

000 001 002 003 004 005 GRADIENT-BASED DIVERSITY OPTIMIZATION WITH 006 DIFFERENTIABLE TOP- k OBJECTIVE 007 008 009

010 **Anonymous authors**
011 Paper under double-blind review
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026

ABSTRACT

027 Predicting relevance is a pervasive problem across digital platforms, covering
028 social media, entertainment, and commerce. However, when optimized solely for
029 relevance and engagement, many machine-learning models amplify data biases and
030 produce homogeneous outputs, reinforcing filter bubbles and content uniformity.
031 To address this issue, we introduce a pairwise top- k diversity objective with a
032 differentiable smooth-ranking approximation, providing a model-agnostic way to
033 incorporate diversity optimization directly into standard gradient-based learning.
034 Building on this objective, we cast relevance and diversity as a joint optimization
035 problem, we analyze the resulting gradient trade-offs, and propose two comple-
036 mentary strategies: direct optimization, which modifies the learning objective, and
037 indirect optimization, which reweights training data. Both strategies can be applied
038 either when training models from scratch or when fine-tuning existing relevance-
039 optimized models. We use recommendation as a natural evaluation setting where
040 scalability and diversity are critical, and show through extensive experiments that
041 our methods consistently improve diversity with negligible accuracy loss. Notably,
042 fine-tuning with our objective is especially efficient, requiring only a few gradient
043 steps to encode diversity at scale.
044

1 INTRODUCTION

045 Modern prediction models are typically evaluated by their ability to maximize accuracy, that is,
046 to rank or classify items in line with ground-truth labels. However, optimizing exclusively for
047 accuracy often yields homogeneous outputs: models repeatedly select similar items, overlook long-
048 tail instances, and fail to provide novel or diverse options. This challenge is particularly evident in
049 recommender systems, which influence decisions in shopping, entertainment, and news (Ricci et al.,
050 2010). However, it also arises broadly in ranking and selection tasks across machine learning (Wang
051 et al., 2023). When diversity is neglected, models risk reinforcing bias, amplifying popularity effects,
052 and reducing the utility of top- k prediction sets (Nguyen et al., 2014; Lambrecht & Tucker, 2019).

053 Diversity has thus emerged as an important complementary criterion. Recent user studies confirm that
054 diverse outputs can improve satisfaction and engagement when achieved without major loss in relevance
055 (Chen et al., 2018a; Holtz et al., 2020; Anderson et al., 2020). The trade-off between relevance
056 and diversity is mainly tackled by either *post-hoc re-ranking methods* or *learning-based methods*.
057 *Post-hoc re-ranking methods* (e.g., MMR, DPP) modify top- k sets to improve diversity (Carbonell &
058 Goldstein, 1998; Chen et al., 2018b), but typically suffer from degrading accuracy when diversity
059 increases (Chen et al., 2017). *Model-specific learning-based methods* integrate diversity into training
060 objectives (Borodin et al., 2017; Hurley, 2013; Wang et al., 2023), achieving strong gains with modest
061 accuracy loss. However, these methods rely on opaque models that obscure the source of diversity
062 gains, exhibit slow convergence, and are sensitive to the choice of trade-off parameter balancing
063 relevance and diversity. Additionally, *data-centric approaches* such as augmentation, reweighting,
064 and debiasing (Wang et al., 2021; Lai et al., 2023; Ren et al., 2018; Rastegarpanah et al., 2019)
065 address bias in data distributions, but do not explicitly target diversity. Despite these efforts, we
066 still lack a *unified, differentiable, and model-agnostic framework* for optimizing both relevance and
067 diversity directly during training.

068 To address these challenges, we propose a unified framework that leverages differentiable ranking to
069 optimize diversity in top- k prediction sets in a scalable and model-agnostic way. At the core, we use
070

054 an effective diversity objective that can be integrated into the gradient-based training without requiring
 055 architecture changes or post-processing. Building on this objective, we introduce two diversification
 056 methods. (i) *direct diversity-guided tuning* (DDT), which augments the loss with a joint relevance-
 057 diversity term, and (ii) *meta-diversity reweighting* (MDR), which preserves relevance-only training
 058 while reweighting data points using the joint loss as a meta-objective. Our approach offers a flexible
 059 alternative to post-hoc or model-specific diversification without compromising efficiency.

060 Our contributions are threefold: (1) We propose a unified differentiable framework for optimizing
 061 relevance and diversity in top- k prediction sets, applicable to both end-to-end training and fine-
 062 tuning; (2) we provide a theoretical analysis of gradient conflicts, deriving feasible intervals for
 063 the trade-off parameter β and showing that an adaptive update coincides with the two-objective
 064 solution of multi-gradient descent algorithm (MGDA), guaranteeing convergence to Pareto-stationary
 065 points; (3) we empirically validate the framework on five benchmark datasets and two model
 066 architectures, demonstrating that DDT and MDR achieve substantial diversity improvements with
 067 minimal relevance loss, outperforming strong baselines. Notably, the diversity gains extend beyond
 068 the explicitly optimized top- k range, reshaping subsequent predictions as well.

069 2 RELATED WORK

070 Our work is related to diverse recommender systems and multi-objective learning.

071 **Diversity in recommender systems.** Among the vast literature on recommender systems, the closest
 072 are post-hoc and learning-based diversification methods (Zhao et al., 2025); see the survey for a
 073 broader overview. *Post-hoc* methods re-rank the output of a relevance-only model to balance relevance
 074 and diversity. Representative approaches include maximal marginal relevance (MMR) (Carbonell
 075 & Goldstein, 1998), diversity-weighted utility maximization (DUM) (Ashkan et al., 2015), and
 076 determinantal point processes (DPP) (Chen et al., 2018b). These methods are model-agnostic and
 077 easy to implement, but their performance is limited by the quality of the initial relevance ranking,
 078 and diversity gain usually comes at a cost of reduced accuracy (Chen et al., 2017).

079 *Learning-based* approaches incorporate diversity objectives directly into training, including penalties
 080 for similarity among recommended items (Hurley, 2013; Wasilewski & Hurley, 2016), formulations
 081 that optimize relevance–diversity trade-offs (Wang et al., 2023) list-wise, and graph-based models
 082 that encourage coverage of item categories or long-tail exposure (Zheng et al., 2021; Yang et al.,
 083 2023). While they often outperform post-hoc re-ranking, they require architectural modifications
 084 or adversarial training, making them model-specific and computationally heavy. In contrast, our
 085 framework is differentiable and model-agnostic: it can be integrated into standard training pipelines
 086 without altering architectures or adding inference overhead.

087 **Multi-objective learning.** Related is the study of multi-objective optimization for balancing goals
 088 such as accuracy, fairness, and revenue (Zheng & Wang, 2022). Classical approaches include
 089 scalarization (Paul et al., 2022; Di Noia et al., 2017), which reduces multiple objectives to a single
 090 weighted loss, and population-based heuristics such as evolutionary algorithms (Cai et al., 2020),
 091 which approximate the Pareto front. While effective in some cases, these approaches either rely on
 092 carefully tuned weights or suffer from high computational cost.

093 More recently, gradient-based methods such as the multi-gradient descent algorithm
 094 (MGDA) (Désidéri, 2012) have been applied to recommendation. For instance, MGDA has been used
 095 to balance accuracy with fairness (Du et al., 2025; Wu et al., 2022) and with revenue (Milojkovic et al.,
 096 2019). These methods guarantee convergence to Pareto-stationary solutions, but their application has
 097 so far been limited to objectives other than diversity.

100 3 PROBLEM FORMULATION

101 In this section, we introduce top- k diversity and our objectives, starting with notation. Given a
 102 set of candidate items $\mathcal{I} = \{i_1, i_2, \dots, i_m\}$ and a collection of users $\mathcal{U} = \{u_1, u_2, \dots, u_n\}$, the
 103 goal is to identify those items in \mathcal{I} that are most relevant for each $u \in \mathcal{U}$. We assume access to a
 104 partially-observed supervision matrix $\mathbf{R} \in \mathbb{R}^{n \times m}$, where entries $\mathbf{R}_{u,i}$ represent relevance scores
 105 (e.g., rating, label, or interaction). The set of all *observed* scores is $\Omega = \{(u, i, \mathbf{R}_{u,i})\}^N$, where
 106 $N \ll nm$ is the number of observations. We use a relevance prediction model \mathcal{F}_{Θ} with parameters

108 Θ to estimate the remaining scores $\tilde{\mathbf{R}}_{u,i} = \mathcal{F}_\Theta(u, i)$. To learn such a model, we consider two
 109 widely-used approaches. First, we consider *matrix factorization* (MF) (Koren et al., 2009) which
 110 predicts ratings $\tilde{\mathbf{R}}_{u,i} = \mathbf{x}_u^\top \mathbf{y}_i$, for user u and item i using the embeddings $\mathbf{x}_u, \mathbf{y}_i \in \mathbb{R}^d$. Second, we
 111 consider *neural network models* (He et al., 2017), which predict ratings $\tilde{\mathbf{R}}_{u,i} = \text{MLP}([\mathbf{x}_u; \mathbf{y}_i])$ with
 112 a multi-layer perceptron (MLP), allowing non-linear interactions.
 113

114 In both cases, the models are trained to minimize the regularized *mean squared error* (MSE)

$$115 \quad \Theta^* = \arg \min_{\Theta} \sum_{(u,i) \in \Omega_T} (\mathbf{R}_{u,i} - \tilde{\mathbf{R}}_{u,i})^2 + \lambda \|\Theta\|_2^2, \quad (1)$$

117 between observed relevance scores and predictions. While this approach yields models that predict
 118 accurately, neither MF nor MLP is optimized for diversity. Top- k diversity seeks to predict relevance
 119 scores $\tilde{\mathbf{r}}_u \in \mathbb{R}^m$ of a user u , so that the top- k scores correspond to a set of diverse items $Z_u(k) =$
 120 $\{z_1, \dots, z_k\}$, indicated by $\mathbf{l}_u = \text{top}_k(\tilde{\mathbf{r}}_u)$. To measure the diversity of the top- k (highest scoring)
 121 items, we take a *distance-based* approach (Hassin et al., 1997). In particular, for a given pairwise
 122 item-item affinity matrix $\mathbf{S} \in \mathbb{R}^{m \times m}$, we define diversity as the *average pairwise dissimilarity*
 123

$$124 \quad D_{\mathbf{S}}(Z_u(k)) = \frac{2}{k(k-1)} \sum_{i=1}^m \sum_{j=1}^m \mathbf{l}_u(i) \mathbf{l}_u(j) (1 - \mathbf{S}_{i,j}) \quad (2)$$

126 of items in $Z_u(k)$. Then, the average top- k diversity is simply
 127

$$128 \quad \mathcal{L}_{\text{DRO}}(k) = \frac{1}{n} \sum_{u=1}^n D_{\mathbf{S}}(Z_u(k)). \quad (3)$$

130 While this serves as a natural diversity objective, it cannot be directly optimized with gradients as
 131 $\text{top}_k(\cdot)$ involves non-differentiable operations. Our goal, however, is to optimize for relevance and
 132 top- k diversity *simultaneously* as part of gradient-based optimization. To achieve this, we relax the
 133 non-differentiable diversity reward objective (DRO) using a *differentiable surrogate* (DDRO) next.
 134

135 **Differentiable diversity.** To overcome the non-differentiability challenge in $\text{top}_k(\cdot)$, we adopt
 136 *differentiable ranking* (Blondel et al., 2020), which is a continuous relaxation of sorting. The key
 137 idea is to replace the discrete permutation \mathbf{z}_u with a soft ranking vector $\tilde{\mathbf{z}}_u^{(\varepsilon)} \in \mathbb{R}^m$, obtained by
 138 projecting the predicted scores $\tilde{\mathbf{r}}_u$ onto the permutohedron \mathcal{P}_m —the convex hull of all permutations
 139 of $(1, 2, \dots, m)$ embedded in an m -dimensional space. This projection is computed by solving the
 140 following entropy-regularized optimization problem (Blondel et al., 2020):

$$141 \quad \tilde{\mathbf{z}}_u^{(\varepsilon)} = \text{softmax}(\tilde{\mathbf{r}}_u) := \arg \min_{r \in \mathcal{P}_m} \left\{ \frac{1}{\varepsilon} \langle \tilde{\mathbf{r}}_u, r \rangle + H(r) \right\}, \quad (4)$$

143 where $H(r)$ is the entropy regularizer and $\varepsilon > 0$ controls the approximation smoothness. This
 144 definition enables the top- k soft indicator $\tilde{\mathbf{l}}_u(i) = \sigma_\tau(k - \tilde{\mathbf{z}}_u(i))$ using a scaled sigmoid function
 145 $\sigma_\tau(x) = [1 + \exp(-x/\tau)]^{-1}$ for user-defined smoothness-sharpness τ , often set to 1 in our experiments.
 146 In turn, this makes it possible to train the prediction model end-to-end using diversity-aware
 147 gradient updates. Like the discrete counterpart, soft ranking operates with $O(n \log n)$ time and $O(n)$
 148 space complexity. The obvious question is, *does soft ranking lead to sufficiently accurate top- k
 149 recommendations?* The better soft ranking approximates the hard ranking, the more reliable are our
 150 top- k recommendations, formally summarized in Lemma. 1.

