MS-Bridge: A Deep Multi-Stakeholder Multi-Objective Recommendation Model for Two-Sided Media Platforms

Moshe Unger[1*], Pan Li[2*], Maxime C. Cohen[3], Brian Brost[4], Alexander Tuzhilin[5]

[1]Coller School of Management, Tel Aviv University, [2]Scheller College of Business, Georgia Institute of Technology, [3]Desautels Faculty of Management, McGill University, [4]Churney, [5]Stern School of Business, New York University

Abstract. Providing personalized recommendations is an important task for two-sided markets in the media industry. However, most media platforms only optimize the objectives of a single stakeholder (i.e., the user) based on the media streaming records, while ignoring other stakeholders, such as artists. We argue in this paper that it would be more beneficial to consider the objectives of multiple stakeholders simultaneously in order to improve the overall welfare of the entire ecosystem, and we formulate it as a multi-stakeholder multi-objective recommendation task. This task, however, is non-trivial and particularly challenging, since there is no well-defined multi-stakeholder utility metric in the prior literature, and the relationships between the objectives of different stakeholders have also not been systematically modeled. To that end, we propose Multi-Stakeholder Nash Social Welfare (**MS-NSW**), a novel utility metric that captures the economic realities of two-sided markets and balances the welfare of both stakeholders, leading to a series of theoretical and empirical benefits over alternative utility metrics.

Since the optimization of MS-NSW is NP-Hard, we subsequently propose **MS-Bridge**, a novel multi-objective multi-stakeholder recommendation framework to optimize MS-NSW in recommendations. Specifically, we develop a deep-learning-based "knowledge bridge" architecture that bidirectionally shares latent information across stakeholders to capture heterogeneous relationships between different objectives. The optimal location of the bridge is determined as the middle layer through a Shapley-value-based approach, which maximizes the effectiveness of the information-sharing process. The final recommendations are produced by aggregating the predicted objective values through ordinal regression or neural networks to select the optimal objective weights.

We empirically demonstrate the benefits of our method through extensive offline evaluations on four industrial-scale media streaming datasets provided by Spotify, Last.fm, and Alibaba, where we achieve significant performance improvements across all objectives for both stakeholders (users and artists) over state-of-the-art media recommendation models, multi-objective recommendation models, and multi-stakeholder recommendation models. We further conduct a simulation experiment to demonstrate the impact of our proposed model on user behaviors, where we observe that the majority of users are better-off, and that the benefits become greater in the long run. Finally, we present the economic value analysis and an interpretable case study to demonstrate that the improvements on our evaluation metrics are economically meaningful and lead to important implications for managers to develop a better media recommendation platform in practice.

Key words: Recommender systems, Multi-Stakeholder Recommendation, Transfer Learning

1. Introduction

Recommender systems (RSes) in the media industry are typically designed to optimize a series of objectives, such as click-through rate, percentage of listening of a song, and retention rate, which capture different aspects of consumer preferences (Adomavicius et al. 2011, Nguyen et al. 2017). In this context, multi-objective RSes have become prevalent in media streaming platforms, including YouTube (Zhao et al. 2019b) and Alibaba-Youku (Li and Tuzhilin 2024a).

However, in these recommender systems, the economic realities of two-sided markets are not systematically captured, as they primarily focus on the stakeholder objectives from the demand side (i.e., consumers), while overlooking the supply side (i.e., artists). The idiosyncratic properties of those markets are summarized as follows (Anderson and Gabszewicz 2006, Rysman 2009, Lin 2020, Cohen and Zhang 2022):

- *Cross-Side Network Effects.* In a two-sided market, we have two types of network effects: same-side (traditional) network effects and cross-side network effects, where the latter refers to the fact that each group's value is tied to the size and engagement *of the other*. For example, if we have more artists producing more songs on the music streaming platform, listeners will earn a higher value; conversely, if we have more listeners who spend more time listening to songs, artists will also be better off.
- *Intermediation.* The media platform in a two-sided market acts as an intermediary that facilitates transactions between the two groups of stakeholders (e.g., users and artists for Spotify), and provides infrastructure, trust, and standardization for them.

As a result of these two unique properties, we argue in this paper that the objectives of both stakeholders need to be *simultaneously taken into account* and *properly balanced* (Mehrotra et al. 2018, Malgonde et al. 2020, Abdollahpouri et al. 2020, Jannach 2022) in order to improve the overall welfare of both stakeholders. To that end, we formulate a multi-stakeholder multi-objective recommendation task for providing recommendations in the media industry. This task, however, is non-trivial and particularly challenging due to two fundamental problems in the existing approaches:

(1) Lack of Utility Metric. To the best of our knowledge, there is no well-defined multi-stakeholder utility metric for recommendations proposed in the prior literature. Instead, prior models optimize the utility of each stakeholder separately, and then aggregate the recommendations according to certain combination methods. The downside is that a recommendation can be beneficial for one stakeholder but harmful for the other. For example, recommending popular songs to the user could perform well in terms of user satisfaction at Spotify, but may not be desirable for new artists, who are trying to get exposure to a larger audience base, therefore, potentially, adversely affecting the overall business performance of the company.

(2) Heterogeneous Objective Relationship. The relationships between objectives across multiple stakeholders are usually heterogeneous, complex, dynamically changing, and at times conflicting, making the task of properly balancing them in recommendations a particularly challenging one. For example, optimizing primarily for the Click-Through Rate objectives at YouTube will lead to a "clickbait" trap, which may drive short-term benefits while destroying long-term user trust and content quality, resulting in the decrease of overall welfare for the company.

As a result of these two challenges, existing methods lead to only suboptimal recommendation performance in practice, and the task of constructing viable multi-stakeholder multi-objective recommender systems for the two-sided markets in the media industry is still open. We also present a motivational example in Section 3.2 to demonstrate the failure of existing utility metrics, and the necessity of a new one.

To address these challenges, in this paper, we propose a new utility metric **MS-NSW** (Multi-Stakeholder Nash Social Welfare), specifically designed for providing multi-objective multi-stakeholder recommendations, which captures economic realities of the two-sided markets. The MS-NSW metric is grounded in both the Multi-Attribute Utility Theory and the Nash Social Welfare theory, and we demonstrate through empirical analysis that optimizing it effectively balances the welfare of both stakeholders in recommendations. We also demonstrate through theoretical analysis that this MS-NSW metric is the unique continuous welfare function that satisfies four economic axioms for two-sided media platforms simultaneously, and that it is theoretically guaranteed to achieve Pareto Dominance over alternative metrics under certain conditions.

Moreover, since the optimization of the MS-NSW metric is an NP-Hard problem and there is no trivial solution, we also present a novel deep learning model **MS-Bridge** to optimize MS-NSW in recommendations. Specifically, MS-Bridge captures the intermediation effect by considering the objectives of both stakeholders in the two-sided market, rather than only focusing on the welfare of one stakeholder. In addition, MS-Bridge captures the cross-side network effect through a “Latent Knowledge Bridge” structure to bidirectionally transfer latent information between different stakeholders. Thus, when optimizing the objectives of one stakeholder, we incorporate the welfare of the other stakeholder by passing appropriate information over that bridge. The optimal location of the bridge is determined using a Shapley-value-based approach, which maximizes the effectiveness of the information-sharing process. The final media recommendations are produced by aggregating the predicted objective values through a neural network-based approach to select the optimal objective weights. As a result, MS-Bridge produces recommendations that effectively *balance* the needs of both stakeholders and significantly improve the welfare of both groups, as validated through both theoretical analyses and empirical evaluations in this paper.

To demonstrate the benefits of our proposed framework, we apply it to multiple media recommendation tasks on four industrial-scale datasets: two provided by Spotify, one by Last.fm, and another one by Alibaba. We consider the list of objectives that are consistent with the common industrial practices of media streaming platforms (Mehrotra et al. 2018, Li and Tuzhilin 2024b). Evaluation results illustrate significant improvements of our proposed method over the state-of-the-art media recommendation models (that focus only on one stakeholder) and multi-stakeholder or multi-objective recommendation models across a wide range of evaluation metrics. We have also conducted a comprehensive set of ablation studies to demonstrate the validity of each design choice in our proposed framework. In addition to the traditional offline evaluation framework, we also conduct a simulation study to examine the impact of our proposed model on user behaviors, where we observe that the majority of users and artists are better off after the adoption, and that the benefits become greater in the long run through iterative interactions with our recommender system. Furthermore, we present an interpretable case study to demonstrate that the improvements on our evaluation metrics are economically meaningful and that they lead to important implications for managers to develop a better media recommendation platform in practice. Thus, we demonstrate the importance of simultaneously taking into account the interests of multiple stakeholders with our proposed model to provide balanced and effective recommendations for the media industry.

In this paper, we make the following contributions. First, we demonstrate the benefits of considering the objectives of all stakeholders (instead of only one) when providing recommendations for two-sided media markets, and we formulate it as a multi-stakeholder multi-objective recommendation task accordingly. Second, we propose a new MS-NSW utility metric to capture the economic realities of two-sided markets and to balance the welfare of both stakeholders in recommendations. Third, we propose a novel MS-Bridge recommendation framework to optimize the MS-NSW metric and provide media recommendations, where we construct a "Latent Knowledge Bridge" architecture to transfer the information of different stakeholders' objectives bidirectionally. The optimal location of the bridge is automatically determined through a Shapley-value-oriented mechanism to ensure the efficiency and effectiveness of the information-sharing process, while the optimal objective weights are determined through ordinal regression for balancing between different stakeholders' objectives. Finally, through extensive theoretical analyses, offline evaluations, ablation studies, simulation studies, and interpretable case studies, we demonstrate that our proposed model significantly improves prediction and recommendation performance over state-of-the-art baseline models, and that such performance improvements are economically meaningful for consumers, artists, and the media streaming platform, ultimately improving the overall welfare, leading to significant managerial implications on the better design of a multi-stakeholder multi-objective recommender system.

2. Related Work

Our work draws on the following topics: recommendations in the two-sided media markets, multi-objective recommendations, multi-stakeholder recommendations, and transfer learning. We will now describe the relevant literature of each topic in detail.

2.1. Recommendations in the Two-Sided Media Markets

While recommender systems (RSes) are central to the media industry, existing methods from the CS community have predominantly focused on the demand side, optimizing user-centric metrics such as accuracy and diversity, and they have been successfully adopted in the major media streaming platforms, such as Spotify (Mehrotra and Carterette 2019), YouTube (Zhao et al. 2019b), and Alibaba (Li and Tuzhilin 2024b).

However, the media market is inherently two-sided, and achieving a balance between the utility of consumers (users) and suppliers (artists/creators) is an important task to improve the overall welfare of the platform, as demonstrated in a series of IS and Marketing research. Specifically, researchers in (Malgonde et al. 2020) highlight that successful matching on digital platforms requires managing the complexity of both supply and demand to prevent provider churn. Similarly, Shi (2023) demonstrates that optimal matchmaking strategies should account for the distinct incentives of both sides to maximize social welfare, rather than merely satisfying immediate user preferences. Cohen and Zhang (2022) further argues that ignoring the supply side in competitive environments can lead to market failure due to the cross-side network effect. Unfortunately, such an effect has not been systematically captured in existing media recommendation models, and, to the best of our knowledge, there is no specific deep learning architecture that can operationalize these economic principles of two-sided optimization in real-time media recommendations.

To address this research gap, we propose a novel "knowledge bridge" in this paper to transfer latent information between stakeholders, which enables us to leverage the cross-side network effects identified in the economic literature, resulting in optimal recommendation performance for the media platform.

2.2. Multi-Objective Recommendations

As an important practice in the industry (Burke and Abdollahpouri 2017, Eide and Zhou 2018, Sener and Koltun 2018, Sürer et al. 2018), multi-objective recommendations are provided using two groups of approaches in the literature. The first group aims at aggregating multiple objective values into one single objective value, and then providing recommendations by optimizing that aggregated value, using the techniques of evolutionary algorithms (Geng et al. 2015), bandits (Mehrotra et al. 2020), or ranking aggregation (Ribeiro et al. 2014). However, these approaches are only applicable to small-scale datasets (Milojkovic et al. 2019), since they require a significant amount of computational effort to determine the optimal configurations for objective aggregation, and as a result, are not scalable to reflect real-world use cases.

The second group, meanwhile, seeks to combine multiple prediction models, each built for one particular objective, into one single prediction model that reflects the balance between multiple objectives (Ruder 2017, Rodriguez et al. 2012, Ma et al. 2018, Sener and Koltun 2018, Lin et al. 2019). Representative methods include Tensor Factorization (Yang and Hospedales 2016), Cross-Stitch Network (Misra et al. 2016), and Relational Network (Zhao et al. 2019a). Researchers have also proposed to use neural networks to facilitate joint optimization of multiple objectives, which include Shared-Bottom (Ruder 2017), where each objective has its own tower after the shared-bottom module, and the loss for each task is computed separately; Multi-Gate Mixture-of-Expert (Ma et al. 2018), which jointly optimizes the modeling of shared information and task-specific information through a gating network and a mixture-of-expert network; and Mixture of Sequential experts (Qin et al. 2020), which explicitly learns sequential user behavior using the LSTM network in the multi-gate mixture-of-expert multi-task modeling framework. However, these methods only focus on the objectives of one single stakeholder (e.g., user), and typically perform well only when the targeted objectives are positively correlated with each other (Lin et al. 2019), while the objective relationships in practice can be highly complex or even conflicting (Milojkovic et al. 2019).

To address these research gaps, we propose a novel multi-objective multi-stakeholder recommendation framework for media recommendations, where we construct multiple neural network “towers” to model each objective individually, and then utilize a “bridge” architecture to bidirectionally transfer latent information between different stakeholders, leading to effective recommendation performance that balances between multiple objectives for different stakeholders.

2.3. Multi-Stakeholder Recommendations

A series of multi-stakeholder recommender systems have been developed to generate satisfying recommendation performance in the literature (Burke and Abdollahpouri 2017, Malgonde et al. 2020, Abdollahpouri et al. 2020, Jannach 2022). For example, researchers have proposed to combine the objectives of multiple stakeholders into a single one to provide recommendations (Nguyen et al. 2017), adopt a constraint learning approach by optimizing the welfare of a single stakeholder and setting all other stakeholders as constraints (Sürer et al. 2018), and perform Pareto optimization for multi-stakeholder objective learning (Zheng et al. 2019). Furthermore, the multi-stakeholder recommendation is also related to existing research on two-sided marketplaces (Shi 2023), where researchers designed multiple matching algorithms to improve consumer satisfaction (Dai and Jordan 2021) and long-term social welfare (Mladenov et al. 2020) simultaneously.

However, existing methods typically assume that each stakeholder is only associated with one single objective, which is explicitly defined based on interaction records. Meanwhile, two-sided media markets typically involve multiple stakeholders with multiple different objectives. In addition, the objectives are generated mostly from users’ implicit interactions with the media streaming platform, such as skipping a song. To bridge this research gap, we propose a novel multi-objective multi-stakeholder recommendation framework in this paper that achieves superior media recommendation performance.

2.4. Transfer Learning

Finally, our paper is related to transfer learning (Pan and Yang 2009), which assumes the existence of a common knowledge structure across multiple data sources for jointly modeling the data distributions. In addition, to exploit the duality between data distributions and to enhance the transfer capability, researchers have proposed the dual transfer learning mechanism (Long et al. 2012, Li and Tuzhilin 2020) in the application of cross-domain recommender systems. In this paper, we leverage the idea of transfer learning and focus on the flow of knowledge sharing (i.e., the bridge) between stakeholders to provide improved predictions for multi-objective multi-stakeholder media recommendations. More specifically, the predicted objectives are simultaneously optimized to balance the trade-off between multiple stakeholders’ objectives. In addition, we present a Shapley-value-oriented mechanism to identify the optimal location for the bridge to maximize its transfer learning effectiveness. We will now describe the details of our proposed framework.

3. MS-NSW: Our Proposed Utility Metric

In this section, we will introduce our proposed MS-NSW utility metric. We will first introduce the preliminaries of the Multi-Attribute Utility Theory (MAUT), which provides the theoretical foundation for a multi-criteria decision-making process. However, through a motivational example, we show that existing MAUT-based utility metrics, which focus either on the user welfare or on the artist welfare, do not perform well in the media recommendation task. Therefore, motivated by the Nash Social Welfare (NSW) theory, we propose the MS-NSW utility metric to balance the welfare of different stakeholders in recommendations. Through theoretical analysis, we demonstrate that MS-NSW is the unique continuous welfare function that satisfies four economic axioms for two-sided media platforms simultaneously, and that it is theoretically guaranteed to achieve Pareto Dominance over alternative metrics under certain conditions. Finally, we illustrate that optimizing MS-NSW is a non-trivial task, and it requires a fundamentally novel recommender system design, which we will present in the next section.

3.1. Preliminaries: Multi-Attribute Utility Theory

Multi-Attribute Utility Theory (MAUT) is a foundational construct in the multi-criteria decision-making literature, which formalizes how a stakeholder evaluates alternatives associated with multiple attributes or objectives (Keeney and Raiffa 1993, Wallenius et al. 2008). In our recommendation task, MAUT provides a disciplined way to combine multiple objectives *of each stakeholder* into a single utility metric for that stakeholder, respectively. For example, a user's experience with a recommendation is not determined by relevance alone, but jointly by listening intensity, engagement, and feedback. Similarly, an artist's value is shaped by multiple growth-related outcomes such as new-fan formation and recurring-fan retention. MAUT allows the media platform to aggregate these dimensions into a single scalar utility for either users or artists when providing recommendations, so that the welfare of that stakeholder can be maximized.

Formally, we denote multiple recommendation objectives associated with the stakeholder x as a vector $\mathbf{x} = (x_1, \ldots, x_d)$. We adopt the standard MAUT assumption of *additive independence* across objectives (Keeney and Raiffa 1993), which implies that the contribution of each objective to the overall utility is linear and independent from other objectives:

$$u(\mathbf{x}) = \sum_{k=1}^{d} u_k(x_k), \tag{1}$$

where $u_k(\cdot)$ denotes the marginal utility function of the k-th objective, and it is usually approximated using linear functions to obtain an additive value model that is both estimable and interpretable (Wallenius et al. 2008). To that end, we can formulate the MAUT utility metric as follows:

$$u(\mathbf{x}) = \sum_{k=1}^{d} \alpha_k \times x_k, \tag{2}$$

where α_k represents the relative importance of the k-th objective.

While MAUT provides a principled objective aggregation mechanism, it remains a *single-stakeholder* construct, as it specifies how one stakeholder aggregates multiple objectives into a scalar utility, without prescribing how a platform should aggregate utilities across different stakeholders. This limitation is particularly undesirable in two-sided media recommendation platforms, where user-side and artist-side objectives are coupled through cross-side network effects and platform intermediation. In fact, we will only be able to provide suboptimal recommendations if we focus only on the MAUT-user or MAUT-artist utility metric, as we demonstrate through the following motivational example.

3.2. A Motivational Example: The Necessity of A New Utility Metric

To motivate the need to propose a new utility metric for the multi-stakeholder, multi-objective recommendation task, we present a motivational example in this section. Specifically, we study a two-sided media market with $N = 1,000$ users and $M = 100$ artists, while the artists are further categorized into two groups to reflect the "long-tail" nature of media markets: *Superstar Artists* (A_{star}, 10% of artists) and *Niche Artists* (A_{niche}, 90% of artists). We sample the objective values based on the following explicit simulation rules:

User Objectives (O_{user}) Users evaluate items based on *Relevance* (o_{rel}) and *Discovery* (o_{disc}).

- **Relevance (o_{rel})**, which is modeled using truncated normal distributions to reflect the heterogeneous user preference towards different artists. For superstar artists, since they typically possess broad appeal (i.e., high mean and low variance), we formulate the objective as $o_{rel}(u, i \in A_{star}) \sim \mathcal{N}(0.95, 0.05)$. For niche artists, we model the consumers' polarized interests as a Gaussian Mixture Model conditioned on a latent indicator variable $I_{u,i}$ (whether user u is a "latent fan" of artist i or not): $o_{rel}(u, i \in A_{niche}) \sim I_{u,i} \times \mathcal{N}(0.85, 0.02) + (1 - I_{u,i}) \times \mathcal{N}(0.10, 0.05)$ where $I_{u,i} \sim \text{Bernoulli}(p = 0.05)$.
- **Discovery (o_{disc})**, which is modeled using uniform distributions: for superstar artists, we assume a low value due to ubiquity: $o_{disc} \sim \mathcal{U}(0.1, 0.3)$, while a high value for niche artists: $o_{disc} \sim \mathcal{U}(0.8, 1.0)$.

Artist Objectives (O_{artist}) Artists seek *Exposure Value* (o_{exp}) and *Fan Growth* (o_{fan}) simultaneously.

