Personalized Fairness-aware Re-ranking for Microlending

Weiwen Liu
The Chinese University of Hong Kong
wwliu@cse.cuhk.edu.hk

Jun Guo
Tsinghua University
eeguojun@outlook.com

Nasim Sonboli
University of Colorado Boulder
nasim.sonboli@colorado.edu

Robin Burke
University of Colorado Boulder
robin.burke@colorado.edu

Shengyu Zhang
Tencent
shengyuzhang@gmail.com

ABSTRACT

Microlending can lead to improved access to capital in impoverished countries. Recommender systems could be used in microlending to provide efficient and personalized service to lenders. However, increasing concerns about discrimination in machine learning hinder the application of recommender systems to the microfinance industry. Most previous recommender systems focus on pure personalization, with fairness issue largely ignored. A desirable fairness property in microlending is to give borrowers from different demographic groups a fair chance of being recommended, as stated by Kiva. To achieve this goal, we propose a Fairness-Aware Re-ranking (FAR) algorithm to balance ranking quality and borrower-side fairness. Furthermore, we take into consideration that lenders may differ in their receptivity to the diversification of recommended loans, and develop a Personalized Fairness-Aware Re-ranking (PFAR) algorithm. Experiments on a real-world dataset from Kiva.org show that our re-ranking algorithm can significantly promote fairness with little sacrifice in accuracy, and be attentive to individual lender preference on loan diversity.

CCS CONCEPTS

• **Information systems → Recommender systems**.

KEYWORDS

Fairness-aware re-ranking; Re-ranking; Loan recommendation

ACM Reference Format:
Weiwen Liu, Jun Guo, Nasim Sonboli, Robin Burke, and Shengyu Zhang. 2019. Personalized Fairness-aware Re-ranking for Microlending. In *Thirteenth ACM Conference on Recommender Systems (RecSys '19), September 16–20, 2019, Copenhagen, Denmark.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3298689.3347016

1 INTRODUCTION

Microlending is the provision of small and low-interest loans (as little as $25) to low-income individuals or small-scale entrepreneurs from under-developed countries [17]. The extremely poor people in rural areas often lack collateral, steady employment, or a verifiable credit history, hence they cannot get access to financial services. Under such circumstances, microlending has attracted an increased attention in the last decade [5], providing impoverished entrepreneurs an opportunity to start their own businesses as well as avoid the vicious cycle of debt.

RecSys '19, September 16–20, 2019, Copenhagen, Denmark

ACM ISBN 978-1-4503-6243-6/19/09…$15.00
https://doi.org/10.1145/3298689.3347016

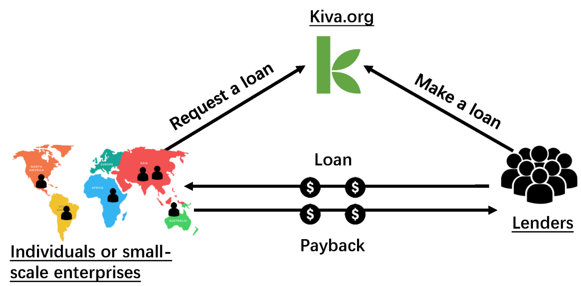


Figure 1: Kiva.org provides an intermediary service for lenders and borrowers.

One of the leading international microlending organizations, Kiva Microfunds (Kiva.org), has crowd-funded about 3 million borrowers with $1.27 billion USD as of February 2019 [13]. Kiva does not collect interest but provides an intermediary service for lenders and borrowers, as illustrated in Figure 1. Borrowers from over 80 countries, divided into 8 regions by Kiva, post their applications for loans on the website for lenders to support. Lenders browse and crowd-fund the loans in the increments of $25 or more.

Loan recommender systems [6, 7] are designed to assist lenders in looking for promising borrowers. Such systems model lenders' historical behaviors and generate personalized recommendations to meet the lenders' interests or needs. While these recommender systems aim to provide efficient and personalized services, two issues are largely overlooked.