151 **Lemma 1** (Soft rank approximation (Blondel et al., 2020)). *Given a rating vector $\tilde{\mathbf{r}}_u \in \mathbb{R}^n$
 152 be the soft rank vector obtained from optimizing (4). Then, as $\varepsilon \rightarrow 0$, the soft ranks converge to
 153 the true ranks of $\tilde{\mathbf{r}}_u$ $\lim_{\varepsilon \rightarrow 0} \tilde{\mathbf{z}}_u^{(\varepsilon)} = \text{rank}(\tilde{\mathbf{r}}_u)$, where $\text{rank}(\tilde{\mathbf{r}}_u) \in \{1, \dots, n\}^n$ denotes the discrete
 154 ranks (breaking ties arbitrarily).*

155 Replacing indicator in Equation (3) with a the soft-ranking-derived indicator $\tilde{\mathbf{l}}_u$ yields the *differentiable*
 156 *diversity reward objective*

$$158 \quad \mathcal{L}_{\text{DDRO}} = \frac{1}{n \cdot N} \sum_u^n \sum_{i=1}^m \sum_{j=1}^m \tilde{\mathbf{l}}_u(i) \tilde{\mathbf{l}}_u(j) (1 - \mathbf{S}_{i,j}). \quad (5)$$

161 To achieve both a high relevance and a high diversity in top- k outputs, we balance both the relevance
 162 objective \mathcal{L}_{rel} and diversity objective \mathcal{L}_{div} introducing our problem below.

162 **Problem 2.** For a given model class \mathcal{F}_Θ , an item affinity matrix \mathbf{S} , a user-defined relevance-diversity
163 trade-off $\beta \in [0, 1]$; find parameters Θ that minimize the joint loss
164

$$165 \quad \mathcal{L}_{\text{JOINT}}(\beta, \Theta) = \beta \mathcal{L}_{\text{rel}}(\Theta) + (1 - \beta) \mathcal{L}_{\text{div}}(\Theta). \quad (6)$$

166 In practice, we take relevance objective $\mathcal{L}_{\text{rel}} = \mathcal{L}_{\text{MSE}}$ and diversity objective $\mathcal{L}_{\text{div}} = -\mathcal{L}_{\text{DDRO}}$.
167 While leading to an efficiently optimizable objective (6), joining them combines two diametrically
168 opposed goals: diversity and relevance. Diverse outputs are not necessarily the most “relevant” ones,
169 and vice versa. We discuss how to deal with this balance in the following section.
170

171 4 JOINT GRADIENT-BASED TRAINING

172 In this section, we discuss how this objective can be incorporated into practical training. We propose
173 two strategies: one *direct*, by optimizing the joint loss explicitly, and one *indirect*, by using the joint
174 loss as a meta-objective to reweigh training samples, starting with the direct approach.
175

176 4.1 BALANCING RELEVANCE AND DIVERSITY

177 The first approach, called *direct diversity tuning* (DDT), optimize the joint loss in Eq. (6). By utilizing
178 the differentiability of our joint loss, we take an efficient gradient-based optimization approach
179 in which the model parameters are updated using the gradient $\nabla_\Theta \mathcal{L}_{\text{JOINT}}$. As diversity opposes
180 relevancy, we need to ensure that the optimization converges to a solution that is both relevant
181 and diverse. That is, if both relevance and diversity objectives have gradients that point in similar
182 directions, we say that they are ‘aligned’. If the two gradient directions are aligned, any β decreases
183 both terms. However, when the gradients $g_{\text{rel}} = \nabla \mathcal{L}_{\text{rel}}$ and $g_{\text{div}} = \nabla \mathcal{L}_{\text{div}}$ are misaligned, any linear
184 combination will necessarily favor one objective at the expense of the other. We want to ensure a
185 descent that ensures a good diversity and accuracy balance for which we realign the gradients. For
186 this, we adaptively compute the optimal balance parameter β^* that ensures simultaneous descent
187 during optimization. Formally, for gradient norms $a = \|g_{\text{rel}}\|$ and $b = \|g_{\text{div}}\|$, we denote the cosine
188 similarity by $\rho = \langle g_{\text{rel}}, g_{\text{div}} \rangle / (ab)$. The combined gradient is $g_\beta = \beta g_{\text{rel}} + (1 - \beta) g_{\text{div}}$. For a step
189 along $-g_\beta$ to decrease both objectives simultaneously, the following conditions
190

$$192 \quad -\delta_{\text{rel}}(\beta) > 0 \iff \langle g_{\text{rel}}, g_\beta \rangle > 0, \quad -\delta_{\text{div}}(\beta) > 0 \iff \langle g_{\text{div}}, g_\beta \rangle > 0 \quad (\text{A-B})$$

193 must be satisfied (Désidéri, 2012). Equivalently, the projections of the combined gradient g_β onto
194 g_{rel} and g_{div} should be positive. This guarantees that a step along $-g_\beta$ decreases both losses at once.
195 We give the feasible region of β that satisfies (A-B):
196

197 **Lemma 3** (Common descent (Désidéri, 2012)). For any $a, b > 0$ and $\rho \in [-1, 1]$, the feasible region
198 of $\beta \in [0, 1]$ satisfying (A-B) is

- 199 1. If $\rho > 0$ (aligned), all $\beta \in [0, 1]$ are feasible.
- 200 2. If $\rho = 0$ (orthogonal), the feasible set is $\beta \in (0, 1)$.
- 201 3. If $\rho < 0$ (opposing), the feasible set is $\beta \in \left(\frac{b|\rho|}{a+b|\rho|}, \frac{b}{b+a|\rho|} \right)$.

202 *Proof sketch.* The result follows from expanding the directional derivatives $\delta_{\text{rel}}(\beta) = -\beta a^2 - (1 - \beta) ab\rho$ and $\delta_{\text{div}}(\beta) = -(1 - \beta) b^2 - \beta ab\rho$, and solving the inequalities (A-B) in the three cases
203 $\rho > 0$, $\rho = 0$, and $\rho < 0$. Full details are provided in Appendix A. \square
204

205 Even when β lies in the feasible interval, different values may still lead to unbalanced overall progress,
206 with one loss improving much more than the other. To avoid this, we select the parameter
207

$$210 \quad \beta^* = \arg \max_{\beta \in [0, 1]} \min \{-\delta_{\text{rel}}(\beta), -\delta_{\text{div}}(\beta)\},$$

211 that maximizes the minimum per-step decrease, yielding a *multiple gradient descent algorithm*
212 (MGDA) (Désidéri, 2012) specialized to two objectives. MGDA finds a convex combination of
213 gradients that minimizes the maximum directional derivative across tasks using the closed-form
214

$$215 \quad \beta^* = \frac{b(b - a\rho)}{a^2 + b^2 - 2ab\rho}. \quad (7)$$

216 When $\rho \leq 0$, the ideal balance β^* lies in the feasible interval described in Lemma 3, ensuring
 217 valid common descent. Intuitively, β^* equalizes the first-order decreases of relevance and diversity,
 218 yielding a balanced update. Finally, when β_t is chosen adaptively by the MGDA rule, we obtain
 219 convergence to Pareto-stationary solutions.

220 **Corollary 4** (Pareto-stationarity with adaptive β_t (Sener & Koltun, 2018)). *Under diminishing step*
 221 *sizes, if each β_t is chosen by the minimax rule (or projected variant), then every accumulation point*
 222 *of $\{\Theta^t\}$ is Pareto-stationary for $(\mathcal{L}_{\text{rel}}, \mathcal{L}_{\text{div}})$.*

224 While the adaptive choice provides balanced progress at each step, a fixed β remains useful, as it
 225 avoids per-iteration computation. To contextualize our results, we introduce an additional alternative
 226 approach which operates with a fixed β next. [We provide details, incl. pseudocode, in Appendix B.](#)

228 4.2 OPTIMIZING DIVERSITY BY EXAMPLE REWEIGHTING.

230 A complementary line of work on machine learning fairness and robustness has shown that *reweighting*
 231 *training examples* can effectively mitigate inherent data bias. The central idea is to retain the standard
 232 prediction objective, but assign *adaptive weights* to individual samples so that the resulting model
 233 better aligns with a criterion (Ren et al., 2018; Rastegarpanah et al., 2019). This approach is naturally
 234 formulated as meta-learning: the inner loop minimizes a weighted relevance loss, while the outer
 235 loop adjusts weights using a meta-objective that encodes the criterion.

236 Inspired by this paradigm, we aim to improve diversity by reducing bias at the *data level* through
 237 per-sample weights. Intuitively, increasing the importance of ‘minor’ items increases the chance of
 238 their exposure. In brief, the idea is to learn the model parameters by optimizing a weighted relevance
 239 loss, while simultaneously learn the weights that balance relevance and diversity. More concretely,
 240 our *meta-diversity reweighting* (MDR) introduces a weight $w_{u,i} \in [0, 1]$ for each user-item pair (u, i)
 241 in a mini-batch \mathcal{B} and optimizes the *reweighted relevance loss*

$$242 \mathcal{L}_{\text{MSE}}^w = \sum_{(u,i) \in \mathcal{B}} w_{u,i} (\mathbf{R}_{u,i} - \tilde{\mathbf{R}}_{u,i})^2. \quad (8)$$

244 The algorithm starts with obtaining a temporary model Θ' (initializing $w_{u,i} = 0$) using a one-step
 245 inner update. We then re-evaluate the predictions with the updated parameters and compute the joint
 246 meta-loss $\mathcal{L}_{\text{JOINT}}$. Next, we compute the gradient of $\mathcal{L}_{\text{JOINT}}$ with respect to w to obtain a utility
 247 score for each sample, and finally normalize the weights of all data points in \mathcal{B} so that they sum to 1.
 248 The joint relevance–diversity objective is used only as a *meta-loss* to update the weights w , and is
 249 never applied directly to model parameters, as detailed in Apx. B.

250 In contrast to direct diversity-guided tuning, which modifies the training objective itself, MDR pre-
 251 serves the standard relevance-oriented loop while implicitly reshaping the effective data distribution.
 252 The meta-objective encourages weights that downplay biased interactions and upweight samples that
 253 contribute to both accuracy and diversity. This reweighting perspective allows us to test whether
 254 diversity gains can be obtained not only by altering the optimization objective, but also by correcting
 255 data imbalance through implicit meta-optimization.

256 We evaluate both DDT and MDR in two settings: *from-scratch training*, where relevance and diversity
 257 are optimized jointly, and *fine-tuning*, where a relevance-trained model is adapted for diversity.