- **Exposure Value (o_{exp})**, which is defined based on the artist's need for visibility and formulated as the inverse of market saturation: $o_{exp}(u, i) = C \times \frac{1}{\sqrt{\text{Popularity}(i)}}$, where C is a normalization constant, and $\text{Popularity}(i)$ represents the number of appearances of i.
- **Fan Growth (o_{fan})**, which is formulated as a binary indicator derived from the user's latent fan status ($I_{u,i}$) to capture the incremental value when a niche artist is successfully matched to a latent fan: $o_{fan}(u, i) = \mathbb{I}(I_{u,i} = 1)$

We then simulate a single-slot recommendation decision for all users ($K = 1$), comparing three utility functions U as the optimization target:

1. **MAUT-User**: Maximizes $U = U_{user} = \alpha_{rel} \times o_{rel} + \alpha_{disc} \times o_{disc}$.
2. **MAUT-Artist**: Maximizes $U = U_{artist} = \alpha_{exp} \times o_{exp} + \alpha_{fan} \times o_{fan}$.

3. **MS-NSW (Proposed)**: Our proposed metric that maximizes the Nash Social Welfare of the user and artist objectives, which is defined as $U = \sqrt{U_{user} \cdot U_{artist}}$. It will be described in detail in Section 3.3.

We determine the optimal values of the objective weights α_{rel}, α_{disc}, α_{exp}, and α_{fan} through Bayesian Hyperparameter Optimization (similar to the one that we conduct in Appendix F), which returns $\alpha_{rel} = 0.9$, $\alpha_{disc} = 0.1$, $\alpha_{exp} = 0.5$, and $\alpha_{fan} = 0.5$. The averaged results over 10 independent runs are presented in Table 1, which shows that the MAUT-User strategy leads to a "Superstar Trap", as it maximizes immediate relevance at the cost of market failure for 90% of suppliers (niche artists) who receive zero exposure. In contrast, the MAUT-Artist strategy achieves equity through "Indiscriminate Exposure", forcing recommendations of niche content to non-fans to satisfy exposure constraints; while this achieves 100% coverage, it degrades user utility to near-random levels (0.214). Our MS-NSW metric successfully addresses the gap in both methods by identifying the specific latent fans in the long tail. It sacrifices a marginal amount of aggregate user utility (-3.1%) to achieve full artist coverage and a massive increase in valid fan matches compared to the artist-centric baseline. We will now describe the details of our MS-NSW utility metric.

Table 1 Comparison of Optimization Strategies (Averaged over 10 Trials)

Metric	MAUT-User	MAUT-Artist	MS-NSW (Proposed)
Avg. User Utility	**0.875**	0.214	0.848
Artist Coverage	10%	**100%**	**100%**
Fan Matches (Total)	0	50	**990**
Gini Coefficient (Exposure)	0.90 (High)	0.05 (Low)	0.28 (Balanced)

Note: "Fan Matches" indicates the number of recommendations where the user was a latent fan of the artist.

3.3. Multi-Stakeholder Nash Social Welfare (MS-NSW)

Our *Multi-Stakeholder Nash Social Welfare (MS-NSW)* utility metric is constructed based on the Multi-Attribute Utility Theory (MAUT) that we describe in Section 3.1. Note that extending MAUT from a single stakeholder to a multi-stakeholder setting is not a trivial generalization. Classical MAUT relies on additive independence across attributes, an assumption that is often reasonable within a stakeholder but generally violated across stakeholders in two-sided media markets, as the objectives of users and artists are coupled through cross-side network effects: the value generated for one stakeholder depends on the outcome of the other. These interactions can be non-linear and complex, making additive aggregation across stakeholders inappropriate. Moreover, once stakeholder utilities interact through discrete recommendation decisions and platform constraints, identifying suitable utility parameters becomes computationally challenging.

To address these challenges, we adopt the concept of *Nash Social Welfare (NSW)* from welfare economics (Kaneko and Nakamura 1979), which has been widely used to evaluate collective outcomes in multi-agent systems. NSW is designed as the geometric mean of stakeholders' utilities to balance efficiency with fairness when allocating resources across agents with potentially conflicting interests (Gong et al. 2023). Formally, for a set of stakeholders $\mathcal{S} = \{1, \ldots, S\}$ with utilities $\{u_s\}_{s \in \mathcal{S}}$, the Nash Social Welfare is given by

$$NSW = \left(\prod_{s \in \mathcal{S}} u_s \right)^{\frac{1}{|\mathcal{S}|}} \tag{3}$$

The geometric aggregation underlying NSW is particularly well-suited for the two-sided platforms: since utilities are combined multiplicatively rather than through a weighted sum, improvements on one side of the platform generate limited welfare gains when the other side's utility remains low. This property naturally captures cross-side network effects that characterize two-sided markets: a recommendation policy that maximizes user engagement but provides little value to suppliers, or vice versa, results in low overall welfare. At the same time, NSW rewards balanced allocations in which all stakeholders receive positive gains and preserve Pareto optimality, while sharply penalizing extreme disparities in outcomes.

Therefore, we propose the Multi-Stakeholder Nash Social Welfare (MS-NSW) utility metric, where we formulate the multi-stakeholder utility aggregation following the NSW theory. Formally speaking, consider two stakeholders denoted as A and B, each evaluating outcomes with their respective MAUT-based utilities $u_A(\mathbf{x}) = \sum_i \alpha_{A,i} \times x_i$ and $u_B(\mathbf{y}) = \sum_i \alpha_{B,i} \times y_i$. The MS-NSW metric is defined as follows:

$$\text{MS-NSW} = \sqrt{u_A(\mathbf{x}) \cdot u_B(\mathbf{y})}. \tag{4}$$

To unleash the full potential of the MS-NSW metric, we present a novel method to determine the objective weights α_k in Equation 2. Specifically, rather than relying on explicit utility labels or exogenously specified weights, we infer these parameters by treating recommendation rankings as revealed-preference signals and estimating the MAUT parameters with ordinal regression. This allows the platform's implicit preference over alternative recommendations to be recovered directly from ranking behavior. By grounding utility estimation in ranking outcomes, our approach links stakeholder-level utility formulation to preference in recommendations, while preserving the interpretability of MAUT and enabling balanced welfare optimization at the platform level. The details of ordinal regression will be introduced in Section 4.5.

Together, our MS-NSW metric provides a principled bridge between stakeholder preference modeling and platform welfare optimization. We will next demonstrate its theoretical benefits in a rigorous manner.

3.4. Theoretical Properties of MS-NSW: Optimality for Two-Sided Media Markets

In this section, we will demonstrate that MS-NSW is the *optimal* utility metric for two-sided media markets under four economic axioms, which any welfare metric for a two-sided media platform should satisfy. Specifically, we will prove that MS-NSW is the unique continuous welfare function satisfying all four axioms simultaneously, and that under the structural conditions of two-sided media markets, MS-NSW dominates all alternative welfare metrics in the class of weighted additive and min-based aggregations.

3.4.1. Four Economic Axioms for Two-Sided Media Platforms Let $\mathcal{W}$ denote the class of continuous welfare functions $W : \mathbb{R}_+^2 \to \mathbb{R}$ mapping the pair of stakeholder utilities (U_A, U_B) to a scalar welfare value used for providing recommendations. We propose the following four axioms, each motivated by the specific economic realities of two-sided media markets:

Axiom 1: Scale Invariance (SI). A welfare metric W is *scale Invariant* if for any constants $\lambda_A, \lambda_B > 0$:

$$\arg\max_{\mathbf{r}} W(\lambda_A U_A(\mathbf{r}),\ \lambda_B U_B(\mathbf{r})) = \arg\max_{\mathbf{r}} W(U_A(\mathbf{r}),\ U_B(\mathbf{r})) \tag{SI}$$

Motivation. In two-sided media markets, user objectives and artist objectives might be defined on different scales and with different natural units. Meanwhile, the optimal recommendation should not change simply because we measure listening time in seconds versus minutes. As we prove in Appendix G, MS-NSW satisfies SI while a weighted-sum metric $W = \lambda U_A + (1-\lambda) U_B$ does not.

Axiom 2: Pareto Optimality (PO). A welfare metric W is *Pareto Optimal* if its maximization always lies on the Pareto frontier of the achievable utility set $\mathcal{F}$, so that there exists no feasible recommendation that simultaneously improves U_A and U_B relative to the optimum.

$$\begin{aligned} \pi^* &= \arg\max_{\pi} W(U_A(\pi), U_B(\pi)) \\ \implies \nexists\, \pi' \text{ s.t. } U_A(\pi') \geq U_A(\pi^*),\ U_B(\pi') &\geq U_B(\pi^*),\ (U_A(\pi'), U_B(\pi')) \neq (U_A(\pi^*), U_B(\pi^*)) \end{aligned} \tag{PO}$$

Motivation. A welfare metric that leads to suboptimal recommendations, where both users and artists could be simultaneously better off, is undesirable for a platform whose goal is to improve overall ecosystem welfare. This directly addresses the failure modes of Table 1, where both MAUT-User and MAUT-Artist produce recommendations that are Pareto-dominated by our proposed MS-NSW metric.

Axiom 3: Independence of Irrelevant Alternatives (IIA). A welfare metric W satisfies *Independence of Irrelevant Alternatives* if the relative ranking of two recommendations r_1 and r_2 under W depends only on their own utility pairs $(U_A(r_1), U_B(r_1))$ and $(U_A(r_2), U_B(r_2))$, and not on the utilities of any other candidate recommendation r_3:

$$W(U_A(r_1), U_B(r_1)) > W(U_A(r_2), U_B(r_2)) \implies W(U_A(r_1), U_B(r_1)) > W(U_A(r_2), U_B(r_2)) \quad \forall\, r_3 \in \mathcal{R} \tag{IIA}$$

Motivation. The recommendation catalog at two-sided media markets, such as Spotify, contains hundreds of thousands of items (as shown in Appendix A). The platform's preference between two specific candidates should not be altered by the presence of a third, lower-quality option in the catalog.

Axiom 4: Cross-Side Complementarity (CSC). A welfare metric W satisfies *Cross-Side Complementarity* if the marginal welfare gain from improving one stakeholder's utility is increasing in the other stakeholder's utility level:

$$\frac{\partial^2 W}{\partial U_A\, \partial U_B} > 0 \tag{CSC}$$

and equivalently, W assigns zero welfare whenever either stakeholder receives zero utility:

$$W(U_A, 0) = W(0, U_B) = 0 \quad \forall\, U_A, U_B \geq 0 \tag{5}$$

Motivation. Cross-side complementarity formalizes the cross-side network effects in two-sided media markets (Anderson and Gabszewicz 2006, Cohen and Zhang 2022): an improvement in user welfare is more economically valuable when artist welfare is also high, because the two sides of the market are interdependent. Conversely, a recommendation that generates zero artist welfare should contribute zero overall welfare regardless of how high user engagement is. This directly captures the "Superstar Trap" failure mode of Table 1, where the MAUT-User metric ($\partial^2 W / \partial U_A \partial U_B = 0$ for an additive metric) cannot penalize the collapse of artist welfare to zero.

3.4.2. Axiomatic Uniqueness of MS-NSW

THEOREM 1 **(Axiomatic Uniqueness of MS-NSW).** *MS-NSW* $= \sqrt{U_A \cdot U_B}$ *is the unique continuous welfare function in* $\mathcal{W}$*, up to monotone transformation, satisfying axioms* (SI)*,* (PO)*,* (IIA)*, and* (CSC) *simultaneously. Moreover:*

(i) No additive welfare function $W = \lambda U_A + (1-\lambda) U_B$ *satisfies* (SI) *and* (CSC) *simultaneously.*

(ii) No min-based welfare function $W = \min(U_A, U_B)$ *satisfies* (SI) *and* (PO) *simultaneously.*

(iii) No max-based welfare function $W = \max(U_A, U_B)$ *satisfies* (PO) *and* (CSC) *simultaneously.*

The full proof of Theorem 1 is provided in Appendix G. As a result, we establish that MS-NSW is the unique axiomatically valid metric based on these four criteria for two-sided media markets.

3.4.3. Dominance of MS-NSW over Alternative Metrics

We now strengthen the results in Theorem 1 by proving that under certain structural conditions of two-sided media markets, namely *Cross-Side Network Effects*, *Long-Tail Content Distribution*, and *Discrete Recommendation Slot Constraints*, MS-NSW strictly *dominates* alternative metrics in terms of the quality of the recommendations it produces.

THEOREM 2 **(Dominance of MS-NSW).** *Let* $\mathcal{W}_{\text{alt}}$ *denote the class of welfare functions that violate at least one of the axioms* (SI)–(CSC)*, including all weighted additive metrics* $W^\lambda = \lambda U_A + (1-\lambda) U_B$ *for any fixed* $\lambda \in [0, 1]$ *and all min-based metrics. Under conditions* (C1)*,* (C2)*, and* (C3) *of Theorem 4 (which will be described in the next section), the following holds:*

(i) ***Welfare Dominance.*** *For any* $W^{\text{alt}} \in \mathcal{W}_{\text{alt}}$*, let* π^*_{NSW} *and* π^*_{alt} *denote the optimal policies under MS-NSW and* W^{alt} *respectively. Then:*

$$\text{MS-NSW}(\pi^*_{\text{NSW}}) \geq \text{MS-NSW}(\pi^*_{\text{alt}}) \tag{6}$$

with strict inequality whenever π^*_{alt} *fails to identify the latent fan matches in the item catalog, i.e., when it falls into either the "Superstar Trap" or "Indiscriminate Exposure" failure modes.*

(ii) ***Market Failure Prevention.*** *MS-NSW is a unique metric in $\mathcal{W}$ that simultaneously prevents both market failure modes identified in the two-sided media market:*

$$\textit{Superstar Trap:} \quad U_B(\pi^*_{\mathrm{NSW}}) > 0 \quad \textit{(niche artists receive positive welfare)} \tag{7}$$

$$\textit{Indiscriminate Exposure:} \quad U_A(\pi^*_{\mathrm{NSW}}) > \underline{U}_A \quad \textit{(user welfare exceeds a minimum threshold)} \tag{8}$$

where $\underline{U}_A > 0$ is the user utility achieved by a random recommendation policy.

(iii) ***Long-Run Ecosystem Growth.*** *Under iterative deployment over T recommendation rounds, the cumulative ecosystem welfare under MS-NSW grows at a rate weakly faster than any $W^{\mathrm{alt}} \in \mathcal{W}_{\mathrm{alt}}$:*

$$\sum_{t=1}^{T} \text{MS-NSW}(\pi^*_t) \geq \sum_{t=1}^{T} \text{MS-NSW}(\pi^{\mathrm{alt}}_t) + \Omega(T \cdot \Delta) \tag{9}$$

where $\Delta > 0$ is the per-round welfare gap established in part (i), and the growth rate strictly increases over time as the recommender system learns from the richer, more diverse feedback signal generated by MS-NSW-optimal recommendations.

The full proof of Theorem 2 is provided in Appendix H, while we would like to point out that Theorems 1 and 2 are directly supported by our empirical results as well. Part (i) of Theorem 2 corresponds to the MS-NSW column improvements in Table 4 (9.98% and 14.64% over the best additive baseline). Part (ii) corresponds to the simultaneous maintenance of user utility (0.848) and artist coverage (100%) in Table 1, which no additive metric achieves. Part (iii) corresponds to the longitudinal simulation results of Figure 7, where MS-Bridge (which optimizes MS-NSW) widens its performance gap over the baselines across all 15 simulation rounds, confirming the strictly positive long-run growth differential Δ in Equation (9).

3.5. Optimization of MS-NSW is Non-Trivial

We will now explain why the task of optimizing MS-NSW in recommendations is a non-trivial one. Specifically, we demonstrate that finding an exact optimal policy π^* is computationally intractable, and that trivial heuristic solutions fail to capture the entanglement between objectives in the two-sided markets.

3.5.1. NP-Hardness of MS-NSW Optimization

THEOREM 3 **(NP-Hardness).** *Maximizing the MS-NSW metric is an NP-Hard problem.*

The full proof of Theorem 3 is provided in Appendix I, which formally states the complexity of our specific optimization problem. Besides the NP-Hardness, recent literature has also established that even for a single stakeholder, there is no polynomial-time algorithm that can approximate the optimal solution for maximizing the Nash Social Welfare (Lee 2017), and that even in restricted settings with binary valuations, the problem remains intractable due to the combinatorial explosion of allocating indivisible items (Cole and Gkatzelis 2015). Therefore, our optimization task is a particularly challenging one.

3.5.2. Infeasibility of Trivial Solutions One might argue that since the MS-NSW objective function is smooth, standard heuristics or relaxation techniques could provide trivial solutions. However, as pointed out in (Ramezani and Endriss 2009), multi-agent resource allocation under NSW is typically complex and non-monotonic. We specifically demonstrate below that two "trivial" solutions, greedy decomposition and linear scalarization, will fail in our two-sided media context:

Failure of Greedy Decomposition. Classical recommender systems assume that recommendations are independent, where the optimal items for User A do not depend on the items chosen for User B. This allows the problem to be decomposed into N parallel sub-problems, one per user. However, in the MS-NSW objective, the marginal gain of recommending artist a to user u depends on the current accumulated utility of artist a from *all other users*: $\frac{\partial(\text{MS-NSW})}{\partial \pi_{u,a}} \propto \frac{1}{U_{artist}(a)}$ Because the gradient for any single user recommendation depends on the global state of the artist, the problem cannot be decomposed, and a greedy approach will only lead to suboptimal local minima (Barman et al. 2018).

Failure of Linear Scalarization. Another trivial solution for multi-objective optimization is to maximize a weighted sum: $\lambda_1 U_{user} + \lambda_2 U_{artist}$. However, MS-NSW is inherently non-linear with the format of the geometric mean, and the *discrete* nature of recommendation slots makes it a non-linear integer programming problem. Unlike linear scalarization, which finds solutions on the convex hull of the Pareto frontier, the NSW optimum often lies in the non-convex interior, balancing equity in a way that weighted sums cannot approximate without dynamic, adaptive weights.

To conclude, the optimization of MS-NSW is NP-Hard, and a trivial solution is not possible. We therefore propose a novel deep-learning-based method to optimize MS-NSW, which we will introduce next.

4. MS-Bridge: Our Proposed Recommendation Framework

4.1. Motivation of Our Design Components

Before describing the model specifics, we first provide the motivations of two crucial design choices: (1) the adoption of deep learning techniques, and (2) the proposed "Knowledge Bridge" mechanism.

4.1.1. Deep Learning is Effective for NP-Hard Problems. As demonstrated in Section 3.5, the optimization of MS-NSW is NP-Hard, and traditional optimization methods such as integer linear programming are not feasible for this task (Dantzig 2002). Meanwhile, recent literature has demonstrated the great potential of deep neural networks for solving NP-Hard problems (Bengio et al. 2021) through *Neural Combinatorial Optimization* (Bello et al. 2016). Unlike traditional optimization methods that are manually designed and often get trapped in local optima (as shown in our motivational example in Section 3.2), deep learning models act as universal function approximators (Hornik et al. 1989) to learn the complex, non-convex landscape of MS-NSW, and discover latent patterns to maximize its value (Bello et al. 2016, Khalil et al. 2017). Therefore, we adopt deep learning in our framework not only for its predictive power but also for its ability to serve as a scalable approximator for our intractable welfare maximization problem.

4.1.2. Knowledge Sharing is Crucial for Optimizing MS-NSW. The second motivation stems from the coupled nature of MS-NSW, where the objectives of both stakeholders are mutually dependent. However, simply merging all objectives into a single shared neural network may lead to the problem of *Negative Transfer* (Wang et al. 2019). As illustrated in Figure 4a, user and artist objectives are often negatively correlated (e.g., "User Satisfaction" vs. "Artist Exposure"). In standard Multi-Task Learning (MTL), forcing conflicting tasks to share the same parameters will result in suboptimal performance, as the optimization of one objective degrades the performance of another (Ruder 2017). To resolve the tension of negative transfer, we develop a *"Latent Knowledge Bridge"* architecture that allows for *selective* information sharing between both stakeholders, where we transfer latent knowledge that is mutually beneficial while isolating conflicting signals. The knowledge bridge acts as a regularization channel that allows the user tower to "borrow" statistical strength from the artist tower (and vice versa) only at specific layers where the latent features are compatible, thereby effectively optimizing MS-NSW without facing the negative transfer problem.

4.1.3. Theoretical Benefits of Knowledge Sharing Bridge We will now establish that the knowledge-sharing bridge is theoretically necessary to optimize MS-NSW under three structural conditions that characterize two-sided media markets. We first introduce these conditions and verify that two-sided media markets inherently satisfy all three.

We denote the utility for users and artists as $u_A(\mathbf{x})$ and $u_B(\mathbf{y})$ accordingly. To optimize the utility metric $MS-NSW = \sqrt{u_A(\mathbf{x}) \cdot u_B(\mathbf{y})}$, we formally define the necessity of the bridge structure as

$$\sup_{\pi \in \Pi_{\text{bridge}}} \mathbb{E}_u[MS-NSW(\pi(u))] > \sup_{\pi \in \Pi_{\text{dec}}} \mathbb{E}_u[MS-NSW(\pi(u))] \tag{10}$$

where u denotes the set of users on the platform, Π_{dec} denotes the class of recommendation policies achievable by any decoupled architecture, and Π_{bridge} denotes policies achievable with a bidirectional cross-stakeholder bridge. We now state three structural conditions on the utility distributions.