(i) *Unfair recommendation.* The existing recommender systems for microlending are all lender-centered. These recommender systems have been demonstrated to favor popular items [4, 14], resulting in extremely unbalanced recommendation results — majority groups are usually over-represented, thereby holding a higher proportion of opportunities and resources, while minority groups barely receive exposure. For example, it is observed that certain geographical regions such as Asia and Africa dominate the recommendation in a designed loan recommender system for Kiva [7]; whereas, others like Oceania and Eastern Europe barely receive

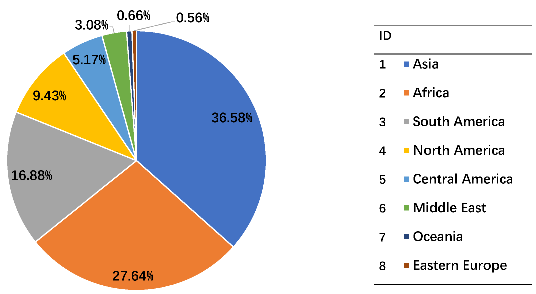


Figure 2: The issue of fairness on regions in a designed loan recommender system [7] for Kiva: the recommendation percentage for each region.

recommendations, as shown in Figure 2. Less exposure means the borrowers from these regions are less likely to be funded.

(ii) *Lender's diversity tolerance.* Another noticeable phenomenon is that lenders' tolerance of diversity varies greatly. Thus, the diversity of recommended lists should be compatible with the level of each lender's interest in diverse recommendations. For instance, some lenders may highly prefer offering loans to certain regions (such as their home countries), while others may be open to diverse regions. In such scenarios, assuming lenders' diversity tolerance is constant and increasing diversity uniformly for all lenders will result in poor recommendations [8]. Therefore, a well-designed recommender system should be personalized by both lenders' interests and the degree of loan diversity in recommended lists.

We aim to design a fairness-aware re-ranking algorithm on top of the existing recommendation algorithms. Our algorithm achieves a balance between recommendation accuracy and borrower-side fairness, and also considers lenders' preferences for diversity. This post-processing step does not depend on any specific recommendation algorithm, and therefore can be widely applied.

2 THE PROPOSED ALGORITHM

In this section, we first propose to formulate this recommendation scenario as a Multi-sided Recommender System (MRS) [1, 2]. Then, we design a personalized re-ranking algorithm to achieve a fair recommendation for microlending.

2.1 Problem Formulation

In Kiva.org, different stakeholders are involved, including lenders, borrowers, and the system (Kiva), which can be modeled by an MRS. Borrowers post their loan applications onto the system from which lenders receive loan recommendations. Besides maximizing the lenders' interests, Kiva also aims to consider the allocation of borrower side recommendation opportunities.

To achieve the above goal, given a set of lenders $\mathcal{U} = \{1, \ldots, n_u\}$, a set of loans $\mathcal{V} = \{1, \ldots, n_v\}$, and an initial ranking list $R(u)$ for lender $u \in \mathcal{U}$, our task is to re-rank $R(u)$ and generate a list of K distinct loans $S(u)$ that is both accurate and fair. For a loan $v \in \mathcal{V}$, $C(v) \in \{1, \ldots, n_c\}$ is the corresponding categorical protected attribute, such as region, race, or gender. Let $\mathcal{V}_c = \{v | C(v) = c, v \in \mathcal{V}\}$ denote the *group* of loans with attribute c. For instance, if the protected attribute is the geographical region and we use the ID specified in Figure 2, then $\mathcal{V}_c$ with $c = 1$ represents the set of all loans applied from Asia.

In some contexts, we may separate borrowers into protected and unprotected groups, where the borrowers from the protected group are the particular concern [18]. This can be viewed as a special case of our problem with $n_c = 2$. In this research, we are looking at more general fairness across all borrower groups and trying to ensure that each re-ranked list $S(u)$ for a specific lender u covers as many borrower groups as possible while considering the personalized constraints on diversity tolerance. However, we are not allowed to sacrifice too much accuracy to achieve an absolutely fair recommendation result. A tradeoff must be made since accuracy and fairness cannot be fully satisfied at the same time [3].

2.2 Algorithm

We first propose a fairness-aware re-ranking algorithm (FAR) and then incorporate a diversity tolerance term τ_u to produce the personalized fairness-aware re-ranking algorithm (PFAR).

2.2.1 *Fairness-aware Re-ranking (FAR).*

Our proposed Fairness-aware Re-ranking (FAR) criterion is defined as Eq.(1), which is the combination of a personalization-induced term and a fairness-induced term, with a hyper-parameter $\lambda \in (0, 1)$ controlling the tradeoff between the two. For any $u \in \mathcal{U}$, we solve

$$\max_{v \in R(u)} \underbrace{(1-\lambda)P(v|u)}_{\text{personalization}} + \underbrace{\lambda \sum_{c} P(\mathcal{V}_c)\mathbb{1}_{\{v \in \mathcal{V}_c\}} \prod_{i \in S(u)} \mathbb{1}_{\{i \notin \mathcal{V}_c\}}}_{\text{fairness}}, \quad (1)$$

where $P(v|u)$ is the personalization score determined by the base recommender, indicating the probability of lender u being interested in loan v. The indicator function $\mathbb{1}_A$ has the value 1 if A is true, and 0 otherwise. The new output list is built iteratively in a greedy manner. At each step, the algorithm selects one loan with the highest re-ranking score from the candidate list $R(u)$ and moves it to the output list $S(u)$.