260 5 EXPERIMENTAL EVALUATION

262 We describe the experimental setup to evaluate the effectiveness of our proposed solutions, introducing
 263 the datasets and evaluation criteria in this section, and provide further details in Appendix C and D.

264 **Datasets.** We evaluate our methods across three domains: *entertainment*, *product*, and *social*
 265 recommendations. For entertainment, we use MovieLens, Netflix, and Yahoo-R2, which contain user
 266 ratings on movies or music with genre/category annotations. For product recommendation we use
 267 the Coat dataset, which includes user ratings and product attributes. For social recommendation we
 268 use KuaiRec, a large-scale mobile video dataset with watch-time based ratings. The item affinity
 269 matrices \mathbf{S} are pre-computed by Jaccard similarity scores based on genre/category information of
 items. Basic statistics and detailed preprocessing are given in the Appendix C.

270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
Table 1: Statistics of datasets and diversity metrics.

Dataset	$ \mathcal{U} $	$ \mathcal{I} $	$ \Omega $	$\text{DRO}(\mathcal{I})$	$ \mathcal{C} $
Coat	290	300	6 960	0.73	33
KuaiRec	1 411	3 327	4 676 570	0.91	31
Netflix	4 999	1 112	557 176	0.83	27
Yahoo-R2	4 050	5 000	684 782	0.26	58
MovieLens	6 040	3 706	1 000 208	0.83	18

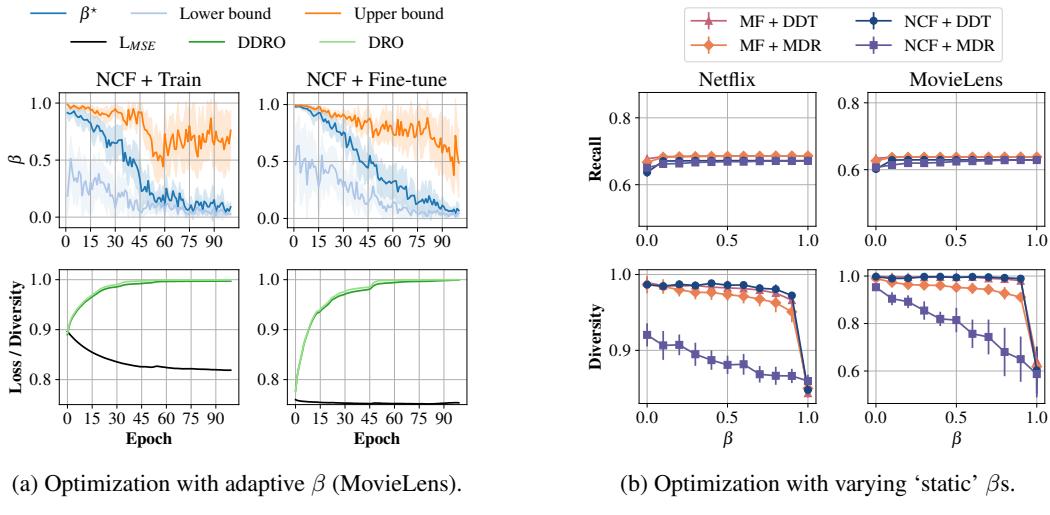


Figure 1: Optimization of relevance-diversity joint objective. Figure 1a shows results of training and fine-tuning of NCF models with adaptive β on the MovieLens dataset. Figure 1b shows the performance comparison of DDT and MDR applied to two models (NMF and NCF) across two datasets with varying fixed $\beta \in [0, 1]$.

Baselines. We compare against relevance- and diversity-aware approaches: (i) *Non-negative Matrix Factorization* (NMF) (Lee & Seung, 1999) as a relevance-only baseline; (ii) greedy post-processing methods such as *Maximal Marginal Relevance* (MMR) (Carbonell & Goldstein, 1998) and *Diversity-weighted Utility Maximization* (DUM) (Ashkan et al., 2015); (iii) probabilistic diversification via *Determinantal Point Processes* (DPP) (Chen et al., 2018b); and (iv) the graph-based *Diversified GNN Recommender* (DGRec) (Yang et al., 2023). We use the authors’ publicly-available implementations.

Evaluation criteria. We evaluate accuracy using *hit rate*, *precision*, and *recall*. For each user u , let \mathcal{R}_u denote the top- k recommended items and \mathcal{T}_u the set of ground-truth relevant items (i.e., rated above a threshold), the *Hit rate* measures whether at least one relevant item appears in \mathcal{R}_u ; *precision* is the fraction of items in \mathcal{R}_u that are in \mathcal{T}_u ; and *recall* is the fraction of relevant items in \mathcal{T}_u that are retrieved in \mathcal{R}_u . When ground-truth is unknown, we report the estimated user satisfaction using preference likelihoods $p(\tilde{\mathbf{R}}_{u,i}) \in [0, 1]$ and threshold $\tau = 0.8$ as the *relevance score*

$$\text{Relevance}(u) = \frac{1}{k} \sum_{i \in \mathcal{R}_u} \mathbb{I}[p(\tilde{\mathbf{R}}_{u,i}) > \tau]. \quad (9)$$

5.1 EXPERIMENTAL ANALYSIS

Having introduced our setup, we now introduce our research questions and experimental analysis.

- 321 **Q1** Does adaptive β behave as predicted by our theory?
- 322 **Q2** How does the choice of β affect the accuracy–diversity trade-off?
- 323 **Q3** How do our approaches compare with established diversification methods?
- 324 **Q4** How do diversity gains evolve with varying k ?

324
 325 Table 2: Performance of DDT, MDR, and five competitors in terms of diversity and relevance across
 326 Coat, Yahoo-R2, Netflix, MovieLens, and KuaiRec. We highlight the the best results with **bold**, and
 327 underline the second best, reporting mean and standard deviation over 10 trials.

Algorithm	Coat		Yahoo-R2		Netflix		MovieLens		KuaiRec	
	Diversity	Relevance								
NMF	0.77 (0.02)	0.41 (0.04)	0.09 (0.05)	0.76 (0.02)	0.84 (0.01)	0.88 (0.03)	0.62 (0.08)	0.98 (0.00)	0.89 (0.01)	0.84 (0.01)
MMR	0.80 (0.01)	0.40 (0.04)	0.80 (0.06)	0.68 (0.03)	0.93 (0.01)	0.85 (0.04)	0.94 (0.03)	0.95 (0.01)	0.99 (0.00)	0.76 (0.01)
DUM	0.81 (0.01)	0.31 (0.04)	0.98 (0.02)	0.60 (0.02)	0.93 (0.00)	0.71 (0.03)	0.93 (0.01)	0.91 (0.01)	0.98 (0.00)	0.42 (0.01)
DPP	0.81 (0.01)	0.39 (0.04)	1.00 (0.00)	0.58 (0.01)	0.96 (0.00)	0.84 (0.04)	0.98 (0.01)	0.95 (0.01)	1.00 (0.00)	0.75 (0.01)
DGRec	0.71 (0.01)	0.69 (0.02)	0.33 (0.01)	0.83 (0.02)	0.76 (0.00)	0.83 (0.01)	0.73 (0.01)	0.47 (0.02)	0.91 (0.02)	0.18 (0.04)
DDT	0.83 (0.01)	0.50 (0.05)	0.98 (0.01)	0.85 (0.01)	0.98 (0.00)	0.97 (0.01)	1.00 (0.00)	1.00 (0.00)	0.98 (0.02)	0.95 (0.00)
MDR	0.82 (0.01)	0.47 (0.03)	0.86 (0.09)	0.82 (0.02)	0.98 (0.01)	0.93 (0.02)	0.97 (0.02)	0.99 (0.00)	0.97 (0.01)	0.85 (0.01)

334
 335 In the following, we experimentally answer all these questions in detail.

336
 337 **Evaluating adaptive β optimization.** We first study the direct optimization approach using both
 338 *from-scratch training* and *fine-tuning* of a relevance-pretrained model. At each step, we adopt the
 339 adaptive β rule, computing both the optimal coefficient β^* and the feasible interval from Lemma 3.
 340 Figure 1a shows from-scratch training and fine-tuning for 100 epochs on MovieLens. We employ the
 341 *neural collaborative filtering* (NCF) model with an MLP architecture, and consider top-10 diversity.
 342

343 In the top row of Figure 1a, we observe that the adaptive coefficient β^* remains consistently within
 344 the feasible interval defined by Lemma 3. We see how β^* gradually shifts from favoring relevance
 345 toward favoring diversity, confirming that the adaptive rule preserves a common descent direction.
 346 The bottom row reports the optimization trajectories. The relevance loss (\mathcal{L}_{MSE}) decreases steadily
 347 while diversity (DRO and DDRO) increases, showing that adaptive β achieves a balanced trade-off
 348 in both training objectives. We see that fine-tuning preserves the relevance of the pretrained model,
 349 whereas from-scratch converges to a worse stationary point.

350 We further observe that fine-tuning preserves the relevance performance of the pretrained model,
 351 whereas training from scratch reduces \mathcal{L}_{MSE} but converges to a worse stationary point. In contrast,
 352 both approaches reach a similar level of diversity, suggesting that regularization may hinder the
 353 convergence towards more ‘relevant’ models during from-scratch optimization. Fine-tuning, on the
 354 other hand, starts with a relevance-optimized model, which we effectively tune for diversity. Finally,
 355 because DRO and DDRO are close, we empirically validated the accuracy of our relaxation. We
 356 discuss the impact of the parameter ϵ on approximation accuracy later. Overall, we saw that adaptive
 357 β effectively delivers simultaneous improvements in relevance and diversity.

358 **Relevance-diversity trade-off.** We next examine how the choice of static β affects relevance and
 359 diversity. Figure 1b reports recall and diversity for DDT and MDR with MF and NCF on Netflix
 360 and MovieLens. Starting from a relevance-trained model, we fine-tune with the joint objective (6)
 361 while varying $\beta \in [0, 1]$, ranging from $\beta = 1$ (pure relevance) to $\beta = 0$ (pure diversity). We
 362 run 10 random trials for each experiment and report the mean and variance. We observe a stable
 363 low-variance recall across a wide range of β , showing that introducing diversity does not substantially
 364 compromise accuracy with fixed β . A noticeable drop occurs only when β approaches zero, where
 365 the objective focuses almost exclusively on diversity. In contrast, diversity improves as β decreases.
 366 Moreover, DDT consistently outperforms MDR in terms of diversity with a neural network model,
 367 while the two strategies behave more similarly with a matrix factorization model, suggesting that
 368 reweighting is particularly beneficial for simpler low-rank models. These findings demonstrate that
 369 fixed- β optimization offers a practical means to explore the relevance–diversity trade-off, enabling
 370 substantial diversity gains with minimal loss of accuracy. Similar patterns are observed for MSE loss,
 371 hit rate, and precision, with detailed results on additional datasets provided in Appendix D.

372 **Diversity-relevance performance of all approaches.** Next, we compare the proposed methods
 373 against alternative diversification approaches across all datasets. Fixing $k = 10$ and $\beta = 0.2$, we
 374 fine-tune a pre-trained relevance-optimized matrix factorization model for 100 epochs and select
 375 the most diverse checkpoint. In Table 2 we report the mean and standard deviation of diversity and
 376 relevance scores for all methods from 10 trials.