(C1) Conditional Utility Dependence. User and artist utilities are not conditionally independent given observable user and artist features:

$$\text{Cov}(u_A(\mathbf{x}), u_B(\mathbf{y})) \neq 0 \quad \text{for some } \mathbf{x} \in \mathcal{X} \text{ and } \mathbf{y} \in \mathcal{Y} \tag{C1}$$

(C2) Cross-Item Covariance Heterogeneity. The conditional covariance structure varies across items:

$$\text{Cov}(u_A(\mathbf{x_i}), u_B(\mathbf{y_i})) \neq \text{Cov}(u_A(\mathbf{x_j}), u_B(\mathbf{y_j})) \quad \text{for some items } i \neq j \tag{C2}$$

(C3) Pareto Non-Linearity. The achievable utility set $\mathcal{F} = \{(u_A(\pi), u_B(\pi)) : \pi \in \Pi\}$ is non-convex, so the MS-NSW optimum cannot be attained by any linear scalarization:

$$\exists\, \pi^* \in \arg\max_\pi \sqrt{u_A(\pi) \cdot u_B(\pi)} \quad \text{s.t.} \quad \pi^* \notin \arg\max_\pi \,[\lambda u_A(\pi) + (1-\lambda) u_B(\pi)] \;\forall\, \lambda \in [0,1] \tag{C3}$$

We can now state our main result on the necessity of the latent knowledge bridge.

THEOREM 4 **(Necessary and Sufficient Conditions for Bridge Necessity).** *The bridge is necessary in the sense of Equation* (10) *if and only if conditions* (C1)*,* (C2)*, and* (C3) *hold simultaneously.*

The full proof of Theorem 4 is provided in Appendix J. Meanwhile, we would like to point out that all three conditions are *structural properties* of two-sided media markets, directly verifiable from the empirical settings of our offline evaluations in Section 5, as we will explain below.

C1 is satisfied. In all four datasets, user and artist utilities are driven by the same underlying user-artist interaction event. Take the Spotify dataset as an example: a single listening event simultaneously determines both the user objective *Percentage of Listening* (how much of the song the user consumed) and the artist objective *New Fan* (whether the user is a first-time listener of the artist. These two quantities are not conditionally independent, as a user who listens to a large fraction of a song is substantially more likely to register as a new fan of that artist than a user who skips it.

C2 is satisfied. The conditional covariance structure between u_A and u_B varies systematically across items in all four datasets, as documented by the objective statistics in Tables 10–13 and the correlation structures of Figure 4. In the Spotify dataset, for example, some songs generate new fans at high rates (for users who have never heard the artist) while others generate none (for already-familiar users).

C3 is satisfied. The non-convexity of the achievable utility set $\mathcal{F}$ is directly evidenced by the Pareto frontier analysis of Figure 11, where the Pareto frontier of MS-Bridge strictly dominates the frontiers of all linear scalarization baselines at every value of α, including at the endpoints $\alpha \to 0$ (artist-only) and $\alpha \to 1$ (user-only). This means the MS-NSW optimum is not attainable by any fixed linear combination of u_A and u_B, confirming that it lies in the non-convex interior of $\mathcal{F}$ as required by (C3).

We therefore conclude that the Latent Knowledge Bridge is a structural necessity for any recommendation architecture deployed in a two-sided media market, which we will describe in detail next.

4.2. Overview of our Recommendation Framework

We now present the overview of our MS-Bridge framework. As shown in Figure 1, based on the inputs of stakeholder information at the bottom layer, our framework predicts the values of each objective through deep learning techniques, and transfers the latent information between stakeholders via a bridge architecture to provide multi-objective multi-stakeholder recommendations. It consists of the following three components that work contingently together: tower network, knowledge transfer bridge, and objective aggregation, and we will now describe the details of them.

4.3. "Tower" Network for Predicting Objective Values

The first element of our framework, **tower network**, constructs the latent representations of the information of every stakeholder by encompassing stakeholder attributes and their historical interactions with the system. As a result, the latent information is represented as an embedding, which is a compact vector representation that captures the essential semantics of the inputs by placing semantically similar inputs close

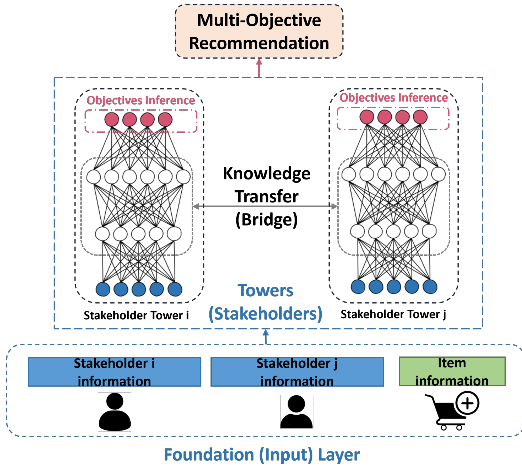


Figure 1 **Deep multi-stakeholder knowledge transfer recommendation framework.**

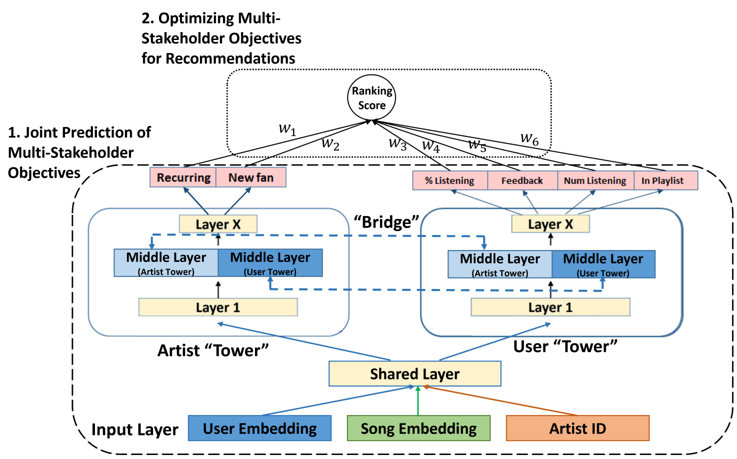


Figure 2 **Deep multi-stakeholder joint recommendation model (illustrated using the example of Spotify).**

together in the latent embedding space (Jin et al. 2017, Zhang et al. 2019). We then build multiple towers, where each leverages a distinct neural network to learn the objectives of its corresponding stakeholder using the provided stakeholder information. The stakeholders' towers are especially efficient for handling a series of objectives that might conflict with each other.

We will use Spotify as an example to explain how the ”Tower Network” works, which is illustrated in Figure 2. We construct two sets of neural networks in a two-tower architecture to account for the two stakeholders at Spotify: the user tower h_{user}, which jointly predicts the user objectives $o_{user} = \{\%Listening, NumTimesListening, FeedbackType, InPlaylist\}$; and the artist tower h_{artist}, which jointly predicts the artist objectives $o_{artist} = \{NewFan, RecurringFan\}$. The details of these objectives will be introduced in Section 5.2.

Next, we construct user embeddings w_{user}, song embeddings w_{song}, and artist embeddings w_{artist} to represent the relevant latent information. For the Spotify dataset in our offline evaluations, these embeddings are generated and provided directly by Spotify (Anderson et al. 2020, Hansen et al. 2021), which constitute

40-dimensional latent vectors and are re-trained every day to account for new content added to the Spotify platform and for the dynamic shift of user preferences. For consistency purposes, we use the set of embeddings provided on the last day of the dataset in our evaluations, while also conducting additional evaluations in Appendix E.2 using the set of embeddings provided on the first day, or on the median day. For the MLHD dataset, however, we use the one-hot encoding of user IDs, media IDs, and artist IDs as their initialized latent embeddings $w_{user}, w_{media}, w_{artist}$ respectively, since we do not have additional feature-level information in those datasets, and will later update these embeddings during the recommendation process through back-propagation in an end-to-end manner. Finally, for the Alibaba-Youku dataset, the embeddings are learned based on the consumer/video/uploader feature information, and they are first randomly initialized based on a uniform distribution, and then learned following the AutoEncoding technique.

Then, to model the interactions between users, media content, and artists, we use the state-of-the-art method of Neural Collaborative Filtering (NCF) (He et al. 2017), which concatenates the user, media, and artist embeddings as a combined vector $w = [w_{user}, w_{media}, w_{artist}]$, and maps this vector in the latent space through shared-bottom network (Ruder 2017) to produce objective value predictions. The latter is commonly used in multi-objective recommendation tasks to improve generalization, speed up training, and reduce memory requirements (Covington et al. 2016, Ma et al. 2018). It is formulated as a fully-connected neural network $f(\cdot)$. The latent element $\hat{w} = f(w)$ that we obtain in the shared-bottom network captures the essence of user-media-artist preference information and serves as the input for subsequent prediction tasks across all stakeholders' objectives o_{user}, o_{artist}. To demonstrate the robustness and flexibility of our proposed framework, we also implemented other embedding approaches, such as Wide & Deep (Cheng et al. 2016), for producing the latent element $\hat{w}$, where we still achieve significant performance improvements.

Above the shared-bottom network, we construct a series of "tower" networks to predict the values of each objective. Unlike existing methods that construct a separate tower network for each objective individually (Ma et al. 2018), we construct a separate tower network for each stakeholder, and jointly predict the associated objectives for each stakeholder within each tower. The motivation comes from the natural correlations and properties between each of the stakeholders' objectives. As shown in (Milojkovic et al. 2019), simultaneously optimizing a multitude of objectives, correlated or not, with potentially different scales, has proven to be difficult. Meanwhile, as shown in the correlation analysis presented in Figure 4a, the artist objectives are significantly positively correlated with each other and negatively correlated with user objectives. Hence, to overcome the negative transfer problem (Wang et al. 2019, Zhang et al. 2022) and utilize the positive correlations between the objectives from the same stakeholder, we propose to build separate towers h_{user} and h_{artist} for each stakeholder. Each tower network consists of three fully-connected neural network layers, denoted by f^{top},f^{middle}, and f^{bottom}. The latent element $\hat{w}$ will be fed into these three layers sequentially to obtain the predictions of corresponding objectives. For example, user objectives o_{user} will be predicted as $o_{\hat{user}} = f^{top}_{user}(f^{middle}_{user}(f^{bottom}_{user}(\hat{w})))$, while the artist objectives o_{artist} will be predicted as $o_{\hat{artist}} = f^{top}_{artist}(f^{middle}_{artist}(f^{bottom}_{artist}(\hat{w})))$.

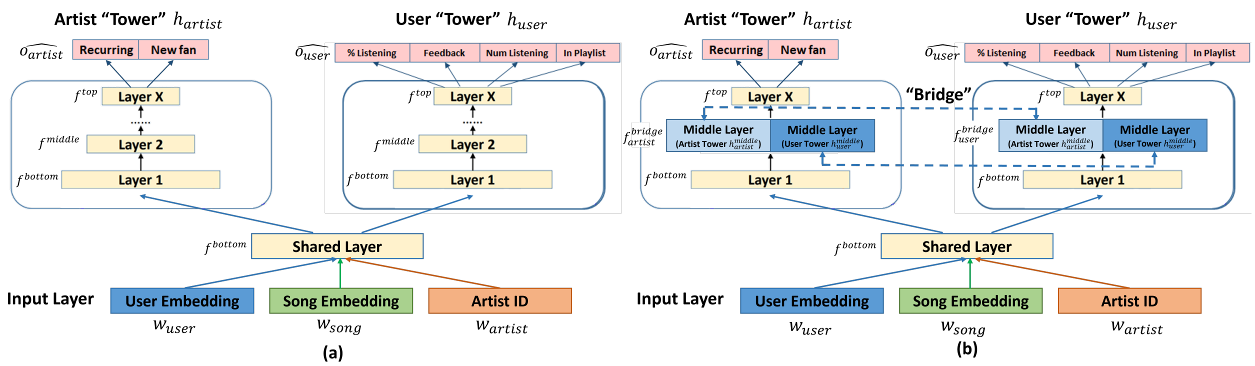


Figure 3 Illustrations of: (a) Joint Prediction Model (No Bridge) and (b) Joint Prediction Model (+Bridge).

4.4. Knowledge Bridge for Joint Multi-Stakeholder Prediction

The second and the most important element in our framework is **Latent Knowledge Bridge**, which facilitates the sharing of underlying latent information learned within each tower network of the stakeholder. As we discussed before, the relationships between different stakeholders are complex but nevertheless useful for the recommendation process. To that end, our knowledge bridge bidirectionally transfers the knowledge between each stakeholder tower and provides a flow of "information" between these towers, as shown in Figure 3(b). The bridge is used to calibrate the learning and optimization processes in the hidden layers and will ultimately generate more comprehensive latent embeddings for recommendation purposes.

We build the bridge by concatenating a certain hidden layer between both towers to jointly estimate the objectives across the two stakeholders. The concatenation technique is particularly effective in the information-sharing task, as shown in the literature (Ruder 2017). Specifically, for the user tower h_{user}, we utilize the transferred information from the artist tower network by concatenating the values of the middle layer of tower h_{user} (denoted as h^{middle}_{user}) with the values of the middle layer of tower h_{artist} (denoted as h^{middle}_{artist} in Figure 3(b)). These values in the middle layer would then be learned and updated via an additional multi-layer perceptron (MLP) network: $h^{\hat{middle}}_{user} = f^{bridge}_{user}(h^{middle}_{user}; h^{middle}_{artist})$ that will be fed into the top layer f^{top} to predict the values of corresponding user objectives and minimize the prediction error. The artist tower network h_{artist} is also constructed and optimized in the same way.

Our design choice of building the bridge at *the middle layer*, rather than other layers of the neural network, is supported by both the nature of the transfer learning task, as well as a Shapley value-based approach that we present in this paper. From the intuition perspective, building it at the bottom is unproductive because this results in sharing "raw" data pertaining to users and artists that has not been properly processed and distilled by a neural network, which would lead to suboptimal results, as was shown in the literature (Ruder 2017). Similarly, building the bridge at the top layer is also not ideal because all the user and artist features are already computed/distilled, and there is not much information that can be properly shared at that point (i.e., it is "too late" to share). Therefore, building the bridge in the middle of the two towers is the best approach because information about the users and artists is already sufficiently distilled, but not "crystallized" yet.

Hence, it can be shared between the two ”towers” in a flow/exchange of valuable information, and can be properly utilized to provide better recommendations for both stakeholders.

Besides these intuitions, we have also conducted a comprehensive study of the optimal location of the bridge, based on a Shapley value-based approach that we present in this paper. Shapley value is the average of all marginal contributions to all possible coalitions in a dataset (Lundberg and Lee 2017), and it characterizes the importance of each component in a learning task. Therefore, by computing the Shapley value of each neural network layer, we are able to understand the relative importance of each layer in the recommendation process. We will then select the layer with the highest Shapley value and transfer its latent information in the most effective manner, since it has the greatest impact on the neural network's outputs (Ancona et al. 2019, Parvez and Chang 2021). Specifically, we compute the Shapley values for each neuron in every hidden layer and aggregate them to determine the average Shapley value for each layer. As we visualize the Shapley values of hidden layers in Figure 6, they are relatively small in the bottom and top layers, while achieving the largest value in the middle layer. This finding matches the intuitions that we previously discussed. Our analyses in Section 5.10 also confirm that the performance is positively associated with the Shapley value of the layer where the bridge is built. Meanwhile, positioning the bridge in the middle layer consistently produces the best performance for all objectives.

4.5. Aggregating Multi-Stakeholder Objectives for Recommendations

The third and final element of our framework involves **objective aggregation** of all the predicted values from multiple stakeholders to produce the final media recommendations (Adomavicius et al. 2011). Its goal is to balance various stakeholders' objectives to improve the overall welfare. In essence, this aggregation process consolidates the predicted objectives of each stakeholder in an end-to-end manner, and it is designed to naturally discern the balance among differing objectives, effectively determining their weights and subsequently generating appropriate recommendations.

Specifically, following the introduction of the MS-NSW metric in the previous section, the utility function for each media i, artist a, and user u is formulated as a weighted geometric mean of the predicted objectives $\hat{o_{artist}} = \{\hat{o_1}, \hat{o_2}\}$, $\hat{o_{user}} = \{\hat{o_3}, \hat{o_4}, \hat{o_5}, \hat{o_6}\}$:

$$Utility_{u,a,i}(W) = \sqrt{MAUT_u \times MAUT_a} = \sqrt{(w_1\hat{o_1} + w_2\hat{o_2}) \times (w_3\hat{o_3} + w_4\hat{o_4} + w_5\hat{o_5} + w_6\hat{o_6})} \qquad (11)$$

The set of objective weights $W = \{w_1, w_2, ..., w_6\}$ is generated from the ordinal regression method based on the ranking of recommendations, as we discussed in Section 3.3 when we formulated the MS-NSW metric. This is motivated by the *Revealed Preference Theory* (Samuelson 2024), where we assume that the ”Ground-Truth Ranking“ observed from actual user interactions reveals the optimal trade-offs between conflicting objectives. The ordinal regression method works as follows.

First, we obtain a list of *Ground-Truth Ranking* (Y_u) for each historical session in our training data. Formally, for a given user u, we assign an ordinal rank $r_{u,i} \in \{1, \ldots, R\}$ to each item i, where $r_{u,i} = 1$

represents the most preferred item and $r_{u,i} = R$ represents the least preferred, where R is the length of that session. This observed ranking serves as the target variable for our optimization.

We then model the relationship between the multi-objective utility and the ground-truth ranking using the *Proportional Odds Model* (McCullagh 1980), which posits that the observed rank $r_{u,i}$ is determined by whether the MS-NSW metric $Utility_{u,a,i}(W)$ exceeds certain learnable thresholds $\theta_1 < \theta_2 < \cdots < \theta_{R-1}$. The probability that item i is ranked at position j or better is given by the cumulative logistic distribution:

$$P(r_{u,i} \leq j|W) = \sigma(Utility_{u,a,i}(W) - \theta_j) \tag{12}$$

where $\sigma(\cdot)$ is the sigmoid function. We estimate the optimal weights W^* by minimizing the negative log-likelihood of the observed ground-truth rankings across all training sessions $\mathcal{D}$:

$$W^* = \operatorname*{argmin}_{W,\Theta} \sum_{(u,i)\in\mathcal{D}} \sum_{j=1}^{R-1} -\log\left[P(r_{u,i} \leq j|W)^{y_{u,i,j}} \cdot (1 - P(r_{u,i} \leq j|W))^{1-y_{u,i,j}}\right] \tag{13}$$

where $y_{u,i,j}$ is a binary indicator that equals 1 if the true rank $r_{u,i} \leq j$.

By optimizing this objective, our framework assigns higher weights w_k to the specific combination of objectives (e.g., ”Listening %” and ”Recurring Fan”) that best reproduces the high-quality rankings actually observed in the system. This approach aligns the theoretical MS-NSW metric with the empirical reality of user behavior, effectively bridging the gap between economic welfare optimization and recommender system design (Hu and Li 2018). In addition, this approach yields a globally concave loss function, which guarantees that the selected objective weights converge to a global optimum regardless of initialization.

Besides the ordinal regression method, there are two additional methods that we can select, including the neural network approach and the Bayesian optimization approach, which we will discuss in detail in the Appendix F. As we will demonstrate through extensive offline evaluations in the next section, both the neural network-based approach and the Bayesian Hyperparameter Optimization method will be able to generate effective objective weights that lead to superior recommendation performance.

4.6. Theoretical Properties of MS-Bridge

THEOREM 5 **(Convergence and Stability of MS-Bridge).** *Let* $\mathcal{U}(\Theta) = \ln\left(\sqrt{u_{user}(\Theta) \cdot u_{artist}(\Theta)}\right)$ *be the log-transformed MS-NSW objective function. Assume that* $\mathcal{U}$ *is* L*-smooth and that the stochastic gradients* $g_t = \nabla\mathcal{U}(\Theta_t;\xi_t)$ *have a bounded variance* σ^2_{bridge} *such that* $\mathbb{E}[\|g_t - \nabla\mathcal{U}(\Theta_t)\|^2] \leq \sigma^2_{bridge}$*. For a constant learning rate* $\eta \leq \frac{1}{L}$*, the expected optimality gap of MS-Bridge parameters* Θ *after* T *iterations of Stochastic Gradient Descent (SGD) is bounded by:*

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=0}^{T-1}\|\nabla\mathcal{U}(\Theta_t)\|^2\right] \leq \frac{2(\mathcal{U}(\Theta^*) - \mathcal{U}(\Theta_0))}{\eta T} + \frac{\eta L \sigma^2_{bridge}}{2} \tag{14}$$

where Θ^* *denotes the optimal parameters maximizing the Nash Social Welfare. Additionally, the longitudinal system of MS-Bridge is asymptotically stable.*

The full proof of Theorem 5 is provided in Appendix K, which guarantees the convergence and the long-term stability performance of MS-Bridge. We will next provide the extensive offline evaluations to demonstrate the benefits of our proposed method across various empirical settings.

