked open data and social network features. *Mobile Information Systems*, 2020, 2020.

[21] Jun Wang, Arjen P. de Vries, and Marcel J. T. Reinders. Unified relevance models for rating prediction in collaborative filtering. *ACM Trans. Inf. Syst.*, 26(3), jun 2008.

[22] Jian Wei, Jianhua He, Kai Chen, Yi Zhou, and Zuoyin Tang. Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69:1339–1351, 3 2017.