377 We see that our direct diversity tuning approach (DDT) demonstrates strong performance in both
 378 relevance and diversity, achieving either the best or second-best results across nearly all metrics and
 379 datasets, showing the effectiveness of jointly optimizing. Similarly, we see that the implicit data

378
 379
 380
 381
 382
 383 Table 3: Diversity gain achieved by Direct Diversity Tuning (DDT) on top-1 $\sim k$ and $k+1 \sim 2k$
 384 recommendations across five datasets Coat, Yahoo-R2, Netflix, MovieLens, and KuaiRec. We vary
 385 the diversity reward parameter $k \in \{5, 10, 20, 30, 40\}$ and apply DDT to pre-trained NMF and NCF,
 386 reporting diversity gain. We report the mean and standard deviation across 10 runs.
 387
 388

Dataset	Diversity gain (%)										
	1 $\sim k$					k+1 $\sim 2k$					
	k=5	k=10	k=20	k=30	k=40	k=5	k=10	k=20	k=30	k=40	
NMF	Coat	10.9 (4.3)	5.4 (1.8)	3.5 (1.7)	2.9 (1.1)	2.6 (0.8)	-0.5 (4.7)	-2.5 (4.4)	-1.8 (1.3)	-1.5 (1.7)	-1.9 (0.8)
	Yahoo-R2	75.0 (14.3)	77.4 (8.7)	78.5 (3.8)	78.4 (2.4)	77.2 (1.6)	54.9 (11.2)	50.8 (9.9)	47.1 (2.7)	45.2 (2.1)	41.9 (1.3)
	Netflix	14.7 (2.3)	13.8 (1.0)	11.3 (0.7)	10.2 (0.6)	9.6 (0.5)	7.8 (2.6)	4.2 (1.7)	3.1 (2.1)	2.4 (2.2)	2.2 (1.8)
	MovieLens	41.4 (13.1)	37.2 (8.3)	28.4 (4.6)	23.8 (3.2)	21.6 (2.6)	25.6 (10.4)	16.2 (5.1)	10.1 (2.8)	8.3 (2.3)	6.8 (1.6)
	KuaiRec	9.2 (1.7)	7.2 (0.5)	9.4 (0.8)	10.6 (0.7)	10.9 (0.8)	5.5 (1.5)	10.3 (2.1)	6.0 (2.0)	3.3 (1.0)	1.4 (1.4)
NCF	Coat	9.7 (2.4)	1.2 (1.8)	6.9 (1.3)	0.3 (0.8)	4.8 (0.5)	-0.9 (0.6)	3.5 (0.4)	-1.1 (0.7)	2.7 (0.5)	-1.0 (0.5)
	Yahoo-R2	79.5 (12.8)	9.4 (17.5)	79.3 (7.6)	4.6 (6.0)	78.2 (8.7)	4.9 (13.7)	77.5 (4.6)	0.6 (9.0)	77.0 (6.3)	1.7 (7.3)
	Netflix	13.7 (2.0)	8.2 (1.9)	13.3 (0.9)	4.5 (1.4)	10.8 (0.7)	3.9 (1.2)	9.5 (0.5)	3.6 (1.1)	9.0 (0.5)	3.2 (1.1)
	MovieLens	46.8 (13.0)	19.6 (6.6)	36.2 (7.8)	12.7 (3.0)	26.9 (4.5)	10.9 (1.9)	22.9 (3.0)	9.6 (1.5)	20.9 (2.5)	9.0 (1.5)
	KuaiRec	8.8 (1.3)	5.4 (1.8)	6.5 (0.3)	10.7 (2.6)	8.7 (0.7)	8.0 (1.2)	10.6 (0.7)	4.2 (0.9)	11.1 (0.5)	2.3 (0.9)

393
 394 reweight approach (MDR) shows highly competitive performance. Its performance is particularly
 395 notable on Netflix, MovieLens and KuaiRec, where it approaches or matches the performance of
 396 DDT. This supports our earlier observation that implicit, data-driven reweighting can offer strong
 397 benefits. On the other hand, while post-hoc diversification of NMF using MMR, DUM, and DPP,
 398 considerably improve top- k diversity, they often notably reduce relevance, aligning with the results
 399 reported in previous studies (Chen et al., 2017). For example, DPP occasionally achieves the highest
 400 diversity—particularly on Yahoo-R2 and KuaiRec—but at a cost of low relevance. Greedy methods
 401 like MMR and DUM yield moderate diversity improvements but underperform in relevance. DGRec
 402 excels in diversity on certain datasets but suffers from severe relevance degradation, especially on
 403 KuaiRec, with a diversity of 0.91 and a relevance of 0.18. In contrast, DDT and MDR maintain a
 404 significantly better trade-off. In summary, fine-tuning with direct and implicit methods both achieve
 405 superior relevance-diversity trade-offs across the board, outperforming the competitors.

406 **Diversity gain with varying k .** We examine how diversity changes across the growing number
 407 of recommended items. Starting with relevance-optimized matrix factorization and neural network
 408 models, we fine-tune with the joint objective $\mathcal{L}_{\text{JOINT}}$ with DDT for 10 epochs at a fixed $\beta = 0.5$,
 409 and vary $k \in \{5, 10, 20, 30, 40\}$, selecting the best results. Since larger top- k sets approach coverage
 410 of the entire item space, it becomes impossible to improve diversity beyond the dataset’s average. To
 411 account for this, we report *relative diversity gains* over the relevance-only baseline, normalized by the
 412 maximum achievable score. Specifically, we measure DRO gains as the difference between fine-tuned
 413 and pre-trained models, normalized to lie in $[0, 1]$ and reported as percentages. We distinguish
 414 between *in-objective* gains (top- k) and *out-of-objective* gains (the subsequent $k+1 \sim 2k$ items).

415 In Table 3, we observe positive *in-objective* diversity gains across all datasets. In Coat, Netflix, and
 416 MovieLens, we see a decreasing gain as k increases. The improvement is especially pronounced on
 417 Yahoo-R2, with diversity gains exceeding $78.5 \pm 3.8\%$ in NMF and 79.3 ± 7.6 in NCF at $k = 20$,
 418 due to its low initial diversity (e.g., 0.09 in the base model). Substantial gains are also observed on
 419 MovieLens (e.g., $41.4 \pm 13.1\%$ and $46.8 \pm 13\%$ at $k = 5$) and Netflix, demonstrating the effectiveness
 420 of our method across both sparse and dense recommendation scenarios.

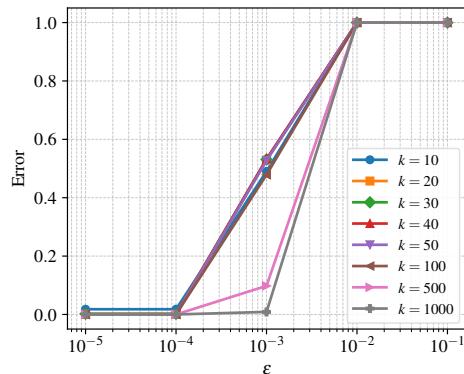
421 **Generality beyond the optimized range.** We next ask whether the benefits of diversity optimization
 422 extend beyond the explicitly encoded objective. To this end, we evaluate diversity not only at the
 423 target top- k set but also in the subsequent $k+1 \sim 2k$ items in Table 3. We observe broadly positive
 424 *out-of-objective* gains, which again diminish as k increases. This effect is especially strong on
 425 Yahoo-R2 and MovieLens, showing that optimization reshapes the ranking itself: improvements are
 426 not limited to the items directly optimized, but generalize to deeper parts of the recommendation list.
 427 In contrast, results on Coat fluctuate around zero, likely due to its small scale and limited item pool.
 428 These findings highlight the generality of our diversity-guided optimization: it improves diversity
 429 not only within the objective’s explicit target (top- k) but also beyond, demonstrating that the learned
 430 updates capture a broader notion of diversity than what is directly optimized.

431 **Approximation error with vary ϵ .** In Figure 2, we report the approximation error between the
 432 exact diversity DRO and its differentiable surrogate DDRO as the top- k size varies. We define the
 433 error as $|\mathcal{L}_{\text{DRO}} - \mathcal{L}_{\text{DDRO}}|/\mathcal{L}_{\text{DRO}}$. For small k (e.g., $k < 100$), setting $\epsilon < 10^{-4}$ already yields a

432

433
434
435
436
Table 4: Comparison of recommendations
lists for a user from DDT and NMF. Movies
appearing in both lists are **bold**.

	Rank	Genres	Movie Title
DDT	1	Horror	Vampyros Lesbos (Las Vampiras)
	2	Thriller	The Spiral Staircase
	3	War	Prisoner of the Mountains
	4	Animation, Musical	Melody Time
	5	Documentary	The Gate of Heavenly Peace
	6	Crime	Lured
	7	Drama	Mamma Roma
	8	Comedy	Smashing Time
	9	Adventure	Ulysses (Ulisse)
	10	Romance	Persuasion
NMF	1	Drama	Mamma Roma
	2	Drama	Foreign Student
	3	Drama	The Apple
	4	Drama, Romance	Leather Jacket Love Story
	5	Comedy	Smashing Time
	6	Documentary	The Gate of Heavenly Peace
	7	Documentary	Modulations
	8	Comedy, Romance, War	Forrest Gump
	9	Drama	Schlafes Bruder (Brother of Sleep)
	10	Drama, War	Schindler's List

443
444
445
446
447
448
449
450
451
452
453
close approximation. As k increases, however, a smaller ϵ is required to maintain the same accuracy. This highlights that the choice of ϵ is critical: an improper value can lead to arbitrarily large errors, rendering the diversity measure unreliable for optimization.454
455
456
457
458
459
460
461
Case study. To demonstrate the utility of our approaches, we examine the recommendations generated from an NMF model and its diversity-optimized variant. In Table 4, we observe that NMF produces homogeneous items where six out of ten items are labeled as *drama*, with remaining entries only marginally extending into *romance*, *documentary*, *comedy* or *War*. In contrast, the DDT-generated list spans a much broader range of genres, including *horror*, *thriller*, *animation*, *crime*, and *adventure*. Despite diversity gains, DDT retains three of the top items in relevance-optimized recommendation (*Mamma Roma*, *Smashing Time*, and *Gate of Heavenly Peace*), which collectively represent the relevant genre themes (*drama*, *comedy*, and *documentary*, respectively), suggesting that DDT successfully preserves relevant while enhancing diversity.462
463
6 CONCLUSION464
465
466
467
468
469
470
471
472
We addressed the issue of limited top- k diversity of relevance prediction models, which contributes to echo chambers, reduced novelty, and social polarization. By integrating diversity into a gradient-based optimization, we presented a unified framework for diversity-aware recommendation by introducing a differentiable diversity objective that enables end-to-end optimization of both relevance and diversity. We proposed two complementary, model-agnostic algorithms to support explicit and implicit integration of diversity into standard recommender systems. With extensive experiments on real-world datasets we demonstrated that our methods consistently improve diversity, converge efficiently, and introduce minimal computational overhead.473
474
475
476
477
478
479
480
481
482
483
Limitations and future work. While our framework effectively promotes both relevance and diversity in top- k recommendations, several challenges remain. First, the adaptive β is applicable only for a direct joint objective optimization, while it is optimal in each step, but the optimization trajectory may converge to a compromised stationary point, which is not globally optimal. The mechanism that leads to a global optimal solution remains unexplored. Future work includes developing new ways to balance our competing goals via multi-objective formulations. Second, we primarily relies on categorical similarity to quantify diversity. While this provides interpretability, it does not capture more nuanced relationships or learned semantic embeddings. Future work extends the research into more expressive and context-sensitive diversity metrics. Finally, as diversity bias stems from the training data—implicit and explicit approaches often achieve similar performance levels—we see significant potential in the data-driven diversification, such as counterfactual data augmentation or diversity-aware sampling.484
485
Despite these research opportunities, we observe that our methods significantly increase diversity with only an imperceptible decrease in relevance.450
451
452
453
Figure 2: Approximation error of DDRO with different ϵ and k on MovieLens.