5. Offline Evaluations

5.1. Datasets

To demonstrate the effectiveness of our proposed method, we leverage the following four different datasets obtained from the media industry. Their detailed statistics will be reported in Appendix A.

1. **Spotify (Next Song Recommendation)**, which is collected during February 14-20, 2020, from Spotify, where each song is recommended in a sequential manner. The dataset consists of 970,013 listening records over 294,469 distinct users who cumulatively listened to 81,948 songs. These users are a random sample from the entire Spotify user pool, which we obtained following Spotify's standard sampling protocols. We filtered out users and songs with less than five interactions, which is consistent with the common practices in recommender systems (Zhang et al. 2019) to ensure sufficient sparsity levels of the dataset and significance levels of our analysis.
2. **Spotify (Session-Based Recommendation)**, which is also collected from Spotify; however, each song is recommended within a playlist of a session. Each session is annotated with detailed information, including the set of tracks played or skipped during the session, and the session timestamp. The dataset consists of 8,892,650 listening records over 82,696 distinct users who listened to 417,932 songs during November 2020. We similarly filtered out users and songs with fewer than five interactions.
3. **Music Listening Histories Dataset (MLHD)**, which is collected from the music streaming platform Last.fm[1]. Since the dataset contains only positive feedback regarding user listening history, we collected information about the song's characteristics (such as track duration) to infer whether the user skipped or finished listening to a song. The dataset consists of 3,393,744 listening records over 1,052 distinct users who cumulatively listened to 407,681 songs, ranging from January 1, 2012, to October 1, 2013. We similarly filtered out users and songs with fewer than five interactions.
4. **Alibaba-Youku**, which is collected from Alibaba-Youku, the major video streaming platform in Asia (Li et al. 2020, Li and Tuzhilin 2024b). The dataset includes both the feature-level information of users and videos, and the session-related information. In total, the dataset includes 99,999 watching records from 51,419 users on 229 unique videos during the week of 07/29/2019 to 08/04/2019. We did not perform the filtering process since this dataset is relatively small.

We apply the five-fold time-based split strategy for all four datasets. Specifically, for the Spotify (Next Song Recommendation) dataset, we use the first five days of data for training ($\approx 72\%$), one day for validation, and the last day for testing ($\approx 14\%$). For the Spotify (Session-Based Recommendation) dataset,

[1] `https://ddmal.ca/research/The_Music_Listening_Histories_Dataset_(MLHD)/`

we use the first three weeks of data for training ($\approx 70\%$), one day for validation, and the last week for testing ($\approx 23\%$). Finally, for both the MLHD and the Alibaba-Youku datasets, we use the first 80% of data for training and validation, and the last 20% for testing. We also implemented the standard five-fold cross-validation strategy, where results in Appendix E.6 show that our method still achieves significant improvements. We further conducted additional evaluations on the long-tail artists and users, where results in Appendix E.1 show that our model still obtains significant improvements, demonstrating its benefits and flexibility regardless of the sparsity level of the data.

Next, we will introduce the details of the objectives that we study in these datasets.

5.2. Multi-Stakeholder Objectives

5.2.1. Spotify. For both the *Spotify (Next Song Recommendation)* and the *Spotify (Session-Based Recommendation)* datasets, we strictly follow the industrial practices at Spotify (Mehrotra et al. 2019) to select a set of objectives that align best with user satisfaction as well as the desire of artists. In addition, the exact formulation of these objectives also follows the internal engineering procedure set up by the then-Director of Research of the company to ensure that they capture meaningful preference information from users and artists, as well as produce effective business performance for Spotify. We further illustrate the business relevance and practical implications of our selected objectives through economic value analysis and a simulation study, which we conduct later in this paper.

Specifically, we consider the following user objectives that are closely related to user satisfaction:

- *Percentage of Listening*, which stands for the percentage of the recommended song that has been listened to by the user. It is a numeric value between 0 and 1.
- *Number of Playing Times*, which records the number of times that the recommended song has been played by the user before. If the user has not played this song, this objective will be 0. We normalize this objective to between 0 and 1 for the benefit of computation efficiency.
- *Type of Feedback*, which includes the labels of like (1), dislike (-1), and neutral/no interaction (0) that are explicitly expressed by the user for each recommended song. We treat this objective as a numerical value (rating), instead of a categorical value, to align with other objectives.
- *In Playlist*, which is a binary variable indicating whether the song was saved by the listener in her/his own playlists. It thus records the interactions of the users with the playlists.

We also consider the following two artist objectives that are essential to the welfare of artists:

- $NewFan(u,a,t)$, which measures the ability to acquire new fans dynamically. This is important for the artists since they need to expose their songs to new users to grow their fan bases. This objective is computed following the standard Spotify practice (Mehrotra et al. 2019) as follows:

$$NewFan(u,a,t) = \begin{cases} 1, & \text{if } NumTimesListenedBefore(u,a,t) = 0, \\ \exp\left[-1 \times \frac{NumTimesListenedBefore(u,a,t)}{\lceil NumDaysFromFirstListening(u,a,t) \rceil}\right], & \text{otherwise.} \end{cases} \tag{15}$$

where $NumTimesListenedBefore(u,a,t)$ is the number of times that user u listened to songs produced by artist a before time t, and $NumDaysFromFirstListening(u,a,t)$ is the number of days (rounded up) elapsed until time t since user u first listened to any song produced by artist a. For users who never listened to the artist before, this objective $NewFan(u,a,t)$ is simply set to 1.

- $RecurringFan(u,a,t)$, which measures the ability to acquire pre-existing listeners who have not recently listened to songs from this artist. This objective is important since fans are often lost between releases, and artists consistently seek to re-engage users who have not recently listened to them. Conversely, users who have recently listened to the artist are assigned a low score, based on the assumption that such users are more likely to listen to this artist again, thus limiting the necessity of a recommendation that would drive these users to this artist. This objective is computed following the standard Spotify practice (Mehrotra et al. 2019):

$$RecurringFan(u,a,t) = \begin{cases} 1, & \text{if } LastTimeListened(u,a,t) = 0, \\ \exp\left[-1 \times \frac{NumTimesListenedBefore(u,a,t)}{\lceil LastTimeListened(u,a,t) \rceil}\right], & \text{otherwise.} \end{cases} \tag{16}$$

where $LastTimeListened(u,a,t)$ is the number of days (rounded up) elapsed until time t since user u last listened to the songs from artist a. $LastTimeListened(u,a,t) = 1$ if the last time of listening was on the same day. $NumTimesListenedBefore(u,a,t)$ is the same as in Equation (15).

These two artist objectives are among the key metrics for which the artist marketing teams routinely spend most of their marketing budget, and therefore, serve as representative objectives in our offline evaluations. We have also included two other artists' objectives in our additional experiments reported in Appendix E.5 to further demonstrate the robustness of our framework. Furthermore, while these two objectives can be calculated in real-time, assuming we have the full user listening history, this is not the case for Spotify since it violates Spotify's privacy policy[2], and also consumes massive amounts of storage. Therefore, it is crucial to predict the values of these two objectives (Fazelnia et al. 2024), which we will further elaborate on.

5.2.2. Last.fm. For the music streaming platform Last.fm, since it is somewhat different from Spotify, we select a slightly different set of objectives for MLHD to accommodate minor idiosyncrasies present in this setting. In particular, the user objectives include the binary variables "*IsFinish*" and "*IsSkip*" measuring user feedback (i.e., whether the user finished listening to or skipped the recommended song), along with the *Percentage of Listening* objective that we use in the Spotify dataset. The artist objectives ("NewFan" and "Recurring Fan") are exactly the same as the Spotify dataset.

5.2.3. Alibaba-Youku. In the video streaming platform at Alibaba-Youku, the company considers the following three user objectives that reflect user satisfaction toward the recommended videos, and are closely related to the business revenues (Li et al. 2020, Li and Tuzhilin 2024b):

- *Video View*, a binary variable indicating whether the user has clicked on the video or not.

[2] `https://www.spotify.com/us/legal/privacy-policy/`

- *Time Spent*, a continuous variable measuring the amount of time that the user has spent on the video. It will be zero if the user does not click on the video.
- *Play Rate*, a continuous variable measuring the percentage of the video that has been played by the user. It will be zero if the user does not click on the video.

In addition, the company also considers the following two artist (i.e., content creator) objectives, ensuring their content reaches the right audiences while maintaining engagement:

- *Relevance*, which indicates the relevance level of the video to the user. It is computed as the inverse of the average Euclidean distance between the latent embeddings of the recommended video, and all videos that the user has watched in the previous session.
- *Novelty*, which indicates the novelty level of the video to the user. It is computed as the Euclidean distance between the embeddings of the recommended video and the last video watched.

The exact formulations of these two objectives strictly follow the industrial practices at Alibaba. Specifically, as explained by the manager of the video recommendation team at Alibaba, relevance and novelty are important objectives for the artists since ”*they provide valuable guidance for the content creators to adjust the topics of their creations, so that they will be able to hold on to their loyal customers (by optimizing the relevance objective), and also attract new customers that are curious about the novel content created by the artists (by optimizing the novelty objective)*.”

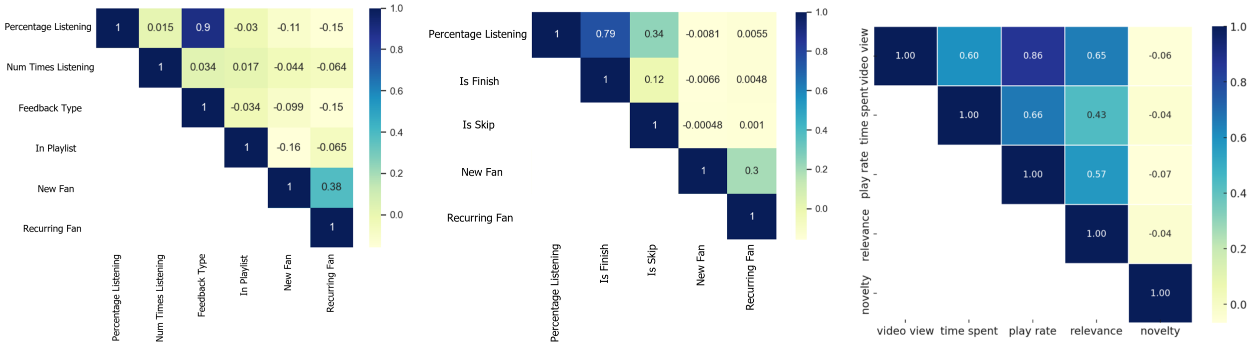
5.2.4. Relationships Between Objectives Note that optimizing these objectives across multiple stakeholders simultaneously is not a trivial task, since many of them conflict with each other. For example, as shown in Figure 4a in the Spotify dataset, the artist objective “Recurring Fan” is positively related to the other artist objective “New Fan,” while it is negatively related to all four listener objectives. We can also observe in Figure 4b and Figure 4c that the objective relationships in the MLHD and the Alibaba datasets are also complicated. This would raise the negative transfer problem (Wang et al. 2019, Zhang et al. 2022), where improving the performance of one objective may hurt the performance of other objectives.

This negative transfer problem is addressed in our proposed framework with the ”knowledge bridge“ architecture to effectively balance between different objectives, which we will show in the experiment results next. In fact, since the distributions of objective values across these four datasets are significantly different from each other, the significant performance improvements that we observe in the offline evaluations further demonstrate the strength and robustness of our proposed model design.

5.3. Baseline Models

We compare our model to state-of-the-art baselines from the recent computer science and business literature:

1. *Classical Recommendation Models*, which include linear regression (**LR**) (Koren 2010), neural matrix factorization (**NeuMF**) (He et al. 2017), and **Wide&Deep** (Cheng et al. 2016). These three models predict one objective at a time, and we implemented multiple models, one for each objective.



(a) (b) (c)

Figure 4 Heatmap of objective correlations in the (a) Spotify, (b) MLHD, and (c) Alibaba datasets.

2. *Multiple-Objective Models*, which include Multi-Criteria Collaborative Filtering **(MCCF)** (Adomavicius et al. 2011), Multi-Gate Mixture of Experts **(MMoE)** (Ma et al. 2018), Multitask Mixture of Sequential Experts **(MoSE)** (Qin et al. 2020), Neural Multi-Task Recommendation **(NMTR)** (Gao et al. 2019), Progressive Layered Extraction **(PLE)** (Tang et al. 2020), Multi-Objective Linear Upper Confidence Bound **(MO-LinCB)** (Mehrotra et al. 2020), and Multi-Objective Reinforcement Learning **(MORL)** (Abels et al. 2019). These seven models consider all objectives for prediction in a single aggregated model.

To ensure a fair comparison with these baseline models, we have devoted the same amount of effort in tuning the best hyperparameters and model training for these models and our MS-Bridge model. In addition, we also set the same optimization target (the MS-NSW utility metric) for all these models as well. Details about hardware conditions and hyperparameter selections are reported in Appendix B.

5.4. Evaluation Tasks and Evaluation Metrics

We implement three evaluation tasks to demonstrate the superiority of MS-Bridge along various dimensions.

Task 1: Joint Prediction of Recommendation Objectives. In this task, we evaluate the prediction accuracy of each objective with the following four popular metrics (Gunawardana et al. 2022): RMSE, MAE, AUC, and F1 score. Since the distributions vary across different objectives, we report the prediction accuracy for each objective separately. We use the Spotify (Next Song Recommendation), MLHD, and Alibaba-Youku datasets for this task.

Task 2: Next Song Recommendation. In this task, we compute Precision@K and Recall@K (Schedl 2019), which measure the quality of top-K song recommendations. Since the average number of users' daily song listening or video watching is relatively low, we select $K = 1, 3, 5$ for evaluation, and we use the Spotify (Next Song Recommendation), MLHD, and Alibaba-Youku datasets for this task.

Task 3: Session-Based Recommendation. In this task, we report the Mean Average Precision (MAP) and the Precision@3,5,10 and Recall@3,5,10 of recommendations in each session (among 50 candidates). A *session* is defined as the period of listening or video watching by a user with no more than five minutes of continuous inactivity (Hansen et al. 2020). We report the average values across all sessions in the test set. We use the Spotify (Session-Based Recommendation) and Alibaba-Youku datasets for this task.

5.5. Experiment Results on Task 1 (Objective Value Prediction)

We first present the results on the objective value prediction task, where we implement MS-Bridge with two backbones for representing latent embeddings: NCF (He et al. 2017) and Wide&Deep (Cheng et al. 2016). We report the results of RMSE and MAE in Table 2, while the results of AUC and F1 score are reported in Appendix C. We observe from these results that our MS-Bridge model significantly outperforms all other baselines across all objectives. Specifically, for the Spotify dataset, we improve the performance by reducing the RMSE by 0.26% to 11.37%, and MAE by 0.84% to 5.67%, relative to the best baseline. Specifically, our model yields the best performance improvement for the Feedback Type objective (11.37% in RMSE and 5.67% in MAE), which can be explained by the fact that knowing past interactions between users and artists can help better estimate whether the user will like or save the artist's song in the playlist. In addition, we can notice that the task of predicting artists' objectives is less challenging than predicting users' objectives, as the prediction accuracies for the artists' objectives in terms of RMSE and MAE are much lower relative to users' objectives. This can be explained by the fact that users' song preferences are often more sensitive and can be affected by many contextual factors, such as users' mood and current activity. We thus conclude that the bridge architecture *effectively learns the joint interactions between stakeholders and improves the multi-objective prediction accuracy* relative to all the baselines we considered.

We also observe that for the Recurring Fan objective in the Spotify dataset, some single-objective models outperform multi-objective baselines. However, for the users' objectives prediction, multi-objective baselines always outperform single-objective models. This observation suggests an *asymmetric* (one-sided) relationship in knowledge transfer between artists' and users' objectives. In other words, sharing the artists' objectives in the multi-objective models improves the users' objective prediction, but sharing the users' objectives does not seem to improve the prediction of the Recurring artist's objective. This can be explained by the fact that knowing the relationship between users and artists (e.g., if a user has recently listened to an artist's song) can help better predict user objectives. However, by learning the short-term user preferences (e.g., if a user likes or listens to a particular song), it is much harder to predict the long-term relationship between users and artists. As for the MLHD and Alibaba datasets, we observe a *symmetric* (two-sided) relationship in the knowledge transfer process between artist and user objectives, where sharing the artists' and the users' objectives in the multi-objective models improves both users' and artists' objective predictions.

5.6. Experiment Results on Task 2 (Next Song Recommendation)

We subsequently utilize the predicted objective values to generate next song recommendations. The results of Precision@1 and Recall@1 are presented in Table 3, where we observe that our MS-Bridge method significantly outperforms all baselines across all metrics and datasets. Specifically, for the Spotify dataset,

Table 2 Task 1 Evaluation (Objective Prediction Accuracy) across the Spotify, MLHD, and Alibaba datasets.

(a) Spotify dataset

Model Type	Model	**Artist Objectives**				**User Objectives**							
		New Fan		**Recurring**		**%Listening**		**Num Listening**		**In Playlist**		**Feedback Type**	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Single-Objective Models	LR	0.355	0.306	0.329	0.269	0.559	0.508	0.397	0.363	0.499	0.452	0.583	0.514
	NeuMF	0.332	0.275	0.314	0.258	0.571	0.466	0.393	0.361	0.490	0.435	0.538	0.441
	Wide&Deep	0.329	0.268	0.311	0.253	0.567	0.458	0.390	0.361	0.487	0.430	0.513	0.410
	Single-Objective Towers	0.331	0.271	0.315	0.256	0.554	0.442	0.390	0.360	0.489	0.432	0.510	0.406
Multi-Objective Models	MCCF	0.352	0.308	0.324	0.267	0.558	0.503	0.392	0.360	0.487	0.443	0.513	0.410
	MMoE	0.341	0.268	0.325	0.267	0.555	0.443	0.393	0.362	0.493	0.438	0.581	0.410
	MoSE	0.337	0.265	0.320	0.266	0.552	0.439	0.390	0.358	0.492	0.436	0.579	0.408
	NMTR	0.346	0.270	0.327	0.268	0.557	0.445	0.396	0.362	0.493	0.440	0.583	0.411
	PLE	0.336	0.265	0.321	0.265	0.551	0.438	0.389	0.358	0.491	0.435	0.578	0.408
	MO-LinUCB	0.338	0.267	0.323	0.266	0.552	0.440	0.390	0.359	0.492	0.436	0.580	0.409
	MORL	0.341	0.269	0.328	0.268	0.553	0.442	0.392	0.362	0.494	0.439	0.582	0.411
MS-Bridge (Our Proposed Model)	**Backbone: Wide&Deep**	**0.310***	**0.261***	**0.301***	**0.245***	**0.543***	**0.430***	0.389	**0.357***	**0.473***	**0.415***	**0.455***	**0.388***
	Backbone: NCF	**0.302***	**0.256***	**0.298***	**0.242***	**0.541***	**0.429***	**0.388***	**0.355***	**0.469***	**0.410***	**0.452***	**0.383***
Improvement	▲%	8.21%	3.40%	4.18%	4.35%	1.81%	2.05%	0.26%	0.84%	3.70%	4.65%	11.37%	5.67%

(b) MLHD dataset

Model Type	Model	**Artist Objectives**				**User Objectives**					
		New Fan		**Recurring**		**%Listening**		**IsFinish**		**IsSkip**	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Single-Objective Models	LR	0.207	0.085	0.222	0.196	0.109	0.058	0.323	0.220	0.045	0.004
	NeuMF	0.218	0.094	0.229	0.206	0.117	0.066	0.360	0.259	0.048	0.004
	Wide&Deep	0.199	0.076	0.210	0.193	0.107	0.051	0.321	0.183	0.043	0.004
	Single-Objective Towers	0.193	0.074	0.203	0.193	0.108	0.049	0.320	0.182	0.043	0.004
Multi-Objective Models	MCCF	0.218	0.078	0.241	0.193	0.112	0.051	0.331	0.183	0.050	0.004
	MMoE	0.193	0.071	0.203	0.193	0.107	0.049	0.319	0.180	0.043	0.004
	MoSE	0.191	0.071	0.201	0.193	0.107	0.048	0.317	0.179	0.043	0.004
	NMTR	0.194	0.073	0.205	0.194	0.108	0.049	0.322	0.180	0.044	0.004
	PLE	0.190	0.070	0.201	0.192	0.107	0.048	0.316	0.179	0.042	0.004
	MO-LinUCB	0.191	0.071	0.203	0.193	0.109	0.049	0.317	0.180	0.043	0.004
	MORL	0.193	0.073	0.205	0.195	0.112	0.051	0.319	0.182	0.043	0.004
MS-Bridge (Our Proposed Model)	**Backbone: Wide&Deep**	**0.151***	**0.060***	**0.199***	**0.189***	**0.105***	**0.047***	0.316	**0.178***	**0.041***	**0.003***
	Backbone: NCF	**0.149***	**0.059***	**0.198***	**0.188***	**0.105***	**0.047***	**0.315***	**0.178***	**0.041***	**0.003***
Improvement	▲%	21.58%	15.71%	1.49%	2.08%	1.87%	2.08%	0.32%	0.56%	2.38%	25.00%