For borrower-side fairness, our idea is to promote the loans that belong to currently uncovered borrower groups. For a loan v that belongs to $\mathcal{V}_c$, we first compute the coverage of $\mathcal{V}_c$ for the current generated re-ranked list $S(u)$ as $\prod_{i \in S(u)} \mathbb{1}_{\{i \notin \mathcal{V}_c\}}$, which is equal to 1 if none of the items in $S(u)$ belong to $\mathcal{V}_c$, and 0 otherwise. If both $\mathbb{1}_{\{v \in \mathcal{V}_c\}}$ and $\prod_{i \in S(u)} \mathbb{1}_{\{i \notin \mathcal{V}_c\}}$ are 1, items that belong to $\mathcal{V}_c$ are promoted by being assigned a higher score, and thus get a larger chance of being selected. The above process is repeated for each borrower group $\mathcal{V}_c$, $c = 1, \ldots, n_c$, and the results are summed up. Since each item may belong to multiple groups, the loans belonging to multiple uncovered borrower groups are favored.

The normalization term $P(\mathcal{V}_c)$ is determined by the system and indicates the importance of $\mathcal{V}_c$. For example, if a borrower group is identified as a protected group and receives few recommendations, then the system can assign a higher $P(\mathcal{V}_c)$ to the corresponding group. For simplicity, we assume a uniform preference over borrower groups and assign an equal $P(\mathcal{V}_c)$ for all borrower groups.

2.2.2 *Personalized Fairness-aware Re-ranking (PFAR).*

Note that Eq.(1) is designed for any lender $u \in \mathcal{U}$. If we simply follow this setting and treat each lender with an equal level of diversity tolerance, then the ranking quality is inevitably suppressed. Actually,

Algorithm 1 (Personalized) Fairness-Aware Re-ranking (FAR/PFAR)

Input: $u, R(u), K, \lambda, \tau_u$
Output: $S(u)$

$S(u) \leftarrow \emptyset$
while $|S(u)| < K$ **do**
  Select the optimal v^* by solving

$$\arg\max_{v \in R(u)} (1-\lambda)P(v|u) + \lambda\tau_u \sum_c P(\mathcal{V}_c)\mathbb{1}_{\{v \in \mathcal{V}_c\}} \prod_{i \in S(u)} \mathbb{1}_{\{i \notin \mathcal{V}_c\}}$$

  $R(u) \leftarrow R(u) \setminus \{v^*\}$
  $S(u) \leftarrow S(u) \cup \{v^*\}$
end while
return $S(u)$

lenders' propensity towards diversity varies and different levels of diversity should be considered.

To address this issue, we personalize the previous re-ranking criterion Eq.(1) by adding a personalized weight τ_u and derive our Personalized Fairness-aware Re-ranking (PFAR) criterion Eq.(2). For any $u \in \mathcal{U}$, we solve

$$\max_{v \in R(u)} \underbrace{(1-\lambda)P(v|u)}_{\text{personalization}} + \underbrace{\lambda\tau_u \sum_c P(\mathcal{V}_c)\mathbb{1}_{\{v \in \mathcal{V}_c\}} \prod_{i \in S(u)} \mathbb{1}_{\{i \notin \mathcal{V}_c\}}}_{\text{personalized fairness}}, \tag{2}$$

where the second term considers personalized fairness. The diversity tolerance τ_u is incorporated to control the weight of the fairness score. If a lender has no special interest to a specific group, the algorithm will focus more on the personalization task. Otherwise, the fairness-induced term can be emphasized.

To calculate τ_u, we first compute a level of interest $P(\mathcal{V}_c|u)$ of lender u for each borrower group $\mathcal{V}_c$, $c = 1, \ldots, n_c$,

$$P(\mathcal{V}_c|u) \triangleq \frac{\sum_v r(u,v)\mathbb{1}_{\{v \in \mathcal{V}_c\}}}{\sum_{c'}\sum_v r(u,v)\mathbb{1}_{\{v \in \mathcal{V}_{c'}\}}}, \tag{3}$$

where $r(u,v)$ is the rating from lender u to loan v. We compute the ratio of summation over rated borrowers that belong to group c over summation over all the rated borrowers.