486 ETHICS STATEMENT
487488 This work is primarily theoretical and focuses on incorporating diversity into machine-learning
489 models for ranking. All datasets used in our experiments are publicly available, and no personally
490 identifiable or sensitive information was collected or processed. No user studies or interventions
491 involving human participants were conducted.492 While our approach is motivated by the goal of improving diversity in ranking outcomes, we acknowl-
493 edge that any deployment of such methods may have broader societal implications. Potential concerns
494 include including fairness considerations, potential biases in the underlying data, or unintended ef-
495 fects depending on the application context. We leave a deeper exploration of these implications to
496 future research.
497498 REPRODUCIBILITY STATEMENT
499500 Our code is available for review in an online anonymous repository.¹ It will also be publicly available
501 upon acceptance.
502503 AI USAGE DISCLOSURE
504505 Large language models (LLMs) were employed during the research phase to assist with surveying
506 related literature, including identifying and summarizing relevant papers and methods, as well as
507 sketching and testing proposed approaches described in prior work. Generative models were also used
508 to draft scripts for data processing (e.g., formatting and visualizing data, preliminary experimental
509 validation). These scripts were only used for exploration and are not part of the final experimental
510 pipeline or released codebase. In addition, ChatGPT and Grammarly were used to assist with grammar
511 and phrasing in the manuscript. All outputs from these tools were reviewed, edited, and verified by
512 the authors, who take full responsibility for the final content.
513514 REFERENCES
515516 Ashton Anderson, Lucas Maystre, Ian Anderson, Rishabh Mehrotra, and Mounia Lalmas. Algorithmic
517 effects on the diversity of consumption on spotify. In *Proceedings of The Web Conference*, pp. 2155–
518 2165. ACM, 2020. URL <https://dl.acm.org/doi/10.1145/3366423.3380281>.
519
520 Azin Ashkan, Branislav Kveton, Shlomo Berkovsky, and Zheng Wen. Optimal greedy diversity for
521 recommendation. In *IJCAI*, volume 15, pp. 1742–1748, 2015.
522
523 James Bennett and Stan Lanning. The Netflix prize, 2007.
524
525 Mathieu Blondel, Olivier Teboul, Quentin Berthet, and Josip Djolonga. Fast differentiable sorting
526 and ranking. In *Proceedings of the 37th International Conference on Machine Learning, ICML*
527 2020, 13–18 July 2020, Virtual Event, Proceedings of Machine Learning Research, pp. 950–959.
528 PMLR, 2020.
529
530 Allan Borodin, Aadhar Jain, Hyun Chul Lee, and Yuli Ye. Max-sum diversification, monotone
531 submodular functions, and dynamic updates. *ACM Transactions on Algorithms (TALG)*, pp. 1–25,
532 2017.
533
534 Xingjuan Cai, Zhaoming Hu, Peng Zhao, Wensheng Zhang, and Jinjun Chen. A hybrid recom-
535 mendation system with many-objective evolutionary algorithm. *Expert Systems with Applications*, 159:
113648, 2020.
536
537 Jaime Carbonell and Jade Goldstein. The use of mmr, diversity-based reranking for reordering
538 documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR*
539 conference on Research and development in information retrieval

1 <https://anonymous.4open.science/r/DDRO-B975>

540 Laming Chen, Guoxin Zhang, and Hanning Zhou. Improving the diversity of top-n recommendation
 541 via determinantal point process. In *Large Scale Recommendation Systems Workshop*, pp. 3, 2017.
 542

543 Laming Chen, Guoxin Zhang, and Eric Zhou. Fast greedy map inference for determinantal point
 544 process to improve recommendation diversity. *Advances in Neural Information Processing Systems*,
 545 31, 2018a.

546 Laming Chen, Guoxin Zhang, and Eric Zhou. Fast greedy map inference for determinantal point
 547 process to improve recommendation diversity. *Advances in Neural Information Processing Systems*,
 548 2018b.

549 Jean-Antoine Désidéri. Multiple-gradient descent algorithm (mgda) for multiobjective optimization.
 550 *Comptes Rendus Mathematique*, 350(5-6):313–318, 2012.
 551

552 Tommaso Di Noia, Jessica Rosati, Paolo Tomeo, and Eugenio Di Sciascio. Adaptive multi-attribute
 553 diversity for recommender systems. *Information Sciences*, 382:234–253, 2017.

554 Gideon Dror, Noam Koenigstein, Yehuda Koren, and Markus Weimer. The yahoo! music dataset and
 555 kdd-cup’11. In *Proceedings of KDD Cup 2011*, pp. 3–18. PMLR, 2012.
 556

557 Yongkang Du, Jieyu Zhao, Yijun Yang, and Tianyi Zhou. Controllable pareto trade-off between
 558 fairness and accuracy. *arXiv preprint arXiv:2509.13651*, 2025.

559 Chongming Gao, Shijun Li, Wenqiang Lei, Jiawei Chen, Biao Li, Peng Jiang, Xiangnan He, Ji-
 560 axin Mao, and Tat-Seng Chua. KuaiRec: A fully-observed dataset and insights for evaluating
 561 recommender systems. In *CIKM*, pp. 540–550, 2022.

562 F Maxwell Harper and Joseph A Konstan. The movielens datasets: History and context. *ACM*
 563 *transactions on interactive intelligent systems (TIIS)*, 2015.

564 Refael Hassin, Shlomi Rubinstein, and Arie Tamir. Approximation algorithms for maximum dispersion.
 565 *Operations Research Letters*, pp. 133–137, 1997.

566 Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural
 567 collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web*,
 568 pp. 173–182, 2017.

569 David Holtz, Ben Carterette, Praveen Chandar, Zahra Nazari, Henriette Cramer, and Sinan Aral. The
 570 engagement-diversity connection: Evidence from a field experiment on spotify. In *Proceedings of
 571 the 21st ACM Conference on Economics and Computation*, pp. 75–76, 2020.

572 Neil J Hurley. Personalised ranking with diversity. In *Proceedings of the 7th ACM Conference on
 573 Recommender Systems*, pp. 379–382, 2013.

574 Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender
 575 systems. *Computer*, pp. 30–37, 2009.

576 Riwei Lai, Li Chen, Yuhang Zhao, Rui Chen, and Qilong Han. Disentangled negative sampling for
 577 collaborative filtering. In *Proceedings of the Sixteenth ACM International Conference on Web
 578 Search and Data Mining*, pp. 96–104, 2023.

579 Anja Lambrecht and Catherine Tucker. Algorithmic bias? an empirical study of apparent gender-
 580 based discrimination in the display of stem career ads. *Management science*, pp. 2966–2981,
 581 2019.

582 Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factoriza-
 583 tion. *nature*, 1999.

584 Nikola Milojkovic, Diego Antognini, Giancarlo Bergamin, Boi Faltings, and Claudiu Musat. Multi-
 585 gradient descent for multi-objective recommender systems. *arXiv preprint arXiv:2001.00846*,
 586 2019.

587 Tien T Nguyen, Pik-Mai Hui, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. Exploring
 588 the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of
 589 the 23rd international conference on World wide web*, pp. 677–686, 2014.

594 Agyemang Paul, Zhefu Wu, Kai Liu, and Shufeng Gong. Robust multi-objective visual bayesian
 595 personalized ranking for multimedia recommendation. *Applied Intelligence*, 52(4):3499–3510,
 596 2022.

597 Bashir Rastegarpanah, Krishna P. Gummadi, and Mark Crovella. Fighting fire with fire: Using
 598 antidote data to improve polarization and fairness of recommender systems. In *Proceedings of the*
 599 *Twelfth ACM International Conference on Web Search and Data Mining*, WSDM '19, pp. 231–239,
 600 New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359405. doi:
 601 10.1145/3289600.3291002. URL <https://doi.org/10.1145/3289600.3291002>.

602 Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. Learning to Reweight Examples for
 603 Robust Deep Learning. In *Proceedings of the 35th International Conference on Machine Learning*,
 604 pp. 4334–4343. PMLR, 2018.

605 Francesco Ricci, Lior Rokach, and Bracha Shapira. Introduction to recommender systems handbook.
 606 In *Recommender systems handbook*, pp. 1–35. Springer, 2010.

607 Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims.
 608 Recommendations as treatments: Debiasing learning and evaluation. In *international conference*
 609 *on machine learning*. PMLR, 2016.

610 Ozan Sener and Vladlen Koltun. Multi-task learning as multi-objective optimization. In *Proceedings*
 611 *of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, pp.
 612 525–536, Red Hook, NY, USA, 2018. Curran Associates Inc.

613 Wenjie Wang, Fuli Feng, Xiangnan He, Liqiang Nie, and Tat-Seng Chua. Denoising implicit feedback
 614 for recommendation. In *Proceedings of the 14th ACM international conference on web search and*
 615 *data mining*, pp. 373–381, 2021.

616 Zihong Wang, Yingxia Shao, Jiyuan He, Jinbao Liu, Shitao Xiao, Tao Feng, and Ming Liu. Diversity-
 617 aware deep ranking network for recommendation. In *Proceedings of the 32nd ACM International*
 618 *Conference on Information and Knowledge Management*, pp. 2564–2573, 2023.

619 Jacek Wasilewski and Neil Hurley. Incorporating diversity in a learning to rank recommender system.
 620 In *FLAIRS*, pp. 572–578, 2016.

621 Haolun Wu, Chen Ma, Bhaskar Mitra, Fernando Diaz, and Xue Liu. A multi-objective optimiza-
 622 tion framework for multi-stakeholder fairness-aware recommendation. *ACM Transactions on*
 623 *Information Systems*, 41(2):1–29, 2022.

624 Liangwei Yang, Shengjie Wang, Yunzhe Tao, Jiankai Sun, Xiaolong Liu, Philip S Yu, and Taiqing
 625 Wang. Dgrec: Graph neural network for recommendation with diversified embedding generation.
 626 In *Proceedings of the sixteenth ACM international conference on web search and data mining*, pp.
 627 661–669, 2023.

628 Yuying Zhao, Yu Wang, Yunchao Liu, Xueqi Cheng, Charu C Aggarwal, and Tyler Derr. Fairness
 629 and diversity in recommender systems: a survey. *ACM Transactions on Intelligent Systems and*
 630 *Technology*, pp. 1–28, 2025.

631 Yong Zheng and David Xuejun Wang. A survey of recommender systems with multi-objective
 632 optimization. *Neurocomputing*, 474:141–153, 2022.

633 Yu Zheng, Chen Gao, Liang Chen, Depeng Jin, and Yong Li. Dgcn: Diversified recommendation
 634 with graph convolutional networks. In *Proceedings of the web conference 2021*, pp. 401–412,
 635 2021.

636

637

638

639

640

641

642

643

644

645

646

647

648 A PROOFS
649650 In this appendix, we provide detailed proofs for the results on the trade-off parameter β .
651652 A.1 PROOF OF LEMMA 3
653654 We minimize both
655

656
$$\mathcal{L}_{\text{rel}} := \mathcal{L}_{\text{MSE}}, \quad \mathcal{L}_{\text{div}} := -\mathcal{L}_{\text{DDRO}},$$

657

so decreasing \mathcal{L}_{div} increases diversity. Denote the gradients
658

659
$$g_{\text{rel}} = \nabla \mathcal{L}_{\text{rel}}, \quad g_{\text{div}} = \nabla \mathcal{L}_{\text{div}}, \quad g_{\beta} = \beta g_{\text{rel}} + (1 - \beta) g_{\text{div}}, \quad \beta \in [0, 1],$$

660

so that direct diversity-guided tuning (DDT) performs
661

662
$$\Theta^+ = \Theta - \eta g_{\beta}.$$

663

Let
664

665
$$a = \|g_{\text{rel}}\| > 0, \quad b = \|g_{\text{div}}\| > 0, \quad \rho = \frac{\langle g_{\text{rel}}, g_{\text{div}} \rangle}{ab} \in [-1, 1].$$

666

667 **1: Exact first-order changes.** For an infinitesimal step along $-g_{\beta}$, the first-order changes in each
668 objective are
669

670
$$\delta_{\text{rel}} = \frac{d}{dt} \mathcal{L}_{\text{rel}}(\Theta - tg_{\beta}) \Big|_{t=0} = -\langle g_{\text{rel}}, g_{\beta} \rangle,$$

671

672
$$\delta_{\text{div}} = \frac{d}{dt} \mathcal{L}_{\text{div}}(\Theta - tg_{\beta}) \Big|_{t=0} = -\langle g_{\text{div}}, g_{\beta} \rangle.$$

673 Substituting $g_{\beta} = \beta g_{\text{rel}} + (1 - \beta) g_{\text{div}}$ gives
674

675
$$\delta_{\text{rel}} = -\beta \|g_{\text{rel}}\|^2 - (1 - \beta) \langle g_{\text{rel}}, g_{\text{div}} \rangle = -\beta a^2 - (1 - \beta) ab\rho,$$

676

677
$$\delta_{\text{div}} = -(1 - \beta) \|g_{\text{div}}\|^2 - \beta \langle g_{\text{div}}, g_{\text{rel}} \rangle = -(1 - \beta) b^2 - \beta ab\rho.$$

678 **2: Common descent conditions.** We require both decreases:
679

680
$$\delta_{\text{rel}} < 0, \quad \delta_{\text{div}} < 0.$$

681

Rearranging each inequality yields the following necessary and sufficient conditions:
682

683
$$\beta a + (1 - \beta) b \rho > 0, \tag{A}$$

684

685
$$(1 - \beta) b + \beta a \rho > 0. \tag{B}$$

686 **3: Case analysis of inequalities.** We now analyze (A) and (B) systematically.
687688 *Inequality (A).*

689
$$\beta a + (1 - \beta) b \rho > 0 \iff \beta(a - b\rho) > -b\rho.$$

690

691

- If $a - b\rho > 0$, then $\beta > \frac{-b\rho}{a - b\rho}$, we need further analyze ρ .
- If $a - b\rho < 0$, then $\beta < \frac{-b\rho}{a - b\rho}$, we need further analyze ρ .
- If $a - b\rho = 0$, then the inequality holds for all β .