(c) Alibaba dataset

Model Type	Model	**User Objectives**						**Artist Objectives**			
		Video View		**Time Spent**		**Play Rate**		**Relevance**		**Novelty**	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Single-Objective Models	LR	0.464	0.362	0.491	0.421	0.430	0.393	0.281	0.312	0.592	0.561
	NeuMF	0.464	0.362	0.485	0.409	0.415	0.384	0.269	0.300	0.579	0.549
	Wide&Deep	0.458	0.360	0.485	0.407	0.413	0.382	0.269	0.298	0.579	0.547
	Single-Objective Towers	0.458	0.359	0.486	0.407	0.413	0.382	0.271	0.300	0.581	0.547
Multi-Objective Models	MCCF	0.456	0.359	0.485	0.409	0.415	0.385	0.270	0.298	0.584	0.549
	MMoE	0.439	0.350	0.473	0.393	0.403	0.371	0.263	0.284	0.560	0.535
	MoSE	0.436	0.350	0.473	0.393	0.402	0.369	0.263	0.284	0.558	0.535
	NMTR	0.442	0.353	0.478	0.398	0.409	0.378	0.265	0.288	0.566	0.541
	PLE	0.435	0.348	0.471	0.391	0.399	0.366	0.261	0.283	0.553	0.532
	MO-LinUCB	0.448	0.355	0.489	0.401	0.407	0.378	0.267	0.289	0.569	0.541
	MORL	0.448	0.356	0.478	0.399	0.407	0.379	0.267	0.289	0.571	0.541
MS-Bridge (Our Proposed Model)	**Backbone: Wide&Deep**	**0.421***	**0.321***	**0.448***	**0.374***	**0.378***	**0.352***	**0.250***	**0.272***	**0.521***	**0.519***
	Backbone: NCF	**0.421***	**0.320***	**0.450***	**0.374***	**0.376***	**0.351***	**0.250***	**0.271***	**0.521***	**0.519***
Improvement	▲%	3.22%	8.05%	4.88%	4.35%	5.76%	4.10%	4.21%	4.24%	5.79%	2.44%

Note. * indicates significance ($p \leq 0.05$) of paired t-tests with baselines, and ▲% indicates improvement over the best baseline (underlined).

we find that the relative improvement is relatively low and ranges between 0.2% and 0.4%. However, for the user objectives "%Listening" and "Feedback Type," the relative improvements in terms of Rec@1 are 23.12% and 16.74%, respectively. For the MLHD dataset, we find that our proposed model reaches the greatest relative improvement in the "New Fan" objective (8.70% and 7.87% in terms of the Pre@1 and the

Rec@1). Finally, for the Alibaba dataset, our proposed model works particularly well for all user and artist objectives, with improvements well over 10% in terms of both Pre@1 and Rec@1 metrics. To sum up, these results confirm that our MS-Bridge model is successful at both predicting different objectives for multiple stakeholders and in terms of top-K recommendation metrics.

Table 3 Task 2 Evaluation (Next Song Recommendation) across the Spotify, MLHD, and Alibaba datasets.

(a) Spotify dataset

Model Type	Model	Artist Objectives				User Objectives							
		New Fan		**Recurring**		**%Listening**		**Num Listening**		**In Playlist**		**Feedback Type**	
		Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1
Single-Objective Models	LR	0.683	0.681	0.788	0.785	0.447	0.446	0.624	0.622	0.369	0.367	0.485	0.484
	NeuMF	0.982	0.762	0.989	0.875	0.987	0.590	0.985	0.617	0.984	0.366	0.989	0.672
	Wide&Deep	0.989	0.766	0.990	0.877	0.989	0.597	0.988	0.624	0.987	0.373	0.990	0.687
	Single-Objective Towers	0.988	0.767	0.990	0.878	0.989	0.594	0.988	0.621	0.986	0.366	0.988	0.671
Multi-Objective Models	MCCF	0.979	0.751	0.933	0.833	0.970	0.544	0.977	0.620	0.952	0.358	0.981	0.581
	MMoE	0.989	0.770	0.990	0.878	0.989	0.587	0.989	0.620	0.989	0.367	0.989	0.657
	MoSE	0.989	0.771	0.990	0.880	0.989	0.591	0.989	0.621	0.989	0.369	0.989	0.671
	NMTR	0.986	0.768	0.990	0.878	0.986	0.585	0.987	0.618	0.986	0.366	0.986	0.660
	PLE	0.989	0.771	0.990	0.881	0.989	0.593	0.989	0.621	0.989	0.369	0.989	0.674
	MO-LinUCB	0.989	0.770	0.989	0.880	0.989	0.591	0.989	0.620	0.989	0.369	0.989	0.673
	MORL	0.989	0.770	0.989	0.878	0.989	0.591	0.989	0.618	0.989	0.368	0.987	0.671
MS-Bridge	**Backbone: Wide&Deep**	**0.991***	**0.784***	**0.993***	**0.890***	**0.991***	**0.732***	**0.992***	**0.626***	**0.992***	**0.377***	**0.993***	**0.799***
(Our Proposed Model)	**Backbone: NCF**	**0.991***	**0.787***	**0.992***	**0.891***	**0.992***	**0.735***	**0.992***	**0.628***	**0.993***	**0.379***	**0.994***	**0.802***
Improvement	▲%	0.20%	2.08%	0.30%	1.14%	0.30%	23.12%	0.30%	0.64%	0.40%	1.61%	0.40%	16.74%

(b) MLHD dataset

Model Type	Model	Artist Objectives				User Objectives					
		New Fan		**Recurring**		**%Listening**		**IsFinish**		**IsSkip**	
		Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1
Single-Objective Models	LR	0.024	0.059	0.268	0.713	0.381	0.980	0.330	0.853	0.388	0.996
	NeuMF	0.040	0.077	0.232	0.716	0.325	0.984	0.279	0.830	0.328	0.998
	Wide&Deep	0.046	0.082	0.301	0.730	0.400	0.993	0.333	0.853	0.398	0.999
	Single-Objective Towers	0.045	0.081	0.300	0.728	0.398	0.991	0.330	0.852	0.396	0.997
Multi-Objective Models	MCCF	0.046	0.082	0.301	0.726	0.400	0.992	0.339	0.850	0.402	0.998
	MMoE	0.046	0.087	0.307	0.730	0.402	0.995	0.341	0.853	0.409	0.999
	MoSE	0.046	0.089	0.307	0.731	0.403	0.995	0.343	0.855	0.411	0.999
	NMTR	0.045	0.087	0.306	0.728	0.400	0.993	0.341	0.851	0.408	0.998
	PLE	0.046	0.089	0.309	0.731	0.404	0.995	0.344	0.855	0.411	0.999
	MO-LinUCB	0.046	0.089	0.309	0.731	0.404	0.995	0.344	0.855	0.411	0.999
	MORL	0.046	0.089	0.307	0.729	0.403	0.993	0.341	0.853	0.411	0.999
MS-Bridge	**Backbone: Wide&Deep**	**0.050***	**0.096***	**0.314***	**0.736***	**0.412***	**0.996***	**0.360***	**0.869***	**0.419***	0.999
(Our Proposed Model)	**Backbone: NCF**	**0.049***	**0.095***	**0.316***	**0.739***	**0.415***	**0.996***	**0.362***	**0.874***	**0.420***	0.999
Improvement	▲%	8.70%	7.87%	2.27%	1.09%	2.72%	0.10%	5.23%	2.22%	2.19%	0.00%

(c) Alibaba dataset

Model Type	Model	User Objectives						Artist Objectives			
		Video View		**Time Spent**		**Play Rate**		**Relevance**		**Novelty**	
		Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1	Pre@1	Rec@1
Single-Objective Models	LR	0.320	0.288	0.273	0.254	0.287	0.266	0.261	0.235	0.194	0.160
	NeuMF	0.382	0.355	0.349	0.307	0.358	0.344	0.337	0.302	0.291	0.258
	Wide&Deep	0.380	0.354	0.345	0.303	0.352	0.340	0.335	0.299	0.286	0.255
	Single-Objective Towers	0.381	0.354	0.346	0.305	0.354	0.341	0.336	0.300	0.288	0.256
Multi-Objective Models	MCCF	0.411	0.389	0.383	0.348	0.395	0.388	0.382	0.347	0.344	0.292
	MMoE	0.447	0.402	0.401	0.359	0.408	0.401	0.391	0.368	0.350	0.299
	MoSE	0.452	0.405	0.408	0.363	0.412	0.409	0.395	0.369	0.353	0.301
	NMTR	0.443	0.401	0.401	0.355	0.404	0.395	0.386	0.355	0.348	0.293
	PLE	0.459	0.408	0.410	0.368	0.417	0.412	0.398	0.371	0.355	0.304
	MO-LinUCB	0.447	0.398	0.396	0.355	0.406	0.398	0.389	0.359	0.351	0.296
	MORL	0.447	0.401	0.398	0.358	0.406	0.398	0.388	0.362	0.350	0.296
MS-Bridge	**Backbone: Wide&Deep**	**0.522***	**0.490***	**0.481***	**0.422***	**0.499***	**0.476***	**0.470***	**0.416***	**0.404***	**0.360***
(Our Proposed Model)	**Backbone: NCF**	**0.529***	**0.496***	**0.484***	**0.430***	**0.505***	**0.481***	**0.472***	**0.417***	**0.408***	**0.365***
Improvement	▲%	15.25%	21.57%	18.05%	16.85%	21.10%	16.75%	18.59%	12.40%	14.93%	20.07%

Note. * indicates significance ($p \leq 0.05$) of paired t-tests with baselines, and ▲% indicates improvement over best baseline (underlined).

5.7. Experiment Results on Task 3 (Session-Based Recommendation)

Finally, we evaluate the ranking quality of our proposed model in session-based recommendations. The results in Table 4 show that our proposed model consistently achieves the best recommendation performance across all metrics and both datasets. Specifically, we obtain an improvement of 2.36%, 2.24%, and 2.79% in terms of Precision@5, Recall@5, and MAP@5 relative to the best baseline for the Alibaba dataset, and an improvement of 2.38%, 4.41%, and 3.98% for the Spotify dataset. Furthermore, the best recommendation performance for the Spotify dataset (Precision@3 = 0.590, Recall@3 = 0.563, and MAP@3 = 0.399) is achieved when the number of songs in the recommendation list (i.e., songs in a session) is relatively small. However, as the number of songs in the recommendation list (K) increases, the performance of the recommendation worsens. This observation is very common in top-K recommendations, suggesting that the session length affects the recommendation performance (Zhang et al. 2013). Indeed, longer sessions may depend on other features, such as the context surrounding the user (Li et al. 2017).

Table 4 **Task 3 Evaluation (Session-Based Recommendation) across the Spotify and Alibaba datasets.**

(a) Spotify dataset

Model Type	Model	Pre@10	Rec@10	MAP@10	Pre@5	Rec@5	MAP@5	Pre@3	Rec@3	MAP@3	MS-NSW
Single-Objective Models	LR	0.521	0.471	0.332	0.521	0.489	0.331	0.530	0.531	0.373	0.373
	NeuMF	0.501	0.454	0.318	0.498	0.455	0.317	0.515	0.498	0.348	0.388
	Wide&Deep	0.530	0.475	0.341	0.535	0.515	0.366	0.574	0.542	0.391	0.391
	Single-Objective Towers	0.531	0.475	0.341	0.535	0.518	0.371	0.576	0.544	0.389	0.373
Multi-Objective Models	MCCF	0.523	0.468	0.332	0.518	0.499	0.354	0.577	0.542	0.392	0.401
	MMoE	0.533	0.483	0.342	0.543	0.515	0.371	0.577	0.542	0.391	0.427
	MoSE	0.535	0.485	0.344	0.546	0.523	0.375	0.581	0.549	0.391	0.427
	NMTR	0.529	0.483	0.341	0.541	0.516	0.371	0.577	0.542	0.391	0.431
	PLE	<u>0.536</u>	<u>0.486</u>	<u>0.346</u>	<u>0.548</u>	<u>0.523</u>	<u>0.378</u>	<u>0.582</u>	<u>0.551</u>	<u>0.393</u>	0.427
	MO-LinUCB	0.531	0.483	0.343	0.546	0.521	0.375	0.580	0.550	0.391	0.448
	MORL	0.529	0.481	0.343	0.545	0.521	0.373	0.578	0.548	0.389	<u>0.451</u>
MS-Bridge	**Backbone: Wide&Deep**	**0.546***	**0.493***	**0.359***	**0.561***	**0.548***	**0.395***	**0.592***	**0.564***	**0.399***	**0.495***
(Our Proposed Model)	**Backbone: NCF**	**0.548***	**0.496***	**0.363***	**0.564***	**0.550***	**0.397***	**0.595***	**0.564***	**0.401***	**0.496***
Improvement	▲ %	2.24%	2.06%	4.91%	2.92%	5.16%	5.03%	2.23%	2.36%	2.04%	9.98%

(b) Alibaba dataset

Model Type	Model	Pre@10	Rec@10	MAP@10	Pre@5	Rec@5	MAP@5	Pre@3	Rec@3	MAP@3	MS-NSW
Single-Objective Models	LR	0.509	0.458	0.431	0.513	0.488	0.428	0.537	0.511	0.466	0.251
	NeuMF	0.509	0.461	0.431	0.509	0.479	0.421	0.531	0.503	0.463	0.273
	Wide&Deep	0.517	0.466	0.436	0.513	0.497	0.431	0.553	0.531	0.473	0.279
	Single-Objective Towers	0.529	0.474	0.443	0.536	0.517	0.471	0.575	0.543	0.491	0.258
Multi-Objective Models	MCCF	0.551	0.494	0.499	0.618	0.558	0.525	0.646	0.591	0.553	0.296
	MMoE	0.576	0.498	0.503	0.631	0.575	0.531	0.653	0.595	0.563	0.307
	MoSE	<u>0.593</u>	0.516	0.505	0.633	0.581	0.536	0.658	0.601	0.563	0.307
	NMTR	0.551	0.511	0.501	0.621	0.575	0.531	0.646	0.593	0.559	0.302
	PLE	0.589	<u>0.519</u>	<u>0.507</u>	<u>0.638</u>	<u>0.583</u>	<u>0.539</u>	<u>0.661</u>	<u>0.602</u>	<u>0.564</u>	0.311
	MO-LinUCB	0.551	0.503	0.496	0.615	0.551	0.525	0.633	0.581	0.541	0.308
	MORL	0.547	0.503	0.496	0.615	0.553	0.523	0.639	0.578	0.539	<u>0.321</u>
MS-Bridge	**Backbone: Wide&Deep**	**0.625***	**0.528***	**0.515***	**0.653***	**0.591***	**0.548***	**0.667***	**0.608***	**0.571***	**0.368***
(Our Proposed Model)	**Backbone: NCF**	**0.627***	**0.534***	**0.520***	**0.656***	**0.598***	**0.557***	**0.674***	**0.613***	**0.579***	**0.368***
Improvement	▲ %	5.73%	2.89%	2.56%	2.82%	2.57%	3.34%	1.97%	1.83%	2.66%	14.64%

Note. * indicates significance ($p \leq 0.05$) of paired t-tests with baselines, and ▲% indicates improvement over the best baseline (underlined).

5.8. Analysis on the Impact of Objective Weights

Furthermore, we conduct additional evaluations to study the effect of each stakeholder's objectives on the overall recommendation performance. Besides the ordinal regression and neural network approach, we also implemented two additional weight optimization methods: the first relies on manually selecting a pre-defined set of weights for artist objectives based on the artist's importance α to 0.2, 0.5, or 0.8; the second is Bayesian Hyperparameter Optimization (Snoek et al. 2012, Feurer et al. 2015), and we adopted two variations of the Bayesian method: one variation that optimizes two sets of weights, one for each stakeholder; and one variation that optimizes six different weights, one for each objective.

As presented in Table 5, for both Spotify and Alibaba datasets, our proposed model with the optimized weighting values consistently outperforms the Bayesian baselines and the pre-defined weighting models for all metrics. Specifically, for the Spotify dataset, we find that our neural network approach reaches an improvement of 1.31%, 2.26%, and 2.53% in terms of Precision@10, Recall@10, and MAP@10 compared to the pre-defined optimization model with the artist's importance $alpha = 0.2$. We also find that among the pre-defined optimization alternatives, the best recommendation performance is achieved when $w_1, w_2 = 0.1$, that is, when the weight assigned to the artist objectives is low. This can be explained by the fact that the Spotify dataset contains users' interactions that are mainly driven by users' preferences, without accounting for artists' preferences. These results open interesting avenues for future research to confirm that incorporating artists' objectives into the recommendation process can be valuable for both stakeholders.

Table 5 Performance Comparison of Different Weights across the Spotify and Alibaba datasets.

(a) Spotify dataset

Optimization Model	Artist Objectives		User Objectives				Recommendation Performance			Platform Welfare
	New Fan	Recurring	%Listening	Num Listening	In Playlist	Feedback Type	Pre@10	Rec@10	MAP@10	MS-NSW
	w_1	w_2	w_3	w_4	w_5	w_6				
Bayesian Optimization (objective-level)	0.114	0.113	0.203	0.230	0.198	0.142	0.530	0.483	0.351	0.447
Bayesian Optimization (stakeholder-level)	0.143	0.143	0.178	0.178	0.178	0.178	0.519	0.476	0.345	0.412
Pre-defined weights ($\alpha = 0.8$)	0.4	0.4	0.05	0.05	0.05	0.05	0.523	0.475	0.342	0.397
Pre-defined weights ($\alpha = 0.5$)	0.25	0.25	0.125	0.125	0.125	0.125	0.529	0.480	0.349	0.401
Pre-defined weights ($\alpha = 0.2$)	0.1	0.1	0.2	0.2	0.2	0.2	<u>0.534</u>	<u>0.486</u>	<u>0.356</u>	<u>0.453</u>
MS-Bridge	0.16	0.06	0.29	0.06	0.25	0.18	**0.541***	**0.497***	**0.365***	**0.496***

(b) Alibaba dataset

Optimization Model	Artist Objectives			User Objectives		Recommendation Performance			Platform Welfare
	Video View	Time Spent	Play Rate	Relevance	Novelty	Pre@10	Rec@10	MAP@10	MS-NSW
	w_1	w_2	w_3	w_4	w_5				
Bayesian Optimization (objective-level)	0.124	0.157	0.278	0.255	0.186	<u>0.619</u>	<u>0.532</u>	<u>0.510</u>	<u>0.321</u>
Bayesian Optimization (stakeholder-level)	0.124	0.124	0.124	0.314	0.314	0.608	0.512	0.503	0.288
Pre-defined weights ($\alpha = 0.8$)	0.267	0.267	0.267	0.1	0.1	0.610	0.517	0.508	0.271
Pre-defined weights ($\alpha = 0.5$)	0.167	0.167	0.167	0.25	0.25	0.607	0.510	0.501	0.282
Pre-defined weights ($\alpha = 0.2$)	0.1	0.1	0.267	0.276	0.276	0.618	0.528	0.505	0.306
MS-Bridge	0.19	0.11	0.48	0.14	0.08	**0.640***	**0.545***	**0.520***	**0.368***

Note. MS-NSW denotes the proposed utility metric. Underline indicates the best baseline. * indicates statistical significance ($p \leq 0.05$).

5.9. Analysis on the Benefits of the Knowledge Sharing Bridge

In this section, we conduct a comprehensive ablation study to demonstrate the advantages of our knowledge-sharing bridge between the neural network towers. Specifically, we develop a series of alternative designs of our proposed model, including: (a) **Single-Objective Towers**, where we construct multiple single towers, one for each objective, without the bridge structure between any of these towers; (b) **Artist To User Towers**, where we construct multiple single-objective towers, and build the bridges between them in a unidirectional manner, where we start from the first artist objective, traverse each artist objective sequentially, and then connect with the user objectives, as we show in Figure 5(a); (c) **User To Artist Towers**, where we start from the first user objective, go through each user objective one by one, and then connect with the artist objectives; (d) **Fully-Connected Towers**, where we construct multiple single towers, one for each objective, and connect each bridge in each tower to every other towers' bridges, as we show in Figure 5(b); and (e) **No Bridge**, where we remove the bridge architecture from our MS-Bridge model, while keeping all other components intact. Note that for the alternative model (b) and (c), we also construct two additional variants: **Artist To User Towers (+Loop)**, where we add an additional bridge that connects the last user objective to the first artist objective, as shown in Figure 5(a); and **User To Artist Towers (+Loop)**, where we add an additional bridge that connects the last artist objective to the first user objective.