The preference $P(\mathcal{V}_c|u) \in [0,1]$ indicates the lender's taste over borrower groups where $\sum_c P(\mathcal{V}_c|u) = 1$. Some lenders may be highly interested in certain borrower groups, while some lenders may have equal preferences over all the borrower groups. To capture this characteristic, we use the information entropy [16] to identify the lender diversity tolerance, namely

$$\tau_u \triangleq -\sum_c P(\mathcal{V}_c|u)\log P(\mathcal{V}_c|u), \tag{4}$$

where a larger τ_u means that the lender is more open to a diverse set of borrower groups.

The algorithm FAR/PFAR is formally given in Algorithm 1. For a lender u, loans are generated iteratively from the initial ranking list. The loan with the highest score is selected from the candidate list $R(u)$ according to our re-ranking criterion Eq.(1) or Eq.(2). The process is repeated until $S(u)$ has reached the desired length. The proposed algorithm automatically balances personalization and fairness by adding a bonus to the loans that belong to the uncovered borrower groups. The generated re-ranked list for each lender tends to cover each borrower group at least once while encouraging personalization.

3 EXPERIMENTS

In this section, we test our proposed algorithms on a real-world dataset from Kiva.org. The performance of our proposed re-ranking algorithms on top of different base recommenders is evaluated in terms of accuracy and fairness. Our implementation is built upon LibRec 2.0 [9]. All results are averages from five-fold cross-validation.

3.1 Dataset

Our algorithms are evaluated on a proprietary dataset obtained from Kiva.org, including all lending transactions over an 8-month period. Each loan is specified by features including borrower's name, gender, borrower's country, loan purpose, funded date, posted date, loan amount, loan sector, and geographical coordinates.

To generate a denser dataset with greater potential for user profile overlap, we apply a content-based technique [15], creating *pseudo-items* that represent large categories of items. In particular, all loans that share the same borrower gender, borrower country, loan purpose, loan amount (binned to 5 equal-sized buckets), and loan sector are combined into a single pseudo-item. Then we apply a 10-core transformation, selecting pseudo-items with at least 10 lenders who had funded at least 10 pseudo-items. The retained dataset has 11,085 pseudo-items, 9,597 lenders and 204,830 ratings.

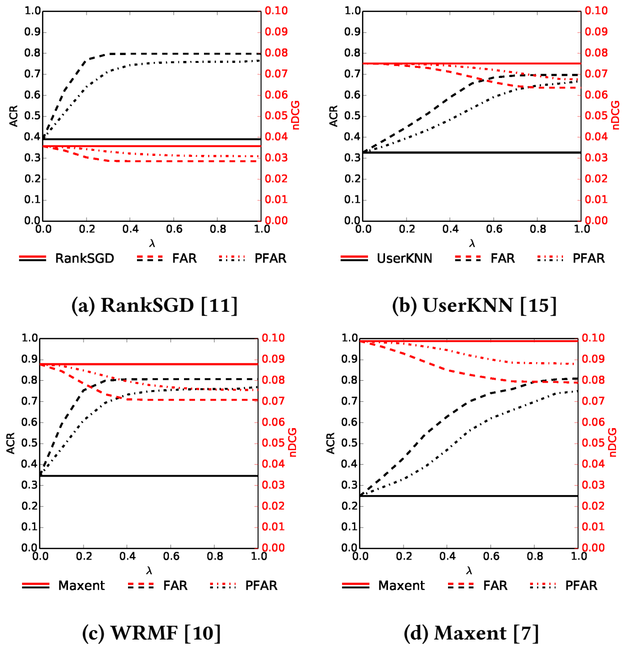
3.2 Comparative Recommenders

As of this writing, Kiva.org does not offer recommendation functionality. In our experiments, we assume a context in which the site provides short lists of recommended loans to lenders for their review. We set the protected attribute as the geographical region, because part of Kiva.org's mission is to achieve equitable access to capital across regions. In order to set up the recommendation scenario for Kiva, we select four representative base recommenders to study their performance in accuracy and fairness, as well as how our proposed algorithms can influence the recommendation results: (a) RankSGD [11] uses stochastic gradient descent to optimize the ranking error; (b) UserKNN [15] is a memory-based collaborative algorithm that computes user similarity; (c) Weighted Regularized Matrix Factorization (WRMF) [10] creates a reduced-dimensionality factorization of the rating matrix; (d) Maximum-entropy distribution (Maxent) [7] is a loan recommender system specially designed for Kiva. Maxent models lending behaviors by estimating a maximum-entropy distribution based on a set of heterogeneous information regarding micro-financial transactions available at Kiva.