695 *Inequality (B).*

696
$$(1 - \beta) b + \beta a \rho > 0 \iff b + \beta(a\rho - b) > 0.$$

697

698

- If $a\rho - b > 0$, then $\beta > \frac{-b}{a\rho - b}$ (RHS negative, vacuous for $\beta \in [0, 1]$).
- If $a\rho - b < 0$, then $\beta < \frac{b}{b - a\rho}$, we need further analyze ρ .
- If $a\rho = b$, then the inequality holds for all β .

702 **4: Regime specialization.** We specialize the above formulas to the three relevant regimes of ρ .
 703

704 *Case I: Aligned gradients ($\rho > 0$).* Here the two gradients point in similar directions. We revisit
 705 inequalities (A) and (B).

706 *Condition (A).*

$$707 \quad \beta a + (1 - \beta)b\rho > 0 \iff \beta(a - b\rho) > -b\rho.$$

709 • If $a - b\rho \geq 0$, then the right-hand side is nonpositive ($-b\rho \leq 0$ since $\rho > 0$). Thus the
 710 inequality is automatically satisfied for all $\beta \in [0, 1]$; no restriction is imposed.

711 • If $a - b\rho < 0$, then the inequality gives $\beta < \frac{-b\rho}{a - b\rho}$. But since $a - b\rho < 0$ and $-b\rho < 0$, the
 712 right-hand side is positive. In fact, one can check that $\frac{-b\rho}{a - b\rho} > 1$. Therefore any $\beta \in [0, 1]$
 713 still satisfies the inequality.

715 *Condition (B).*

$$716 \quad (1 - \beta)b + \beta a\rho > 0 \iff b + \beta(a\rho - b) > 0.$$

718 • If $a\rho - b \geq 0$, then the right-hand side is increasing in β and at $\beta = 0$ equals $b > 0$, so the
 719 inequality holds for all $\beta \in [0, 1]$.

720 • If $a\rho - b < 0$, then the inequality becomes $\beta < \frac{b}{b - a\rho}$. Since $a\rho - b < 0$, the denominator
 721 $b - a\rho > 0$ and thus the right-hand side exceeds 1. Hence the condition imposes no
 722 restriction within $\beta \in [0, 1]$.

724 *Conclusion.* Both inequalities (A) and (B) are therefore automatically satisfied when $\rho > 0$. Thus

$$725 \quad \delta_{\text{rel}} < 0 \quad \text{and} \quad \delta_{\text{div}} < 0 \quad \text{for every } \beta \in [0, 1].$$

727 In other words, when gradients are aligned, any convex combination yields a valid common descent
 728 direction. At $\beta = 0$ or $\beta = 1$, one loss still strictly decreases while the other is nonincreasing; for
 729 $0 < \beta < 1$, both decrease strictly.

730 *Case II: Orthogonal gradients ($\rho = 0$).* Plugging $\rho = 0$ into (A) and (B) gives

$$732 \quad \beta a > 0 \iff \beta > 0, \quad (1 - \beta)b > 0 \iff \beta < 1.$$

$$733 \quad \implies \delta_{\text{rel}} < 0, \delta_{\text{div}} < 0 \quad \text{for all } \beta \in (0, 1).$$

735 *Case III: Opposing gradients ($\rho < 0$).* Let $\rho = -|\rho| < 0$. Then

$$737 \quad \text{From (A): } \beta > \frac{b|\rho|}{a + b|\rho|} =: L, \quad \text{From (B): } \beta < \frac{b}{b + a|\rho|} =: U.$$

739 Thus the feasible interval is

$$741 \quad \beta \in (L, U) = \left(\frac{b|\rho|}{a + b|\rho|}, \frac{b}{b + a|\rho|} \right).$$

743 This interval is nonempty whenever $|\rho| < 1$, since

$$745 \quad U - L = \frac{ab(1 - |\rho|^2)}{(b + a|\rho|)(a + b|\rho|)} > 0.$$

747 If $|\rho| = 1$ (exactly opposite gradients), then any $\beta \in (0, 1)$ still makes both $\delta_{\text{rel}}, \delta_{\text{div}} < 0$.

749 **Conclusion.** We have characterized the feasible region of β :

751 • If $\rho > 0$: every $\beta \in [0, 1]$ yields common descent.
 752 • If $\rho = 0$: every $\beta \in (0, 1)$ yields common descent.
 753 • If $\rho < 0$: feasible region is (L, U) as defined above.

755 This completes the proof of Lemma 3. □

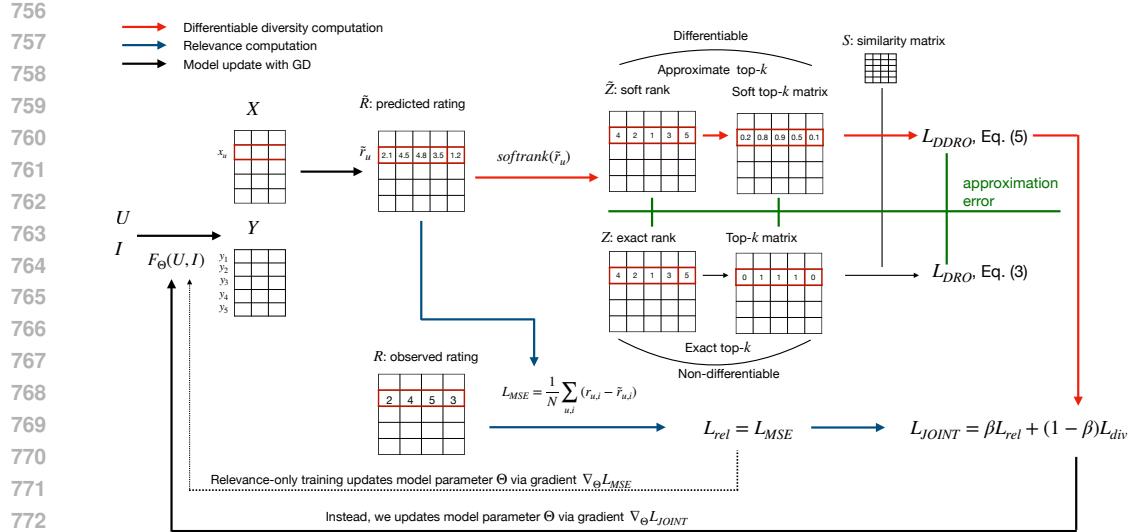


Figure 3: DDT workflow

A.2 PROOF OF β^*

We seek β^* such that $-\delta_{\text{rel}}(\beta) = -\delta_{\text{div}}(\beta)$. Expanding:

$$\beta a^2 + (1 - \beta)ab\rho = (1 - \beta)b^2 + \beta ab\rho.$$

Rearranging,

$$(a^2 + b^2 - 2ab\rho)\beta = b^2 - ab\rho.$$

Thus

$$\beta^* = \frac{b(b - a\rho)}{a^2 + b^2 - 2ab\rho}.$$

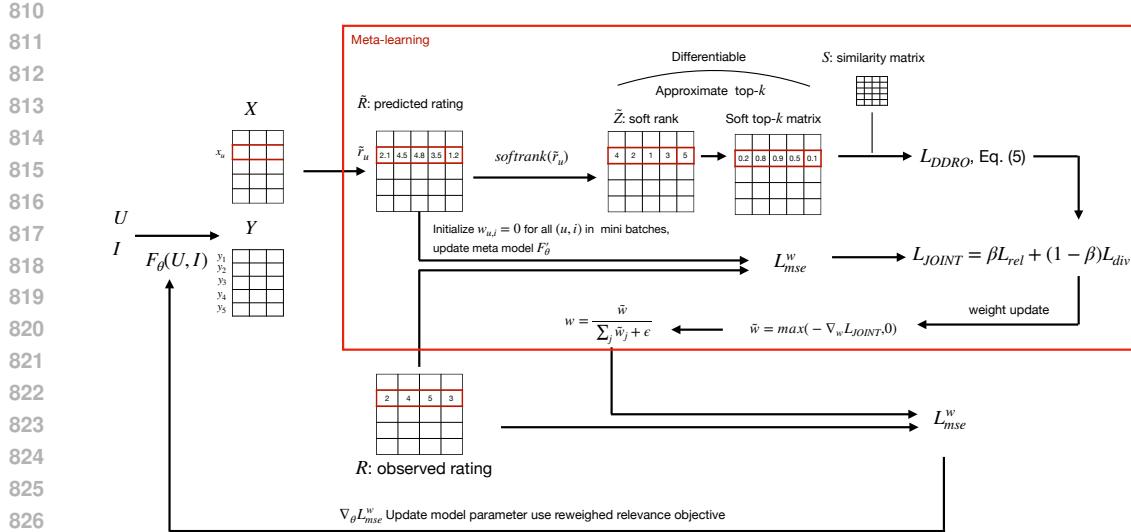
When $\rho \leq 0$, substitution verifies $L < \beta^* < U$, where (L, U) is the interval from Lemma 3. Hence β^* is feasible and maximizes the minimum first-order decrease. \square

B ALGORITHMS

Algorithm workflow diagram Figure 3 and 4 illustrate the workflow of DDT and MDR.

DDT directly optimizes the joint loss $\mathcal{L}_{\text{JOINT}} = \beta \mathcal{L}_{\text{MSE}} + (1 - \beta) \mathcal{L}_{\text{DDRO}}$ via standard gradient descent using innovative soft-ranking, top- k selection, pairwise diversity penalties, and adaptive diversity weights.

For MDR, in a nutshell, the meta learning procedure reweights using two different stages: the model stage and the design stage. While the design stage updates w using gradients $\nabla_w \mathcal{L}_{\text{JOINT}}$ from loss + diversity; the model stage updates the model Θ through gradients $\nabla_\Theta \mathcal{L}_{\text{MSE}}^w$ of the w -weighted relevance objective. In detail, we consider the outer loop to be the standard mini-batch training procedure that updates the model parameters for each mini-batch of data. For each data point in the mini-batch, we initialise the weight $w_{u,i}$ and copy the current model $\Theta^{(t)}$ as a meta-model. The inner loop is a one-step optimisation of the $\nabla_w \mathcal{L}_{\text{JOINT}}$ where $\mathcal{L}_{\text{JOINT}}$ is evaluated using the meta-model. The value of weight is the gradient direction that maximises the $\mathcal{L}_{\text{JOINT}}$. Then we rescale and normalize the weight vector to obtain the weight distribution that implicitly encodes the contribution of the data point towards diversity gain. Then the outloop updates the model parameter by the weighted MSE loss (which is relevance-only).