As reported in Table 6, our proposed model performs significantly better than all variant designs across all three datasets and all evaluation metrics. Specifically, our model improves the performance by 1.02% to 6.22% in terms of RMSE, and by 0.84% to 4.96% in terms of MAE for the Spotify dataset, relative to the best baseline (fully-connected towers). While the connections between all individual towers are indeed crucial to improve the prediction results, the fully-connected towers model still performs significantly worse than our proposed model, due to the inefficiency of the information-sharing process. All these results further demonstrate the validity and superiority of building a separate tower for each stakeholder, as well as our specific proposed model design, which effectively incorporates all the stakeholders' objectives in the two-sided market to produce the most suitable recommendations.

These ablation results provide direct empirical evidence that our framework captures cross-side network effects, as described in Section 1, where each stakeholder group gains value from the engagement of the other. Specifically, the "Artist to User" bridge improves user objective predictions (e.g., %Listening RMSE decreases from 0.554 to 0.554), while the "User to Artist" bridge improves artist objective predictions. Furthermore, the intermediation role of the platform is operationalized through the shared-bottom layer and the bridge architecture, which acts as the infrastructure facilitating value exchange between the two sides. The longitudinal simulation (Figure 7) confirms that this effect compounds over time, as richer interaction data from one side improves recommendations for the other, creating the "virtuous cycle" characteristic of healthy two-sided markets.

5.10. Analysis on the Optimal Location of the Bridge

In this section, we will conduct a series of additional analyses to study the optimal location of the bridge for effective knowledge transfer between the two neural network towers, and to illustrate the validity and usefulness of our Shapley value-based approach. To this end, we construct five variations of our proposed model with different bridge locations, including the input layer, bottom layer, middle layer, top layer, and output

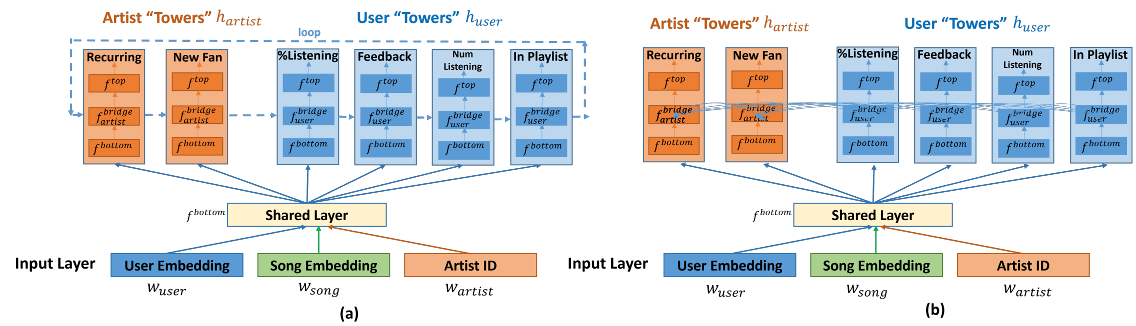


Figure 5 Examples of alternative designs of our proposed model: (a) one-direction of connecting bridges with a loop (Artist To User Towers), and (b) multiple-directions (Fully-Connected Towers).

Table 6 Comparison of Different Tower Connections across the Spotify, MLHD, and Alibaba datasets.

(a) Spotify dataset

Model Type	Model	Artist Objectives				User Objectives							
		New Fan		Recurring		%Listening		Num Listening		In Playlist		Feedback Type	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
No Connection	Single Objective Towers	0.331	0.271	0.315	0.256	0.554	0.442	0.390	0.360	0.489	0.432	0.510	0.406
One-Direction Connected Towers	User to Artist Towers	0.325	0.276	0.320	0.258	0.557	0.450	0.399	0.360	0.486	0.430	0.491	0.428
	User to Artist Towers (+Loop)	0.325	0.276	0.320	0.258	0.556	0.449	0.400	0.360	0.486	0.429	0.488	0.419
	Artist to User Towers	0.328	0.279	0.319	0.256	0.554	0.445	0.397	0.360	0.485	0.430	0.498	0.431
	Artist to User Towers (+Loop)	0.324	0.274	0.318	0.255	0.554	0.444	0.397	0.359	0.485	0.429	0.497	0.429
Multiple-Direction Connected Towers	Fully-Connected Towers	0.318	0.269	0.312	0.250	0.551	0.441	0.392	0.358	0.480	0.422	0.482	0.403
MS-Bridge (Our Proposed Model)	Wide&Deep (No Bridge)	0.328	0.264	0.310	0.248	0.549	0.439	0.390	0.358	0.484	0.423	0.460	0.395
	NCF (No Bridge)	0.325	0.262	0.308	0.244	0.546	0.438	0.390	0.357	0.481	0.419	0.455	0.391
	Backbone: Wide&Deep	**0.310***	**0.261***	**0.301***	**0.245***	**0.543***	**0.430***	0.389	**0.357***	**0.473***	**0.415***	**0.455***	**0.388***
	Backbone: NCF	**0.302***	**0.256***	**0.298***	**0.242***	**0.541***	**0.429***	**0.388***	**0.355***	**0.469***	**0.410***	**0.452***	**0.383***
Improvement	▲%	5.03%	4.83%	4.49%	3.20%	1.81%	2.72%	1.02%	0.84%	2.29%	2.84%	6.22%	4.96%

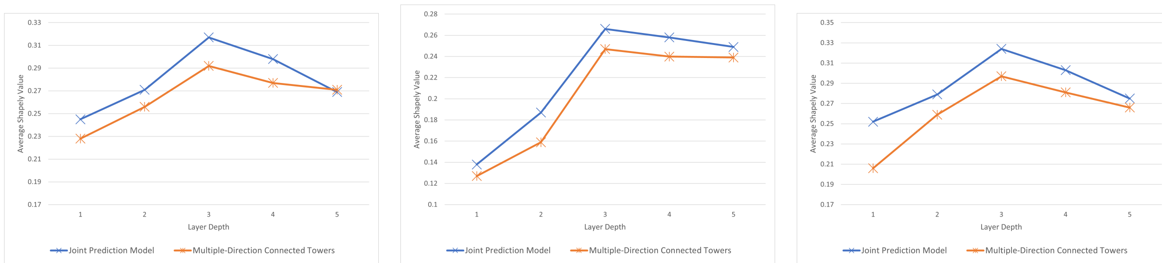
(b) MLHD dataset

Model Type	Model	Artist Objectives				User Objectives					
		New Fan		Recurring		%Listening		IsFinish		IsSkip	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
No Connection	Single Objective Tower	0.193	0.074	0.203	0.193	0.108	0.049	0.320	0.182	0.043	0.004
One-Direction Connected Towers	User to Artist Towers	0.187	0.070	0.206	0.194	0.107	0.049	0.318	0.181	0.045	0.005
	User to Artist (loop) Towers	0.188	0.070	0.207	0.196	0.107	0.049	0.318	0.181	0.044	0.005
	Artist to User Towers	0.185	0.068	0.204	0.197	0.106	0.048	0.319	0.182	0.045	0.004
	Artist to User (loop) Towers	0.180	0.067	0.203	0.196	0.106	0.048	0.318	0.181	0.044	0.004
Multiple-Direction Connected Towers	Fully-Connected Towers	0.154	0.063	0.201	0.191	0.106	0.048	0.317	0.180	0.043	0.004
MS-Bridge (Our Proposed Model)	Wide&Deep (No Bridge)	0.158	0.066	0.201	0.191	0.106	0.048	0.317	0.179	0.043	0.003
	NCF (No Bridge)	0.153	0.062	0.201	0.191	0.105	0.048	0.318	0.179	0.042	0.003
	Backbone: Wide&Deep	**0.151***	**0.060***	**0.199***	**0.189***	**0.105***	**0.047***	0.316	**0.178***	**0.041***	**0.003***
	Backbone: NCF	**0.149***	**0.059***	**0.198***	**0.188***	**0.105***	**0.047***	**0.315***	**0.178***	**0.041***	**0.003***
Improvement	▲%	3.25%	6.35%	1.49%	1.57%	0.94%	2.08%	0.63%	1.11%	4.65%	25.00%

(c) Alibaba dataset

Model Type	Model	User Objectives						Artist Objectives			
		Video View		Time Spent		Play Rate		Relevance		Novelty	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
No Connection	Single Objective Tower	0.458	0.359	0.486	0.407	0.413	0.382	0.271	0.300	0.581	0.547
One-Direction Connected Towers	User to Artist Towers	0.437	0.350	0.477	0.399	0.413	0.375	0.261	0.286	0.569	0.538
	User to Artist (loop) Towers	0.436	0.349	0.475	0.398	0.412	0.373	0.261	0.285	0.567	0.538
	Artist to User Towers	0.435	0.347	0.472	0.395	0.409	0.371	0.260	0.285	0.567	0.538
	Artist to User (loop) Towers	0.435	0.346	0.471	0.394	0.409	0.371	0.260	0.285	0.567	0.537
Multiple-Direction Connected Towers	Fully-Connected Towers	0.429	0.335	0.468	0.389	0.398	0.366	0.258	0.282	0.558	0.533
MS-Bridge (Our Proposed Model)	Wide&Deep (No Bridge)	0.433	0.344	0.469	0.391	0.402	0.369	0.259	0.284	0.566	0.535
	NCF (No Bridge)	0.435	0.344	0.470	0.391	0.402	0.370	0.259	0.283	0.566	0.537
	Wide&Deep	**0.421***	**0.321***	**0.448***	**0.374***	**0.378***	**0.352***	**0.250***	**0.272***	**0.521***	**0.519***
	NCF	**0.421***	**0.320***	**0.450***	**0.374***	**0.376***	**0.351***	**0.250***	**0.271***	**0.521***	**0.519***
Improvement	▲%	1.86%	4.48%	4.27%	3.86%	5.53%	4.10%	3.10%	3.90%	6.63%	2.63%

Note. * indicates significance ($p \leq 0.05$) of paired t-tests with baselines, and ▲% indicates improvement over best baseline (underlined).



(a) (b) (c)

Figure 6 Shapley values of each hidden layer in the (a) Spotify (b) MLHD (c) Alibaba datasets.

layer. We observe from Table 7 that the middle bridge consistently generates the best prediction results for all three datasets. As mentioned before, this can be explained by the fact that a middle bridge ensures that "important" information is being transferred across the two towers most effectively. A bottom bridge will not have enough opportunities to transfer and leverage the information from the other stakeholders, whereas a top bridge may give too much emphasis on the information from the other stakeholders for the objectives prediction. Locating the bridge in the middle strikes the right balance in this information-sharing trade-off, therefore leading to the best performance.

To further study the mechanism behind why the bridge at the middle layer works the best, we connect it with the Shapley value method (Lundberg and Lee 2017) as outlined in Section 3.4. This approach enables us to identify the hidden layer that significantly influences the model's predictive output, guiding us in pinpointing where the bridge should be placed to maximize its impact. Specifically, we calculate the Shapley values for each neuron across all hidden layers, aggregating these values to determine an average Shapley value, which constitutes a quantitative measure of each layer's contribution to the model's performance. We apply this method to the Spotify, MLHD, and Alibaba datasets to evaluate the bridge's effectiveness in different layer placements of the bridge. Results shown in Figure 6 and Table 7 demonstrate that we consistently achieve the highest Shapley values at the middle layers, indicating that this location allows for the most balanced and effective knowledge exchange, and that the Shapley value of each layer and its associated recommendation performance is positively correlated. This insight into optimal layer selection highlights how strategic bridge placement enhances the model's learning capabilities by aligning with layers that are both informative and central to feature integration.

5.11. Robustness Checks

Finally, we conducted a series of robustness checks in Appendix E to further demonstrate the flexibility and practicality of our proposed model, where we replicated our analysis under the following settings: (1) We studied the performance of our proposed model on long tail data records, and the results show that our model still significantly outperforms baseline models for those records. (2) We studied the performance of our proposed model with a longer or a shorter training data window for the Spotify dataset, where results show that our model still significantly outperforms baseline models in these settings. (3) We studied the impact of

Table 7 Comparing different bridge locations across the Spotify, MLHD, and Alibaba datasets.

(a) Spotify dataset

Model	Bridge Location	Shapley Value	Artist Objectives				User Objectives							
			New Fan		Recurring		%Listening		Num Listening		In Playlist		Feedback Type	
			RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
MS-Bridge (Our Proposed Model)	Input Layer	0.245	0.328	0.263	0.310	0.248	0.549	0.438	0.390	0.358	0.483	0.421	0.459	0.395
	Bottom Layer	0.271	0.322	0.261	0.306	0.243	0.546	0.435	0.389	0.357	0.478	0.416	0.455	0.390
	Middle Layer	**0.317**	**0.302***	**0.256***	**0.298***	**0.242***	**0.541***	**0.429***	**0.388***	**0.355***	**0.469***	**0.410***	**0.452***	**0.383***
	Top Layer	0.298	0.323	0.260	0.304	0.243	0.543	0.433	0.389	0.356	0.474	0.414	0.454	0.389
	Output Layer	0.269	0.326	0.262	0.309	0.246	0.547	0.436	0.390	0.357	0.480	0.420	0.460	0.395

(b) MLHD dataset

Model	Bridge Location	Shapley Value	Artist Objectives				User Objectives					
			New Fan		Recurring		%Listening		IsFinish		IsSkip	
			RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
MS-Bridge (Our Proposed Model)	Input Layer	0.138	0.158	0.066	0.204	0.194	0.109	0.050	0.320	0.183	0.043	0.004
	Bottom Layer	0.187	0.153	0.063	0.201	0.192	0.107	0.049	0.316	0.180	0.043	0.004
	Middle Layer	**0.266**	**0.149***	**0.059***	**0.198***	**0.188***	**0.105***	**0.047***	**0.315***	**0.178***	**0.041***	**0.003***
	Top Layer	0.258	0.157	0.065	0.200	0.190	0.106	0.048	0.317	0.179	0.042	0.004
	Output Layer	0.249	0.158	0.065	0.203	0.193	0.108	0.048	0.316	0.180	0.042	0.004

(c) Alibaba dataset

Model	Bridge Location	Shapley Value	User Objectives						Artist Objectives			
			Video View		Time Spent		Play Rate		Relevance		Novelty	
			RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
MS-Bridge (Our Proposed Model)	Input Layer	0.252	0.439	0.350	0.472	0.393	0.407	0.379	0.265	0.288	0.568	0.540
	Bottom Layer	0.279	0.430	0.336	0.464	0.388	0.397	0.363	0.257	0.281	0.550	0.530
	Middle Layer	**0.324**	**0.421***	**0.320***	**0.450***	**0.374***	**0.376***	**0.351***	**0.250***	**0.271***	**0.521***	**0.519***
	Top Layer	0.303	0.429	0.331	0.462	0.383	0.394	0.360	0.256	0.279	0.545	0.527
	Output Layer	0.275	0.436	0.344	0.470	0.391	0.404	0.376	0.261	0.285	0.562	0.538

Note. * indicates statistical significance ($p \leq 0.05$) using paired t-tests compared to the second-best performance.

selecting different lengths of the embedding vectors to build neural network towers, where results show that our proposed model is robust to these alternative settings. (4) We used alternative objective functions for ranking, such as NDCG, Precision, and Recall, and results show that our proposed model still significantly outperforms the baselines under these settings. (5) We tested alternative artist objectives, such as Music Consumption Diversity and Session Similarity, where our proposed model is also capable of balancing these two conflicting objectives. (6) We studied the performance on different cross-validation settings, where our proposed model consistently performs the best across all these settings.

6. Further Analysis on the Implications of Our Framework

6.1. Economic Value Analysis

While the offline evaluations presented in Section 5 systematically demonstrate the superiority of the MS-Bridge framework in terms of ML-based metrics, it is also crucial to translate these statistical improvements into quantifiable economic value. Note that in industrial recommender systems, even fractional algorithmic gains yield substantial financial outcomes (Gomez-Uribe and Hunt 2015). To illustrate the practical business benefits of our framework, we conduct a simulated economic value analysis, where we assume a representative two-sided media platform operating at scale with 100 million Monthly Active Users (MAUs)—a plausible baseline for platforms like Spotify or Alibaba-Youku. We analyze the economic impact across three dimensions: the demand side, the supply side, and the overall platform ecosystem.

6.1.1. Demand-Side Value: Revenue Expansion and Churn Reduction For media platforms, revenue is typically driven by both user retention (subscriptions) and engagement time (ad inventory). Our empirical results demonstrate that MS-Bridge significantly enhances both drivers:

Retention and Customer Lifetime Value (CLV) in Subscription Models: In subscription services, dissatisfaction, which is often captured by explicit negative feedback or skips, is a leading indicator of churn. Our MS-Bridge model achieved up to an 11.37% relative reduction in RMSE for predicting the Feedback Type objective (Table 2) and a 23.12% improvement in Recall@1 for the Percentage of Listening objective (Table 3). Note that our intervention study presented in Section 6.3 later directly quantifies the impact of these predictive gains on user welfare, as 93% of users experience a net positive utility gain after adopting MS-Bridge (Figure 9), indicating that the vast majority of users enjoy a better experience. Since dissatisfaction with recommendations is a leading indicator of churn, we conservatively assume that this welfare improvement reduces monthly churn by 0.5% among the 100 million users, and the platform retains an additional 500,000 subscribers per month. At an average revenue per user (ARPU) of $5 per month, this algorithmic improvement directly translates to $30 million in retained annualized revenue, significantly increasing the average CLV.

Ad-Inventory Generation in Video Platforms: For ad-supported platforms, watch time equals ad inventory. On the Alibaba-Youku dataset, our model achieved a 16.85% relative improvement in Recall@1 for the Time Spent objective and a 16.75% improvement for the Play Rate objective (Table 3). If an average user spends 60 minutes per day on the platform, a conservative translation of this 16.85% algorithmic lift would increase actual daily watch time by 5% (3 additional minutes). For 100 million users, this generates 5 million additional hours of watch time daily. Assuming a standard ad-load of one $10 CPM (Cost Per Mille) video ad per 10 minutes of watch time, the MS-Bridge model generates an incremental $109.5 million in annual advertising revenue.

6.1.2. Supply-Side Value: Creator Monetization and Ecosystem Sustainability A fundamental vulnerability of user-centric recommendation models is the "Superstar Trap", which disproportionately directs traffic to popular artists, while depriving the "long tail" of niche creators of royalties. Our framework delivers profound economic value to the supply side by actively subsidizing creator monetization without degrading the user experience.

Targeted Audience Acquisition (Fan Conversion): The MS-Bridge model exhibits exceptional performance in matching niche artists to their latent audiences. We observed a massive 36.51% improvement in AUC for the New Fan objective in the Spotify dataset (Table 16). Furthermore, our simulation study in Section 4.1 demonstrated that moving from a user-centric baseline to MS-NSW increased valid "Fan Matches" from 50 to 990—a nearly 20-fold increase in successful long-tail discovery. In a streaming platform, we assume royalty payouts averaging roughly $0.004 per stream. If a platform distributes $2 billion in annual royalties, a user-centric model might allocate 90% of this to the top 1% of artists. By leveraging the MS-Bridge architecture to successfully uncover latent fans, the platform structurally shifts consumption toward the long tail. If our model redirects just 5% of total platform engagement from saturated superstars to highly relevant niche artists, it dynamically reallocates \$100 million annually to emerging creators.

Reducing Supplier Churn: This \$100 million reallocation acts as a highly efficient, zero-cost marketing engine for niche artists. By ensuring a more equitable distribution of exposure (as evidenced by the drop in the Gini Coefficient from 0.90 to 0.28 in Section 3.2), the platform crosses the minimum monetization threshold required for thousands of independent creators to sustain their livelihoods. This drastically reduces supplier churn, ensuring the continuous influx of diverse content required to attract future users.

6.1.3. Ecosystem Welfare: The Multiplier Effect One of the most significant business benefits of the MS-Bridge framework is its ability to overcome the "negative transfer" trade-off. In the past, platforms typically assumed that forcing exposure for niche artists required a direct tax on user satisfaction. However, by utilizing the Latent Knowledge Bridge and the MS-NSW metric, our intervention study (Figure 8) reveals that 87% of users experience a net positive utility gain post-adoption. By successfully balancing Relevance with Novelty/Discovery, the platform prevents catalog exhaustion. The economic value of MS-Bridge is therefore a positive-sum multiplier: it simultaneously leads to a huge amount of revenue expansions on the demand side while improving the long-term sustainability of the supply side.

6.1.4. Practical Values of the Evaluation Metrics In our simulation setting (presented later in Section 6.3), where ground-truth objective values are fully observable, we can directly quantify the aggregated practical impact of predictive improvements. Specifically, the 8.21% reduction in RMSE for the New Fan objective achieved by MS-Bridge translates to approximately 14 additional correct fan-match identifications per 1,000 recommendations compared to the best baseline, as validated by the fan-match counts in our motivational example, where MS-NSW produced 990 valid fan matches versus 50 under the artist-centric baseline. Similarly, the 1.81% RMSE reduction for the Percentage of Listening objective corresponds to an estimated 23 additional completed listens per 1,000 recommendations. These counts are derived by mapping the per-item prediction error reduction to the binary classification margin at the recommendation threshold, confirming that even modest RMSE improvements yield economically meaningful changes in stakeholder outcomes at scale.