3.3 Evaluation Metrics

We propose to utilize Normalized Discounted Cumulative Gain (nDCG) [12] and Average Coverage Rate (ACR) to evaluate recommendation accuracy and borrower-side fairness, respectively.

$$\text{ACR} = \frac{\sum_{u \in U_t} N_{S(u)}}{N_{\text{bg}}|U_t|}, \tag{5}$$



(a) RankSGD [11]

(b) UserKNN [15]

(c) WRMF [10]

(d) Maxent [7]

Figure 3: Tendencies of ACR and nDCG with increasing λ.

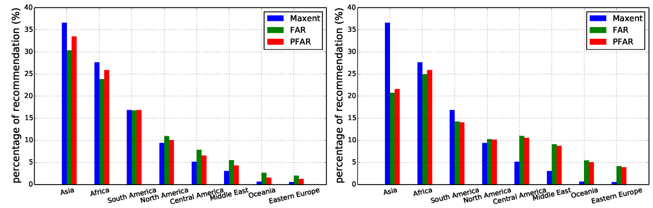
where U_t is the test lender set, $|U_t|$ is the number of lenders in the test set, N_{bg} is the total number of borrower groups and $N_{S(u)}$ is the number of borrower groups covered in the list $S(u)$. A larger ACR indicates a fairer system regarding borrower-side fairness.

3.4 Results and Analysis

We first study the performance of applying FAR and PFAR to the base recommenders. We vary the hyper-parameter λ from 0 to 1 in steps of 0.1 and record the corresponding ACR and nDCG, where a larger λ means the weight of fairness is larger. The results are shown in Figure 3.

Considering the performance of all base recommenders (the solid lines), rankSGD has the lowest accuracy (nDCG = 0.0358); UserKNN comes next with nDCG=0.0752 by finding the nearest neighbors; WRMF performs better than UserKNN by learning the latent factors of lenders and loans (nDCG=0.0878); Maxent obtains the highest nDCG of 0.0988 since Maxent is specially designed for loan recommendation and additional features of loans, *e.g.*, loan sector and geographical coordinates, are utilized. However, accurate recommenders tend to favor items from certain groups, thus resulting in fairness issues.

Effectiveness of the re-ranking algorithms. By applying our proposed algorithms (when $\lambda \in (0, 1)$), all recommenders tend to achieve fairer recommendation results by promoting the loans that belong to less-popular groups, showing the flexibility and effectiveness of our proposed algorithms. Our re-ranking algorithms can be deployed to any base recommender, with the weight of the fairness tunable. The accuracy slightly decreases as λ increases, as there is a price to pay for obtaining a fairer system. Take $\lambda = 0.1$ for instance, Maxent can obtain a gain of 25.1% in ACR with a loss of merely 1.6% in nDCG. Moreover, RankSGD and WRMF converge faster than UserKNN and Maxent with the increase of λ, indicating that the behavior of our proposed algorithms depends on the initial ranking list to some extent.



(a) $\lambda = 0.1$, $\mathbf{nDCG_{FAR}} = 0.0962$, $\mathbf{nDCG_{PFAR}} = 0.0983$.

(b) $\lambda = 0.99$, $\mathbf{nDCG_{FAR}} = 0.0709$, $\mathbf{nDCG_{PFAR}} = 0.0756$.

Figure 4: Recommendation percentage of each region.

Comparison between FAR and PFAR. The recommendation accuracy of PFAR is higher than FAR, since PFAR limits the amount of loan diversity that the re-ranking imposes, based on the individual tolerance. We can also observe from Figure 3 that fairness of PFAR is lower, which is consistent with our previous discussion and demonstrates the tradeoff between accuracy and fairness.

Visualization of the re-ranking results. We compute the percentage of recommendations for each group with and without the proposed re-ranking algorithms, and study the corresponding allocation distribution. Due to the 4-page limitation, we chose Maxent as an example, and the results are shown in Figure 4. Similar trends can be observed for other base recommenders.

(i) The blue bars show the distribution of the base recommender Maxent, *i.e.*, $\lambda = 0$. We observe that Maxent focuses on a few major borrower groups, namely Asia and Africa (making up 36.58% and 27.64% of total recommendations, respectively), while paying less attention to the others.