864 **Algorithm 1** Direct Diversity-Guided Tuning (DDT)

865

866 **Require:** Initialization model parameter Θ , training data Ω_T , item distance matrix \mathbf{D} , trade-off β ,
867 learning rate η

868 1: **for** each epoch from 1 to T **do**

869 2: **for** each mini-batch $\mathcal{B} \in \Omega_T$ **do**

870 3: $\tilde{\mathbf{R}}_{u,i} \leftarrow \mathcal{F}_\Theta(u, i)$, for all $(u, i, \mathbf{R}_{u,i}) \in \mathcal{B}$

871 4: $\mathcal{L}_{\text{MSE}} \leftarrow \text{Compute } \mathcal{L}_{\text{MSE}}(\tilde{\mathbf{R}}_{u,i}, \mathbf{R}_{u,i}; \mathcal{B}) \text{ with (1)}$

872 5: $\mathcal{L}_{\text{DDRO}} \leftarrow \text{compute DDRO with (5)}$

873 6: $\mathcal{L}_{\text{JOINT}} \leftarrow \beta \cdot \mathcal{L}_{\text{MSE}} - (1 - \beta) \cdot \mathcal{L}_{\text{DDRO}}$

874 7: $\Theta \leftarrow \Theta - \eta \cdot \nabla_\Theta \mathcal{L}_{\text{JOINT}}$

875 8: **return** Θ

876 **Algorithm 2** Meta Diversity Reweighting (MDR)

877

878 **Require:** Initialization model parameter Θ , training data Ω_T , item distance matrix \mathbf{D} , trade-off β ,
879 learning rate η

880 1: **for** each epoch from 1 to T **do**

881 2: **for** mini-batches $\mathcal{B} \in \Omega_T$ **do**

882 3: $\tilde{\mathbf{R}}_{u,i} \leftarrow \mathcal{F}_\Theta(u, i)$, for all $(u, i, \mathbf{R}_{u,i}) \in \mathcal{B}$

883 4: $\mathbf{w} \leftarrow \vec{0}$

884 5: $\mathcal{L}_{\text{MSE}}^w \leftarrow \mathcal{L}_{\text{MSE}}^w(\tilde{\mathbf{R}}_{u,i}, \mathbf{R}_{u,i}, \mathbf{w}; \mathcal{B}) \text{ with (10)}$

885 6: $\Theta' \leftarrow \Theta - \eta \cdot \nabla_\Theta \mathcal{L}_{\text{MSE}}^w$, update meta model.

886 7: $\tilde{\mathbf{R}}'_{u,i} \leftarrow \mathcal{F}_{\Theta'}(u, i)$, for all $(u, i, \mathbf{R}_{u,i}) \in \mathcal{B}$

887 8: $\mathcal{L}_{\text{JOINT}} \leftarrow \beta \cdot \mathcal{L}_{\text{MSE}}^w - (1 - \beta) \cdot \mathcal{L}_{\text{DDRO}}$

888 9: $\tilde{\mathbf{w}} \leftarrow \max(-\nabla_{\mathbf{w}} \mathcal{L}_{\text{JOINT}}, 0)$;

889 10: $\mathbf{w} \leftarrow \frac{\tilde{\mathbf{w}}}{\sum_j \tilde{w} + \epsilon}$; normalization according to Ren et al. (2018)

890 11: $\Theta \leftarrow \Theta - \eta \cdot \nabla_\Theta \mathcal{L}_{\text{MSE}}^w(\tilde{\mathbf{R}}_{u,i}, \mathbf{R}_{u,i}, \mathbf{w}; \mathcal{B})$

891 12: **return** Θ

892

893 **C DETAILED EXPERIMENT SETTING.**

894

895 **Datasets.** To evaluate our methods across different recommendation scenarios, we consider datasets
896 from three domains: entertainment, product, and social recommendations, as detailed in Tab. 1. To
897 cover the *entertainment* domain, we use Netflix (Bennett & Lanning, 2007) and MovieLens (Harper
898 & Konstan, 2015) for user–movie recommendations, as well as Yahoo-R2 (Dror et al., 2012) for
899 user–music recommendations. These datasets contain user ratings on a 5-point scale [1, 5], along
900 with genre or category annotations. For Netflix (respectively, Yahoo-R2), we randomly sample 3 000
901 items (respectively, 5 000) and retain users with at least 20 ratings (respectively, 100+ ratings). In the
902 *product recommendation* setting, we use the Coat (Schnabel et al., 2016), which captures user-coat
903 interactions in e-commerce. It contains [1, 5] ratings and item ‘meta’ attributes. Finally, for the *social*
904 *recommendation* scenario, we consider the KuaiRec (Gao et al., 2022), which is collected from a
905 mobile video-sharing platform, which includes play duration, video length, and ‘watch ratios’ from 0
906 (never watched) to 2 (twice watched), which we linearly interpolate to 5-star ratings for consistency.

907 **Baselines.** We compare our algorithm against a broad set of state-of-the-art recommender-systems
908 methods, as well as diversification techniques covering greedy, probabilistic, and graph-based
909 strategies. To study the impact of diversification, we employ classical *Non-negative MF* (NMF) (Lee &
910 Seung, 1999) as a relevance-only baseline that does not use any diversity mechanisms. We also include
911 two baselines from the post-processing family: *Maximal Marginal Relevance* (MMR) (Carbonell &
912 Goldstein, 1998) and *Diversity-weighted Utility Maximization* (DUM) (Ashkan et al., 2015), both
913 greedy diversification techniques applied on top of NMF as the underlying model.

914 MMR greedily selects top- k items that maximizes a weighted combination of relevance and dissimi-
915 larity with previously selected items. DUM, on the other hand, uses a submodular combination of
916 relevance and category-based diversity reward. *Determinantal Point Processes* (DPP) (Chen et al.,
917 2018b) estimates the likelihood of item sets to be diverse and relevant as the determinant of an
item-item similarity kernel matrix, from which we select the top- k using a greedy selection. Finally,

918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971

Table 5: Diversity gain achieved by Direct Diversity Tuning (DDT) on top-1 $\sim k$ and $k+1 \sim 2k$ recommendations across five datasets Coat, Yahoo-R2, Netflix, MovieLens, and KuaiRec. We vary the diversity reward parameter $k \in \{5, 10, 20, 30, 40\}$ and apply MDR to pre-trained NMF and NCF, reporting diversity gain. We report the mean and standard deviation across 10 runs.

Dataset	Diversity gain (%)										
	1 $\sim k$					k+1 $\sim 2k$					
	k=5	k=10	k=20	k=30	k=40	k=5	k=10	k=20	k=30	k=40	
NMF	KuaiRec	8.5 (1.6)	5.2 (2.0)	6.5 (0.6)	9.9 (1.3)	8.7 (0.8)	5.8 (1.3)	10.1 (0.7)	2.9 (0.9)	10.3 (0.8)	0.9 (0.8)
	Coat	5.2 (3.4)	2.7 (4.0)	2.3 (1.1)	-1.2 (3.7)	1.1 (1.3)	0.4 (1.2)	1.4 (0.9)	-0.4 (2.0)	1.8 (0.7)	-1.2 (1.1)
	MovieLens	36.4 (11.1)	22.0 (9.9)	32.3 (6.9)	12.1 (4.4)	25.4 (4.1)	7.3 (2.9)	21.7 (3.1)	6.8 (1.7)	19.7 (2.6)	6.4 (1.4)
	Netflix	14.2 (2.3)	4.8 (3.2)	13.0 (1.3)	2.1 (2.1)	10.3 (0.9)	1.6 (1.7)	9.2 (0.6)	1.4 (2.0)	8.7 (0.5)	1.4 (1.5)
	Yahoo-R2	49.8 (19.1)	36.0 (10.8)	59.0 (12.4)	34.2 (7.7)	65.1 (5.8)	33.1 (2.4)	66.5 (3.2)	31.4 (1.3)	66.5 (2.3)	29.6 (1.0)
NCF	KuaiRec	1.3 (1.8)	-0.0 (1.2)	0.3 (0.4)	1.2 (1.8)	1.1 (0.9)	0.9 (1.2)	1.3 (0.9)	0.2 (0.9)	1.4 (0.6)	0.0 (0.8)
	Coat	0.8 (2.3)	0.9 (1.2)	0.8 (0.9)	0.3 (0.7)	0.7 (0.9)	-0.3 (0.8)	0.3 (0.6)	-0.1 (0.3)	0.2 (0.4)	0.1 (0.6)
	MovieLens	33.2 (11.8)	6.1 (4.3)	17.8 (5.9)	3.6 (2.3)	9.5 (3.8)	1.9 (2.0)	4.9 (2.8)	1.1 (1.3)	3.4 (1.7)	0.6 (1.0)
	Netflix	4.6 (2.4)	1.4 (1.4)	2.8 (1.3)	-0.1 (1.1)	1.1 (0.9)	-0.5 (0.8)	0.4 (0.4)	-0.2 (0.3)	0.2 (0.3)	-0.0 (0.3)
	Yahoo-R2	27.0 (15.3)	-0.2 (8.8)	20.8 (10.2)	2.3 (8.3)	19.5 (8.5)	7.3 (7.1)	16.7 (8.1)	2.0 (5.4)	16.0 (6.9)	3.2 (3.2)

we include a recent embedding-based method, *Diversified GNN Recommender* (DGRec) (Yang et al., 2023), which introduces a diversity-aware aggregation mechanism into graph neural networks by selecting neighbors that maximize coverage over item categories.

Evaluation criteria. We evaluate model accuracy using metrics suitable for top- k recommendations, *hit rate*, *precision*, and *recall* metrics. For each user u , let \mathcal{R}_u denote the top- k recommended items and \mathcal{T}_u the set of ground-truth relevant items (i.e., rated above 4 in the test set). *Hit rate* measures whether at least one relevant item appears in \mathcal{R}_u ; *precision* is the fraction of items in \mathcal{R}_u that are in \mathcal{T}_u ; and *recall* is the fraction of relevant items in \mathcal{T}_u that are retrieved in \mathcal{R}_u .

However, the above requires known ground-truth. To evaluate the relevance of unseen items, we compute the potential user satisfaction as the *relevance score*

$$\text{Relevance}(u) = \frac{1}{k} \sum_{i \in \mathcal{R}_u} \mathbb{I}[p(\tilde{\mathbf{R}}_{u,i}) > \tau], \quad (11)$$