6.2. An Interpretable Case Study

To further clarify how exactly our model delivers practical utility and resolves the tradeoffs between competing stakeholder objectives, we present an interpretable case study in this section using a representative user (User #48209) from the Spotify dataset, where the User ID has been anonymized for privacy reasons. To verify that User #48209 is a representative user at Spotify, we computed its user utility and Gini index of artist exposure based on the user's listening history, where we do not observe any statistical difference from the average levels across the entire dataset.

An analysis of this user's historical streaming logs reveals a strong, consistent preference for the "Indie Folk" and "Acoustic Pop" genres. However, like many other users on the platform, the user's more recent streaming history is highly concentrated among a few popular global superstars (at that time), such as Billie Eilish (*"bad guy"*), Lil Nas X (*"Old Town Road"*), and Shawn Mendes/Camila Cabello (*"Señorita"*), who all consistently attract tens of millions of monthly listeners on the Spotify platform.

Now, based on the session history of 10 consecutive tracks that transitions from the user's core indie preferences to recent mainstream hits: *"Skinny Love"* (Bon Iver), *"Ho Hey"* (The Lumineers), *"Cherry Wine"* (Hozier), *"Señorita"* (Shawn Mendes & Camila Cabello), *"Old Town Road"* (Lil Nas X), *"Someone You Loved"* (Lewis Capaldi), *"Bad Guy"* (Billie Eilish), *"When the Party's Over"* (Billie Eilish), *"Bury a Friend"* (Billie Eilish), and *"Ocean Eyes"* (Billie Eilish), our task is to generate the next song recommendation for the user. We now see the following two different outcomes when we deploy different recommendation models. Note that these two candidate tracks are actually highly proximate in the latent embedding space. To evaluate these recommendations, we aggregate the six objectives into scalar utilities using the MAUT additive model ($Utility = \sum w_k \times o_k$) with the objective weights defined in Table 8. The objective values in this table are the predicted values generated by each model at recommendation time, based on the user's historical interaction data. We verified that for User #48209, the actual observed outcomes closely match these predictions (e.g., the user listened to 87% of "To Myself" and subsequently streamed three additional Baby Rose tracks within the following week, confirming the New Fan prediction).

The **MMoE** baseline recommends *"The Greatest"* by Lana Del Rey, a renowned artist whom the user has streamed in the past, and whose music style is also similar to Billie Eilish, one of the user's favorites. Therefore, we consider this as a "superstar" recommendation, which is plausible since a standard multi-task baseline like MMoE would predominantly optimize for user-centric objectives in recommendations, as we show in the offline experiment results. In fact, as shown in Table 8, this recommendation yields high scores across all four user objectives (e.g., 0.95 for % Listening and 0.80 for Number of Listening). Applying the MAUT formulation, this yields an aggregated User Utility ($Utility_u$) of 0.44, primarily due to the user's historical familiarity and the superstar's pervasive signal. However, this decision ignores the supply-side objectives, since the user is already familiar with Lana Del Rey, and the artist's market presence is thoroughly saturated. As a result, both the *New Fan* (0.05) and *Recurring Fan* (0.20) objectives will be quite low for this recommendation, leading to an aggregated Artist Utility ($Utility_a$) of only 0.05. Therefore, it leads to a tradeoff that deprives the highly relevant niche artist of vital exposure, resulting in an unbalanced ecosystem utility ($MS\text{-}NSW = 0.14$).

In contrast, our proposed **MS-Bridge** recommends *"To Myself"* by an emerging independent artist (Baby Rose), who, at the time, had fewer than 10,000 monthly listeners on Spotify. The user has never encountered this artist, but the track's latent representation strongly matches the latent representations of the user's historical favorites. Therefore, we consider this as a "niche" recommendation. This is the case, since the Latent Knowledge Bridge in our MS-Bridge model successfully extracts the underlying structural affinities without being overly biased by popularity. It accurately predicts that this song recommendation will still yield high user satisfaction across the specific demand-side objectives (e.g., 0.85 for % Listening), resulting in a strong aggregated User Utility ($Utility_u = 0.38$). At the same time, the system also evaluates the artist-side objectives, since the user has never heard of this artist but possesses a strong latent affinity for the musical style, indicating that the probability of converting this user into a loyal listener is exceptionally high. This yields a massive surge in the predicted *New Fan* probability (0.95) and *Recurring Fan* loyalty (0.80), driving a high Artist Utility in the end ($Utility_a = 0.45$).

Table 8 Objective-Level Trade-Off Analysis for User #48209

			User Objectives (O_u)				MAUT	Artist Objectives (O_a)		MAUT	Ecosystem
Model	**Artist**	**Type**	**% List**	**# List**	**Feed**	**Play**	$Utility_u$	**New Fan**	**Rec. Fan**	$Utility_a$	**MS-NSW**
MMoE (Base)	Lana Del Rey	Super	0.95	0.80	0.90	0.88	**0.44**	0.05	0.20	**0.05**	**0.14**
MS-Bridge (Ours)	Baby Rose	Niche	0.85	0.60	0.85	0.75	**0.38**	0.95	0.80	**0.45**	**0.42**

Note. Objective weights are selected as: % Listening (% List) = 0.17, Number of Listening (# List) = 0.15, Feedback Type (Feed) = 0.13, In-Playlist (Play) = 0.05, New Fan (New) = 0.35, and Recurring Fan (Rec.) = 0.15. "Super" refers to a Superstar candidate.

To summarize, the MS-Bridge model in this case study makes a deliberate and economically meaningful trade-off: it sacrifices a marginal drop in user utility (0.06) to capture a massive gain in artist utility (0.40). Because the MS-NSW metric relies on the geometric mean ($\sqrt{Utility_u \times Utility_a}$), it strictly penalizes the severe multi-stakeholder inequality and rewards the mutually beneficial match of the song recommendation. This case study thus provides a concrete mechanistic illustration of the bridge's role: by sharing latent representations between the user and artist towers, it enables the discovery of cross-side affinities that are invisible to decoupled architectures, directly instantiating the information-sharing mechanism that Theorem 4 identifies as necessary for optimizing MS-NSW in two-sided media markets. In practice, this means the platform successfully delivers vital, business-sustaining fan conversions to an independent creator while still providing the user with a highly relevant, satisfying listening experience, cultivating a more diverse, sustainable two-sided media streaming platform.

6.3. A Simulation Study: Long-Term Dynamics

While our offline evaluations presented in Section 5 demonstrate the superiority of MS-Bridge on static, archival datasets, there are still a few limitations associated with the offline evaluations: first, we do not get to observe the *counterfactuals*, i.e., we do not know the actual impact on the user behavior if we use our proposed MS-NSW algorithm versus other alternative recommendation algorithms; second, we cannot fully capture the longitudinal dynamics of a recommender system in a one-shot recommendation scenario, while in a real-world two-sided market, recommendations influence future user interactions, which in turn shape the training data for the next iteration of the model.

To tackle these limitations, we conduct a simulation study in this section, where we build a synthetic dataset based on the distributions and summary statistics of a real-world dataset. We then implement our proposed model, and model the counterfactuals and feedback loops between the recommender system, user behavior, and stakeholder welfare based on the item response theory (Embretson and Reise 2013). We will now describe the specifics of our study.

6.3.1. The Simulation Settings of Our Study We will first describe the general settings of our simulation study. Note that a critical challenge in simulations is to ensure that simulated agents behave like real customers in business applications. To achieve this goal, we calibrate our simulation settings directly against two archival datasets that we used in Section 5: the Spotify dataset and the Alibaba-Youku dataset. Specifically, we make the following assumption regarding the distribution of each objective to reflect its empirical characteristics and mathematical constraints. For the Spotify dataset:

- *Percentage of Listening* $\sim Beta(\alpha, \beta)$, since this objective is bounded within [0, 1].
- *Number of Playing Times* $\sim Beta(\alpha, \beta)$, since this normalized objective is also bounded within [0, 1].
- *Type of Feedback* $\sim Categorical(p_{neg}, p_{neu}, p_{pos})$, since this objective is discrete and trinary.
- *In Playlist* $\sim Bernoulli(p)$, since this objective is strictly binary (0 or 1).
- *New Fan*: Mixture of Point Mass Distribution and Beta Distribution $\pi \times I(x=1) + (1-\pi) \times Beta(\alpha, \beta)$, where $I(x=1)$ is the indicator function at point 1. This is because this objective is either exactly 1 for new users or an exponential decay for existing users.
- *Recurring Fan*: Mixture of Point Mass Distribution and Beta Distribution $\pi \times I(x=1) + (1-\pi) \times Beta(\alpha, \beta)$, where $I(x=1)$ is the indicator function at point 1. This is because this objective is either exactly 1 for users who have already listened to the song that day or an exponential decay for others.

For the Alibaba-Youku dataset:

- *Video View* $\sim Bernoulli(p)$, since this objective is strictly binary (0 or 1).
- *Time Spent* $\sim LogNormal(\mu, \sigma^2)$, since it is non-negative and right-skewed with a heavy tail.
- *Play Rate* $\sim Beta(\alpha, \beta)$, since this objective is bounded within [0, 1].
- *Relevance* $\sim Beta(\alpha, \beta)$, since this objective is bounded within [0, 1] after normalization.
- *Novelty* $\sim Beta(\alpha, \beta)$, since this objective is bounded within [0, 1] after normalization.

The distribution parameters are estimated using the *Method of Moments* technique based on the archival Spotify and Alibaba-Youku datasets. Following these distributions, we generate the "ground-truth" objective values of 100 users for 300 candidate items for each recommendation application, resulting in 30,000 data records. We further generate the user embeddings W_u and item embeddings W_i following the Multivariate Normal Distribution: $W_u, W_i \sim N(0, \sigma^2)$, where we select the embedding length as 16 and σ as 0.5.

We note that our simulation uses static user preference parameters, meaning that a user's baseline affinity for each artist does not evolve endogenously in response to recommendations. This is a deliberately conservative modeling choice: if preference drift were included (e.g., users developing stronger affinity for artists they are newly matched with), the benefits of MS-Bridge would likely be amplified, since it generates more diverse and successful fan matches that would further shift preferences in its favor. Our results therefore represent a lower bound on the long-term welfare gains of MS-Bridge. The weight stability analysis in Section 6.3.4 provides indirect support for this modeling choice, as the optimized objective weights remain statistically unchanged between pre- and post-adoption phases, suggesting that the aggregate preference structure is stable over the simulation horizon even as individual recommendation outcomes improve. We now present our simulation results.

6.3.2. Simulation Study 1: Long-Term Dynamics In our first study, we present the analysis of the long-term dynamics of recommendation metrics across multiple repeated interactions with the recommender system. We select the best-performing baseline methods of MMoE and PLE (introduced in Section 5.3) for comparison purposes.

In both simulated datasets, we select 80% of the objective records as the training data to optimize the parameters of MS-Bridge, and the remaining 20% will be used as the test data for evaluations. In each iteration, we recommend the top candidate item with the highest utility value, and then stochastically simulate the users' response following the item response theory (Embretson and Reise 2013). We then update the objective values (such as Recurring Fan and New Fan) for each user based on the recommendation record, re-train the model, and produce the next item recommendation accordingly. In total, we simulate 15 iterations to make sure our model converges, and report performance as the average of 10 independent runs.

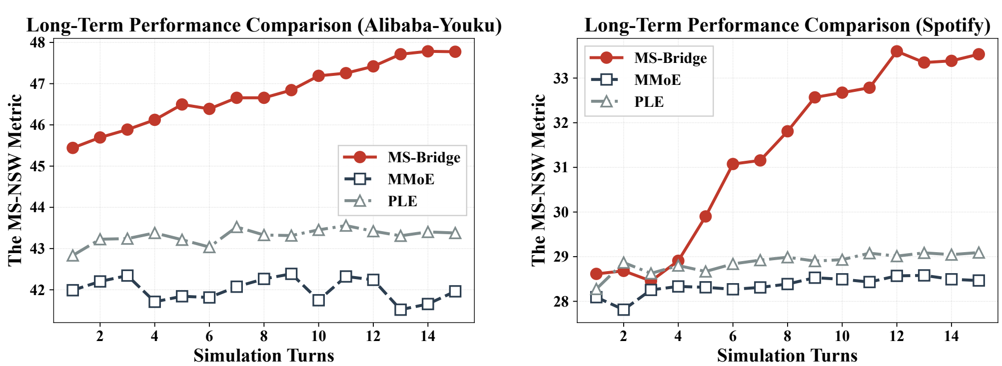


Figure 7 Longitudinal impact of MS-Bridge on the Alibaba-Youku dataset (left) and Spotify dataset (right). Y-axis shows the Nash Social Welfare of the ecosystem, while X-axis shows the round of iterations.

The results, visualized in Figure 7, illustrate the longitudinal superiority of our proposed MS-Bridge method over alternative baselines (specifically, MMOE and PLE), as it not only starts with a higher MS-NSW but also widens the performance gap over time, demonstrating a consistent upward trajectory over the 15 simulation turns. Meanwhile, both baseline models remain relatively stagnant throughout the simulation as they fail to properly capture cross-side network effects over the long term, whereas our MS-Bridge method creates a sustainable “Virtuous Cycle” of overall ecosystem welfare by successfully balancing user and artist utilities in the two-sided media market.

This monotonically widening MS-NSW gap across simulation rounds also serves as evidence of improved user retention: in each successive round, MS-Bridge users continue to receive high-utility recommendations that sustain engagement, whereas baseline users experience diminishing recommendation quality as the catalog of unexplored high-affinity content is exhausted. This pattern is consistent with the churn-reduction mechanism described in Section 6.1, where a 0.5 percentage point reduction in monthly churn was projected to retain 500,000 additional subscribers annually.

6.3.3. Simulation Study 2: User Welfare Before and After Adoption To quantify the specific impact of MS-Bridge on consumer behaviors, we also conduct an intervention study, where we simulate 16 recommendation iterations in both datasets, divided into two phases:

- **Pre-Adoption (first 8 iterations):** the system uses the MMoE baseline for recommendations.
- **Post-Adoption (last 8 iterations):** the system switches to MS-Bridge for recommendations.

Similar to the settings in Study 1, we select 80% of the data to optimize MMoE and MS-Bridge, and the remaining 20% for evaluation. In each iteration, we recommend the top candidate item with the highest utility value produced by MMoE/MS-Bridge, and then stochastically simulate the users’ response following the item response theory (Embretson and Reise 2013). We then update the objective values, re-train the model, and produce the next item recommendation accordingly. We also report performance as the average of 10 independent runs, where we calculate the *MS-NSW Utility Gain* (the utility delta between the Post-Adoption and Pre-Adoption) for both users and artists to determine the net impact of the algorithmic switch.

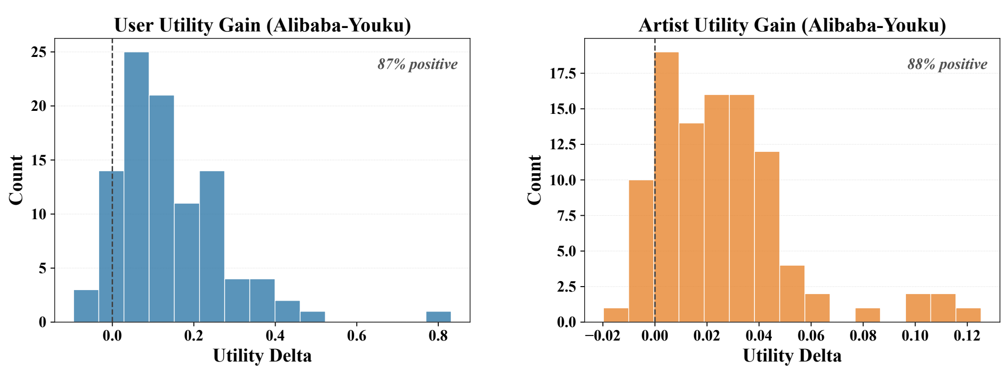


Figure 8 **Distribution of Utility Gain for Users/Artists after adopting MS-Bridge (Alibaba-Youku dataset).**

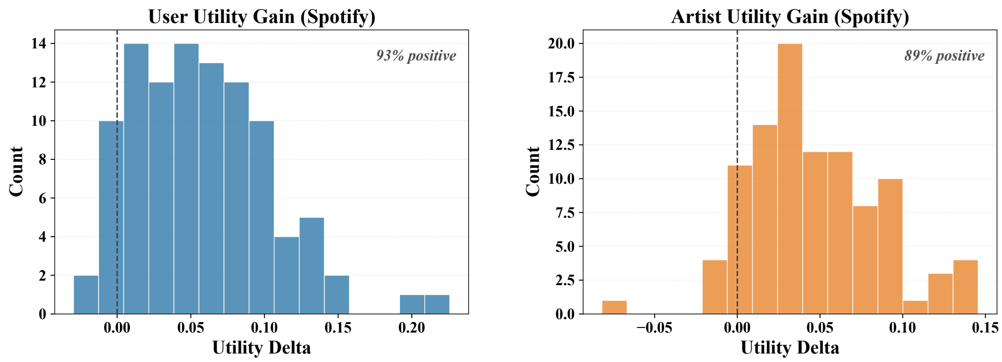


Figure 9 Distribution of Utility Gain for Users/Artists after adopting MS-Bridge (Spotify dataset).

The results, visualized in the histograms in Figures 8 and 9, demonstrate the welfare improvements across both sides of the market: the vast majority of users experience a net positive utility gain after the platform switches to the MS-Bridge model, and the distribution is heavily skewed to the right of the zero line; meanwhile, artists also heavily benefit from the intervention, as the distribution of artist utility gain is predominantly situated in the positive region. These results confirm that multi-stakeholder optimization via MS-Bridge is not a zero-sum game; instead, it effectively balances and enhances the welfare of both consumers and content creators simultaneously.

6.3.4. Weight Stability Analysis To further examine whether the adoption of MS-Bridge leads to structural changes in the relative importance of recommendation objectives, we conduct an additional analysis in this section, where we apply the ordinal regression technique described in Section 4.5 to learn the optimal objective weights separately for the pre-adoption phase (iterations 1-8) and the post-adoption phase (iterations 9-16). The learned weights, averaged over 10 independent runs, are reported in Table 9, where we observe that the optimized objective weights remain stable across both phases for all objectives in both datasets with no significant differences. This finding indicates that while MS-Bridge substantially improves the welfare of both stakeholders, it does so by producing better recommendations under the same implicit preference structure, rather than by altering the fundamental trade-off between objectives. In other words, the user preferences remain consistent across both algorithms for the same recommendation task, and the performance gains of MS-Bridge are attributable to its superior architectural design rather than a shift in the underlying objective priorities. We also note that some objectives receive near-zero weights in the simulation setting (Table 9) despite receiving positive weights in the archival data (Table 5). This discrepancy arises because the simulated data generated via the Method of Moments preserves marginal statistics but may amplify collinearity among objectives, and consequently, the ordinal regression concentrates weight on fewer objectives. Importantly, this does not affect the validity of the simulation study as the same weight structure is applied consistently across both the pre- and post-adoption phases, ensuring a fair comparison.

Table 9 Learned objective weights before and after MS-Bridge adoption.

	Pre-Adoption (MMoE)		Post-Adoption (MS-Bridge)		
Objective	Mean	Std	Mean	Std	p-value
(a) Spotify dataset					
New Fan	0.1967	0.0314	0.1958	0.0291	0.9494
Recurring	0.3507	0.0306	0.3822	0.0374	0.0666
%Listening	0.0000	0.0000	0.0000	0.0000	0.9498
Num Listening	0.4147	0.0561	0.3988	0.0570	0.5563
In Playlist	0.0379	0.0175	0.0223	0.0140	0.0667
Feedback Type	0.0000	0.0000	0.0000	0.0000	0.5788
(b) Alibaba dataset					
Video View	0.5902	0.0046	0.5904	0.0038	0.9126
Time Spent	0.1393	0.0157	0.1331	0.0150	0.4052
Play Rate	0.2303	0.0083	0.2348	0.0079	0.2513
Relevance	0.0403	0.0032	0.0417	0.0037	0.3903
Novelty	0.0000	0.0000	0.0000	0.0000	0.5924

Note. p-values from two-sample t-tests: no significant differences are observed.

Table 10 Sensitivity analysis of artist-side weight α on recommendations.

	Spotify dataset				Alibaba dataset			
α	U_{user}	U_{artist}	MS-NSW	Fan Rate	U_{user}	U_{artist}	MS-NSW	Fan Rate
0.1	0.4564	0.1007	0.2138	0.0300	0.2120	0.9093	0.4269	0.0250
0.3	0.4631	0.1008	0.2156	0.0240	0.2064	0.9100	0.4187	0.0250
0.5	0.4597	0.1032	0.2174	0.0210	0.2131	0.9094	0.4272	0.0240
0.7	0.4448	0.1007	0.2110	0.0260	0.2165	0.9096	0.4322	0.0200
0.9	0.4439	0.0992	0.2094	0.0230	0.2139	0.9103	0.4351	0.0180
MS-NSW	**0.4647**	**0.1041**	**0.2194**	**0.0330**	**0.2265**	**0.9113**	**0.4361**	**0.0260**

Note. α controls the weight in the linear utility $\alpha \cdot U_{\text{artist}} + (1-\alpha) \cdot U_{\text{user}}$. "Fan Rate" denotes the fraction of recommendations where the user is a new fan of the artist.