(ii) The ideal fairness is to give each group an equal chance of being recommended. However, accuracy will be significantly downgraded as lenders' preferences are not learned. As a compromise, we find a balance between the two and select $\lambda = 0.1$, where the growth rate of ACR per unit nDCG loss is the largest. As illustrated by the green and red bars in Figure 4a, nDCG still remains at a high level after the re-ranking (nDCG=0.0962 for FAR and nDCG=0.0983 for PFAR), while fairness of the recommendation is significantly improved, as loans belonging to less-popular groups are promoted.

(iii) In Figure 4b, a larger $\lambda = 0.99$ is applied, and the distribution can be even more balanced, while the accuracy is lower.

4 CONCLUSION

In this work, we proposed a personalized fairness-aware re-ranking algorithm for microlending that can balance accuracy and fairness. We increase the coverage rate of borrowers' regions for Kiva.org to achieve borrower-side fairness, and we show that our algorithm can do so with minimal loss in ranking accuracy. In addition, our algorithm includes lender-specific weights that can be used to personalize the degree of loan diversity.

ACKNOWLEDGMENTS

Jun Guo is partly supported by the 2018 Tencent Rhino-Bird Elite Training Program and Tsinghua-Berkeley Shenzhen Institute Grant.

REFERENCES

[1] Robin Burke. 2017. Multisided fairness for recommendation. *arXiv preprint arXiv:1707.00093* (2017).

[2] Robin Burke and Himan Abdollahpouri. 2017. Patterns of Multistakeholder Recommendation. *arXiv preprint arXiv:1707.09258* (2017).

[3] Burke, Robin. 2017. Multisided Fairness for Recommendation. In *Workshop on Fairness, Accountability and Transparency in Machine Learning (FATML)*. arXiv, Halifax, Nova Scotia, Article arXiv:1707.00093 [cs.CY], 5 pages.

[4] Òscar Celma and Pedro Cano. 2008. From hits to niches?: or how popular artists can bias music recommendation and discovery. In *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition*. ACM, 5.

[5] Juanyi Chen, Amber Y Chang, and Garry D Bruton. 2017. Microfinance: Where are we today and where should the research go in the future? *International Small Business Journal* 35, 7 (2017), 793–802.

[6] Jaegul Choo, Changhyun Lee, Daniel Lee, Hongyuan Zha, and Haesun Park. 2014. Understanding and promoting micro-finance activities in kiva. org. In *Proceedings of the 7th ACM international conference on Web search and data mining*. ACM, 583–592.

[7] Jaegul Choo, Daniel Lee, Bistra Dilkina, Hongyuan Zha, and Haesun Park. 2014. To gather together for a better world: Understanding and leveraging communities in micro-lending recommendation. In *Proceedings of the 23rd international conference on World wide web*. ACM, 249–260.

[8] Farzad Eskandanian, Bamshad Mobasher, and Robin Burke. 2017. A Clustering Approach for Personalizing Diversity in Collaborative Recommender Systems. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 280–284.

[9] Guibing Guo, Jie Zhang, Zhu Sun, and Neil Yorke-Smith. 2015. LibRec: A Java Library for Recommender Systems. In *UMAP Extended Proceedings*. CEUR, Dublin, Ireland, Article 9, 4 pages.

[10] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*. Ieee, 263–272.

[11] Michael Jahrer and Andreas Töscher. 2012. Collaborative Filtering Ensemble for Ranking. In *Proceedings of KDD Cup 2011 (Proceedings of Machine Learning Research)*, Vol. 18. PMLR, 153–167.

[12] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems (TOIS)* 20, 4 (2002), 422–446.

[13] kiva. 2018. Kiva - Loans that change lives. https://www.kiva.org/

[14] Eric L Lee, Jing-Kai Lou, Wei-Ming Chen, Yen-Chi Chen, Shou-De Lin, Yen-Sheng Chiang, and Kuan-Ta Chen. 2014. Fairness-aware loan recommendation for microfinance services. In *Proceedings of the 2014 International Conference on Social Computing*. ACM, Beijing, China, 3.

[15] Paul Resnick and Hal R Varian. 1997. Recommender systems. *Commun. ACM* 40, 3 (1997), 56–58.

[16] Claude Elwood Shannon. 2001. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review* 5, 1 (2001), 3–55.

[17] Muhammad Yunus. 1998. *Banker to the Poor*. Penguin Books India.

[18] Indre Zliobaite. 2015. A survey on measuring indirect discrimination in machine learning. *arXiv preprint arXiv:1511.00148* (2015).