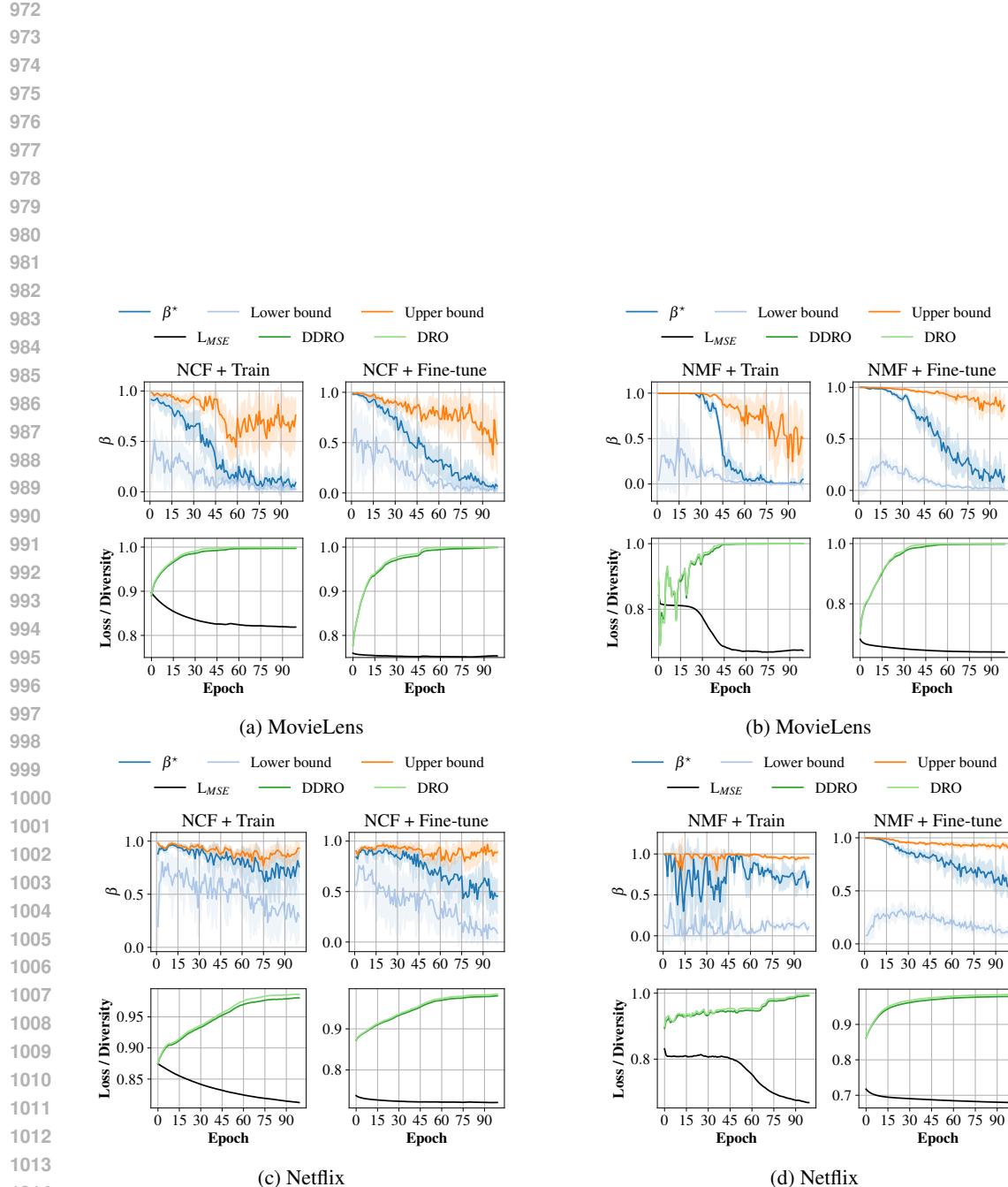
using preference likelihoods $p(\tilde{\mathbf{R}}_{u,i}) \in [0, 1]$.

Experiment environment. We conduct all our experiments on 2 AMD Epyc 7742 CPUs, 1 TB of RAM and 1 NVIDIA DGX-A100 GPU. Our code is written in Python v3.11.7. All results are averaged over 10 independent runs. We adopt the Adam optimizer during the optimization.

Detailed Parameter setting. In all experiment, we set $\epsilon = 10^{-4}$ for a exact rank approximation. During the optimization with adaptive β , we choose a learning rate $l = 0.01$ for the matrix factorization model, and $l = 0.001$ for the neural network model. We keep the Adam weight decay parameter as the default.

D ADDITIONAL EXPERIMENT

We report the additional experiment results that are omitted in the main content. Figure 5 shows the adaptive β based optimization of two datasets (MovieLens and Netflix) and two models (NMF and NCF). The patterns are consistently aligned with the analysis. Figure 6 shows the 4 accuracy metric results: recall, mean squared error, hit rate, and precision, along with the diversity reward objective score (DRO). Similar to Table 3, Table 5 presents the *in-objective* and *out-objective* diversity gain of NMF and NCF with MDR.

Figure 5: Optimization of relevance-diversity joint objective with adaptive β .

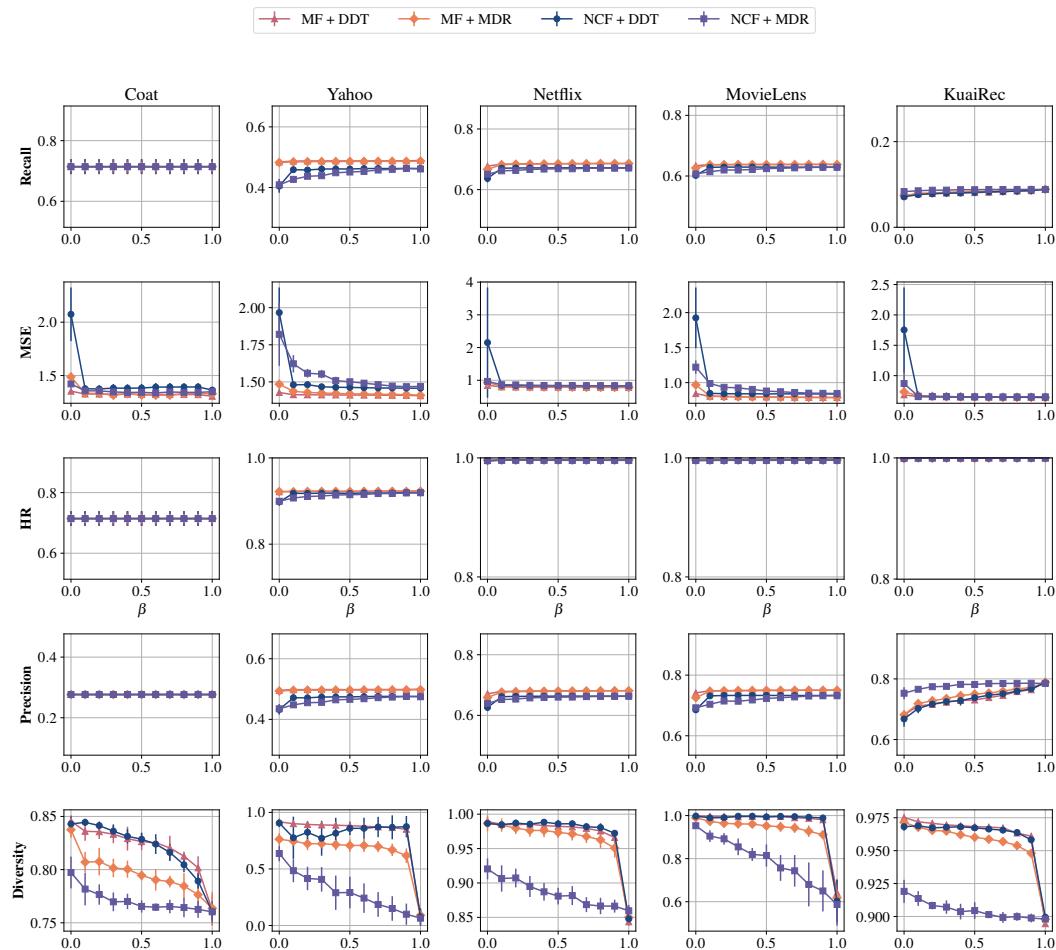


Figure 6: Performance comparison of DDT and MDR applied to two recommender models (NMF and NCF) across five datasets: Coat, Yahoo-R2, Netflix, MovieLens, and KuaiRec. We vary the parameter $\beta \in [0, 1]$ to control the trade-off between relevance and diversity, where $\beta = 1$ corresponds to optimizing only the relevance loss \mathcal{L}_{MSE} and $\beta = 0$ corresponds to optimizing only the diversity loss $\mathcal{L}_{\text{DDRO}}$. The x -axis indicates the value of β . The y -axis shows recall (top row), mean squared error (second row), hit rate (third row), precision (fourth row), and diversity reward objective (DRO) score with $k = 10$ (bottom row). In each setting, we initialize from a pre-trained model (using \mathcal{L}_{MSE} only), then fine-tune with the joint loss $\mathcal{L}_{\text{JOINT}}$ for 10 epochs, selecting the best result by diversity score. All experiments are repeated 10 times, and we report the mean and standard deviation.

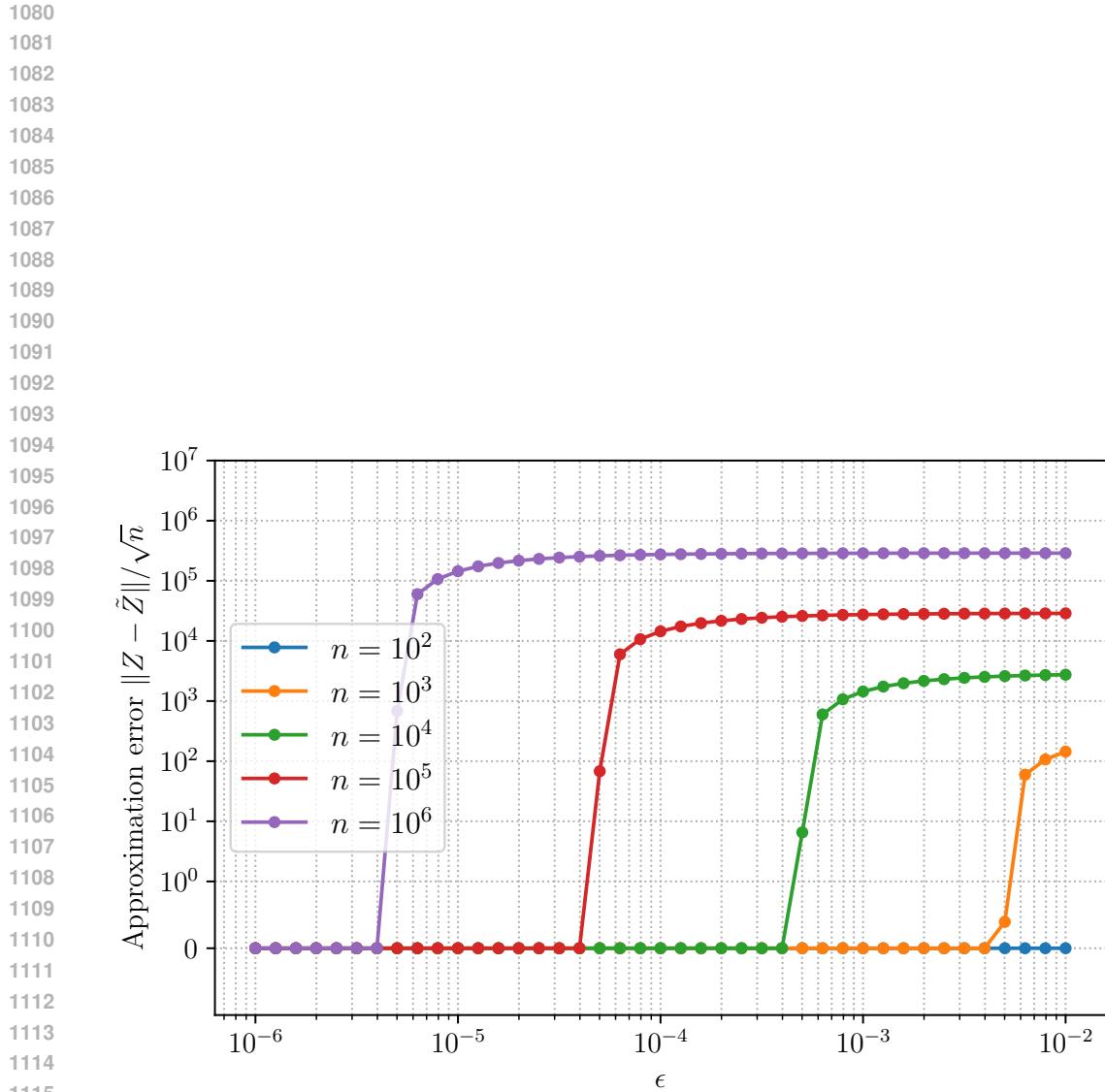


Figure 7: Rank approximation error with varying parameter ϵ and item size. The y -axis is the approximation error between exact rank and softrank approximation $\|Z - \tilde{Z}\|_2 / \sqrt{n}$, x -axis report the number of items. It clearly shows that setting $\epsilon = \frac{1}{n}$ yield zero approximation error, and there is a tunable space to achieve small error with larger ϵ .