6.3.5. New Fan Weight Sensitivity Analysis Finally, to study the consequences of overweighting a single artist's objective beyond its optimized value, we conduct a sensitivity analysis in this section, where we systematically increase the artist's weight α from 0.1 to 0.9 in the linear scalarization score, while keeping the MS-Bridge model fixed. The results, reported in Table 10, reveal a clear trade-off: as alpha increases beyond the MS-NSW equilibrium, the average user utility degrades monotonically, while artist utility initially improves but eventually plateaus. In particular, at $\alpha = 0.9$, user utility drops by approximately 4.48% relative to the MS-NSW optimum, confirming that forcefully overweighting the New Fan objective leads the system to recommend less preferred (niche) content to non-affinity users, degrading the user experience without proportional artist-side gains. Crucially, the MS-NSW scoring function achieves the highest overall ecosystem welfare among all weight configurations, further validating that our proposed metric identifies the optimal balance between artist exposure and user satisfaction without requiring manual weight tuning.

6.4. Managerial Implications

Our proposed technical artifacts and the empirical findings in this paper offer significant strategic and operational insights for the design and governance of two-sided media platforms, as we provide a comprehensive toolkit for platform operators to maintain a healthy, sustainable ecosystem. We summarize the key managerial implications across the following dimensions.

6.4.1. Moving from User-Centric to Ecosystem-Centric Utility Metrics Historically, media platforms have heavily prioritized demand-side metrics, such as click-through rate and watch time, with the assumption that maximizing user engagement naturally benefits the supply side. However, we demonstrate that this deprives niche and emerging artists of exposure, potentially leading to supplier churn and a stagnant content catalog. Our proposed MS-NSW metric, meanwhile, offers managers a theoretically rigorous and empirically effective KPI that penalizes extreme disparities in outcomes, forcing the recommendation engine to find mutually beneficial matches. For platform operators, adopting the MS-NSW metric enables them to capture cross-side network effects and ensure that the platform cultivates a diverse supply side without indiscriminately sacrificing the user experience. As Table 1 demonstrates, platforms using additive user-only metrics achieve only 10% artist coverage and zero fan matches, while MS-NSW achieves 100% coverage and 990 fan matches with only a 3.1% reduction in user utility.

6.4.2. Overcoming the Trade-Off in Algorithmic Design A critical challenge in media platforms is the "negative transfer" problem, where optimizing for artist objectives (e.g., generating new fans for niche creators) often degrades user objectives (e.g., listening time). Our MS-Bridge method provides a tangible operational solution to this conflict through the design of "Latent Knowledge Bridge", which bidirectionally transfers relevant preference information between users and artists while isolating irrelevant noise. As a result, platforms achieve superior prediction accuracy across both sets of objectives simultaneously, effectively pushing the Pareto Frontier outward to obtain a "win-win" scenario for both stakeholders. Table 6 shows that the bridge architecture reduces New Fan RMSE by 5.03% over the best alternative, confirming that the negative transfer problem can be resolved without sacrificing user-side performance.

6.4.3. Automating Policy Decisions via Revealed Preferences Determining the weight of each objective in a recommendation is a delicate decision that often involves arbitrary setting or tedious A/B testing. Our MS-Bridge framework, meanwhile, alleviates this burden by employing ordinal regression and neural networks to dynamically learn the optimal objective weights directly from ground-truth rankings. From a managerial perspective, this data-driven approach operationalizes the economic theory of revealed preferences. It allows the platform to automatically infer the natural equilibrium of the market based on how users actually consume ranked content without the need for manual parameter tuning. As shown in Table 5, the ordinal regression approach outperforms the best manually-tuned weight configuration by 1.31% in Precision@10 and 9.53% in MS-NSW, eliminating the need for costly manual A/B tests.

6.4.4. Cultivating a Long-Term "Virtuous Cycle" of Platform Growth The results of our longitudinal simulation study highlight a critical implication for platform sustainability: while baseline models tend to plateau user satisfaction, MS-Bridge gathers richer, more diverse feedback data over time by deliberately injecting high-quality niche content to users who possess latent affinities for those items, which in turn drives higher long-term user satisfaction and total ecosystem welfare, leading to a "Virtuous Cycle". Platform operators may leverage this insight to justify investments in multi-stakeholder optimization: while balancing artist welfare might sacrifice certain performance in the short term, it is actually a prerequisite for preventing catalog exhaustion, reducing user churn, and driving sustainable long-term revenues. Figure 7 shows that the MS-NSW gap between MS-Bridge and the best baseline grows from approximately 1.5 to 4.0 over 15 rounds, confirming that the long-term benefits compound rather than diminish.

6.4.5. Generalizability Beyond Media Platforms While our evaluations center on media platforms, the core economic tensions addressed by our framework are prevalent across the digital economy. Managers of e-commerce marketplaces (balancing buyer conversion with third-party seller exposure), gig-economy applications (balancing rider wait times with driver earnings), and social media networks (balancing user engagement with creator monetization) may also adopt our proposed MS-Bridge and MS-NSW paradigms to balance the welfare of both the supply and demand sides, avoid market failures and foster more equitable, profitable digital ecosystems.

7. Conclusions

In this paper, we examine the economic realities of two-sided media markets, and we demonstrate through both motivational examples and empirical analysis that it is beneficial to consider the objectives of both stakeholders in recommendations, rather than focusing only on the user objectives, as was done in the prior literature. To that end, we develop a novel MS-NSW utility metric to capture the economic realities of the two-sided media markets and to balance the trade-off between both stakeholders.

Since optimizing the MS-NSW metric is an NP-Hard, non-trivial task in practice, we subsequently propose a novel neural network architecture of "Latent Knowledge Bridge", where we bidirectionally transfer the preference information between two stakeholders to strike an effective balance between them. The optimized location of the bridge, meanwhile, is determined as the middle layer based on both anecdotal evidence and a Shapley-value-analysis that we conduct in the paper. This enables us to avoid the *Negative Transfer* problem commonly encountered in the prior literature. In addition, we present the ordinal regression method and a neural network method to select the most suitable objective weights, which produce significantly better performance over the existing Bayesian and weighted average methods.

Using four large-scale industrial datasets from media streaming platforms, we show that our proposed model yields significant improvements across various baselines and evaluation metrics, and that it is effective at properly balancing the objectives of multiple stakeholders, which is crucial for providing recommendations in two-sided markets in a more equitable manner. We further conduct the economic value analysis and an interpretable case study to show that these performance improvements are economically meaningful and potentially lead to significant economic impact for the two-sided media platform that adopts our proposed framework. Finally, we conduct a simulation experiment to study the impact on consumer behaviors, where we observe that the majority of consumers and artists will be better off by adopting our model, and that the benefits become greater in the long term when they consistently interact with the model.

We also acknowledge certain limitations of our evaluation framework. Our offline evaluations rely on logged interaction data, where the observed outcomes are shaped by the platforms' existing recommendation policies, meaning that users' responses to items that were never recommended remain unobserved. While our simulation study partially addresses this by modeling counterfactual responses and iterative feedback loops, it relies on distributional assumptions calibrated from the same archival data. Furthermore, although online controlled experiments (A/B tests) represent the gold standard for establishing the causal impact of recommendation algorithms on stakeholder welfare (Nandy et al. 2021), they were not feasible in our study due to the strict internal policies at Spotify and Alibaba governing deployment and public reporting. Consequently, while our results collectively provide strong evidence for the superiority of MS-Bridge, the economic projections in Section 6.1 should be interpreted as indicative estimates under reasonable assumptions, as the realized impact will depend on platform-specific factors such as the existing recommendation policy, user population, and monetization model.

These limitations point to several promising directions for future work. First, performing online experiments at a partnering media platform would allow us to directly measure the causal effects of MS-Bridge on stakeholder welfare and confirm that incorporating artist objectives as part of the recommendation process is valuable to both stakeholders. Second, while our analysis has largely focused on the two-sided markets in the media industry, future research may conduct similar experiments in other types of two-sided marketplaces, such as food delivery services, ride-hailing, short-term rentals, tutoring services, e-commerce, digital advertising, and many others. Due to the heterogeneity of the market structure, the differences in users' preferences, and the competitive landscape of these two-sided markets, it would be interesting to explore how to adjust our recommendation strategies to accommodate these unique properties accordingly.

References

Abdollahpouri H, Adomavicius G, Burke R, Guy I, Jannach D, Kamishima T, Krasnodebski J, Pizzato L (2020) Multistakeholder recommendation: Survey and research directions. *User Modeling and User-Adapted Interaction* 30(1):127–158.

Abels A, Roijers D, Lenaerts T, Nowé A, Steckelmacher D (2019) Dynamic weights in multi-objective deep reinforcement learning. *International Conference on Machine Learning*, 11–20 (PMLR).

Adomavicius G, Manouselis N, Kwon Y (2011) Multi-criteria recommender systems. *Recommender systems handbook*, 769–803 (Springer).

Ancona M, Oztireli C, Gross M (2019) Explaining deep neural networks with a polynomial time algorithm for shapley value approximation. *International Conference on Machine Learning*, 272–281 (PMLR).

Anderson A, Maystre L, Anderson I, Mehrotra R, Lalmas M (2020) Algorithmic effects on the diversity of consumption on spotify. *Proceedings of the web conference 2020*, 2155–2165.

Anderson SP, Gabszewicz JJ (2006) The media and advertising: a tale of two-sided markets. *Handbook of the Economics of Art and Culture* 1:567–614.

Barman S, Krishnamurthy SK, Vaish R (2018) Finding fair and efficient allocations. *Proceedings of the 2018 ACM Conference on Economics and Computation*, 557–574.

Bauschke HH, Bolte J, Teboulle M (2017) A descent lemma beyond lipschitz gradient continuity: first-order methods revisited and applications. *Mathematics of Operations Research* 42(2):330–348.

Bauschke HH, Combettes PL (2017) Convex cones and generalized interiors. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 111–132 (Springer).

Bello I, Pham H, Le QV, Norouzi M, Bengio S (2016) Neural combinatorial optimization with reinforcement learning. *arXiv preprint arXiv:1611.09940* .

Bengio Y, Lodi A, Prouvost A (2021) Machine learning for combinatorial optimization: a methodological tour d'horizon. *European Journal of Operational Research* 290(2):405–421.

Burke R, Abdollahpouri H (2017) Patterns of multistakeholder recommendation. *arXiv preprint arXiv:1707.09258* .

Cheng HT, Koc L, Harmsen J, Shaked T, Chandra T, Aradhye H, Anderson G, Corrado G, Chai W, Ispir M, et al. (2016) Wide & deep learning for recommender systems. *Proceedings of the 1st workshop on deep learning for recommender systems*, 7–10.

Cohen MC, Zhang R (2022) Competition and coopetition for two-sided platforms. *Production and Operations Management* 31(5):1997–2014.

Cole R, Gkatzelis V (2015) Approximating the nash social welfare with indivisible items. *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, 371–380.

Covington P, Adams J, Sargin E (2016) Deep neural networks for youtube recommendations. *Proceedings of the 10th ACM conference on recommender systems*, 191–198.

Dai X, Jordan MI (2021) Learning strategies in decentralized matching markets under uncertain preferences. *The Journal of Machine Learning Research* 22(1):11806–11855.

Dantzig GB (2002) Linear programming. *Operations research* 50(1):42–47.

Eide S, Zhou N (2018) Deep neural network marketplace recommenders in online experiments. *Proceedings of the 12th ACM Conference on Recommender Systems*, 387–391.

Embretson SE, Reise SP (2013) *Item response theory for psychologists* (Psychology Press).

Fazelnia G, Gupta S, Keum C, Koh M, Anderson I, Lalmas M (2024) Generalized user representations for transfer learning. *arXiv preprint arXiv:2403.00584* .

Feurer M, Springenberg J, Hutter F (2015) Initializing bayesian hyperparameter optimization via meta-learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29.

Gao C, He X, Gan D, Chen X, Feng F, Li Y, Chua TS, Yao L, Song Y, Jin D (2019) Learning to recommend with multiple cascading behaviors. *IEEE Transactions on Knowledge and Data Engineering* .

Geng B, Li L, Jiao L, Gong M, Cai Q, Wu Y (2015) Nnia-rs: A multi-objective optimization based recommender system. *Physica A: Statistical Mechanics and its Applications* 424:383–397.

Gomez-Uribe CA, Hunt N (2015) The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)* 6(4):1–19.

Gong G, Li K, Zha Q (2023) A maximum fairness consensus model with limited cost in group decision making. *Computers & Industrial Engineering* 175:108891.

Gunawardana A, Shani G, Yogev S (2022) Evaluating recommender systems. *Recommender Systems Handbook: Third Edition*, 547–601 (Springer US).

Hansen C, Hansen C, Maystre L, Mehrotra R, Brost B, Tomasi F, Lalmas M (2020) Contextual and sequential user embeddings for large-scale music recommendation. *Fourteenth ACM Conference on Recommender Systems*, 53–62.

Hansen C, Mehrotra R, Hansen C, Brost B, Maystre L, Lalmas M (2021) Shifting consumption towards diverse content on music streaming platforms. *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, 238–246.

He X, Liao L, Zhang H, Nie L, Hu X, Chua TS (2017) Neural collaborative filtering. *Proceedings of the 26th international conference on world wide web*, 173–182.

Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. *Neural networks* 2(5):359–366.

Hu J, Li P (2018) Collaborative filtering via additive ordinal regression. *Proceedings of the eleventh ACM international conference on web search and data mining*, 243–251.

Jannach D (2022) Multi-objective recommendation: Overview and challenges. *Proceedings of the 2nd Workshop on Multi-Objective Recommender Systems co-located with 16th ACM Conference on Recommender Systems (RecSys 2022)*, volume 3268.

Jin L, Yuan D, Zhang H (2017) Music recommendation based on embedding model with user preference and context. *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, 688–692 (IEEE).

Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. *Journal of Global optimization* 13(4):455–492.

Kaneko M, Nakamura K (1979) The nash social welfare function. *Econometrica: Journal of the Econometric Society* 423–435.

Keeney RL, Raiffa H (1993) *Decisions with multiple objectives: preferences and value trade-offs* (Cambridge university press).

Khalil E, Dai H, Zhang Y, Dilkina B, Song L (2017) Learning combinatorial optimization algorithms over graphs. *Advances in neural information processing systems* 30.

Koren Y (2010) Factor in the neighbors: Scalable and accurate collaborative filtering. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 4(1):1–24.

Lee E (2017) Apx-hardness of maximizing nash social welfare with indivisible items. *Information Processing Letters* 122:17–20.

Li J, Ren P, Chen Z, Ren Z, Lian T, Ma J (2017) Neural attentive session-based recommendation. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 1419–1428.

Li P, Que M, Jiang Z, Hu Y, Tuzhilin A (2020) Purs: Personalized unexpected recommender system for improving user satisfaction. *Proceedings of the 14th ACM Conference on Recommender Systems*, 279–288.

Li P, Tuzhilin A (2020) Ddtcdr: Deep dual transfer cross domain recommendation. *Proceedings of the 13th International Conference on Web Search and Data Mining*, 331–339.

Li P, Tuzhilin A (2024a) Deep pareto reinforcement learning for multi-objective recommender systems. *arXiv preprint arXiv:2407.03580* .

Li P, Tuzhilin A (2024b) When variety seeking meets unexpectedness: Incorporating variety-seeking behaviors into design of unexpected recommender systems. *Information Systems Research* 35(3):1257–1273.

Lin S (2020) Two-sided price discrimination by media platforms. *Marketing Science* 39(2):317–338.

Lin X, Chen H, Pei C, Sun F, Xiao X, Sun H, Zhang Y, Ou W, Jiang P (2019) A pareto-efficient algorithm for multiple objective optimization in e-commerce recommendation. *Proceedings of the 13th ACM Conference on Recommender Systems*, 20–28.

Long M, Wang J, Ding G, Cheng W, Zhang X, Wang W (2012) Dual transfer learning. *Proceedings of the 2012 SIAM International Conference on Data Mining*, 540–551 (SIAM).

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. *Advances in neural information processing systems* 30.

Ma J, Zhao Z, Yi X, Chen J, Hong L, Chi EH (2018) Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1930–1939.

Malgonde O, Zhang H, Padmanabhan B, Limayem M (2020) Taming complexity in search matching: Two-sided recommender systems on digital platforms. *MIS Quarterly* 44(1).

McCullagh P (1980) Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)* 42(2):109–127.

Mehrotra R, Carterette B (2019) Recommendations in a marketplace. *Proceedings of the 13th ACM Conference on Recommender Systems*, 580–581.

Mehrotra R, Lalmas M, Kenney D, Lim-Meng T, Hashemian G (2019) Jointly leveraging intent and interaction signals to predict user satisfaction with slate recommendations. *The World Wide Web Conference*, 1256–1267.

Mehrotra R, McInerney J, Bouchard H, Lalmas M, Diaz F (2018) Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems. *Proceedings of the 27th acm international conference on information and knowledge management*, 2243–2251.

Mehrotra R, Xue N, Lalmas M (2020) Bandit based optimization of multiple objectives on a music streaming platform. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3224–3233.

Milojkovic N, Antognini D, Bergamin G, Faltings B, Musat C (2019) Multi-gradient descent for multi-objective recommender systems. *arXiv preprint arXiv:2001.00846* .

Misra I, Shrivastava A, Gupta A, Hebert M (2016) Cross-stitch networks for multi-task learning. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3994–4003.

Mladenov M, Creager E, Ben-Porat O, Swersky K, Zemel R, Boutilier C (2020) Optimizing long-term social welfare in recommender systems: A constrained matching approach. *International Conference on Machine Learning*, 6987–6998 (PMLR).

Nandy P, Venugopalan D, Lo C, Chatterjee S (2021) A/b testing for recommender systems in a two-sided marketplace. *Advances in Neural Information Processing Systems* 34:6466–6477.

Nguyen P, Dines J, Krasnodebski J (2017) A multi-objective learning to re-rank approach to optimize online marketplaces for multiple stakeholders. *arXiv preprint arXiv:1708.00651* .

Pan SJ, Yang Q (2009) A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22(10):1345–1359.

Parvez MR, Chang KW (2021) Evaluating the values of sources in transfer learning. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 5084–5116.

Qin Z, Cheng Y, Zhao Z, Chen Z, Metzler D, Qin J (2020) Multitask mixture of sequential experts for user activity streams. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3083–3091.

Ramezani S, Endriss U (2009) Nash social welfare in multiagent resource allocation. *International Workshop on Agent-Mediated Electronic Commerce*, 117–131 (Springer).

Ribeiro MT, Ziviani N, Moura ESD, Hata I, Lacerda A, Veloso A (2014) Multiobjective pareto-efficient approaches for recommender systems. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5(4):1–20.

Rodriguez M, Posse C, Zhang E (2012) Multiple objective optimization in recommender systems. *Proceedings of the sixth ACM conference on Recommender systems*, 11–18.

Ruder S (2017) An overview of multi-task learning in deep neural networks. *arXiv:1706.05098* .

Rysman M (2009) The economics of two-sided markets. *Journal of economic perspectives* 23(3):125–143.

Samuelson PA (2024) A note on the pure theory of consumer's behaviour. *The Foundations of Price Theory Vol 4*, 101–116 (Routledge).

Schedl M (2019) Deep learning in music recommendation systems. *Frontiers in Applied Mathematics and Statistics* 5:44.

Sener O, Koltun V (2018) Multi-task learning as multi-objective optimization. *Advances in Neural Information Processing Systems*, 527–538.

Shi P (2023) Optimal matchmaking strategy in two-sided marketplaces. *Management Science* 69(3):1323–1340.

Snoek J, Larochelle H, Adams RP (2012) Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems* 25.

Sürer Ö, Burke R, Malthouse EC (2018) Multistakeholder recommendation with provider constraints. *Proceedings of the 12th ACM Conference on Recommender Systems*, 54–62.

Tang H, Liu J, Zhao M, Gong X (2020) Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations. *Fourteenth ACM Conference on Recommender Systems*, 269–278.

Valizadegan H, Jin R, Zhang R, Mao J (2009) Learning to rank by optimizing ndcg measure. *Advances in neural information processing systems* 22.

Wallenius J, Dyer JS, Fishburn PC, Steuer RE, Zionts S, Deb K (2008) Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management science* 54(7):1336–1349.

Wang Z, Dai Z, Póczos B, Carbonell J (2019) Characterizing and avoiding negative transfer. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11293–11302.

Yang Y, Hospedales T (2016) Deep multi-task representation learning: A tensor factorisation approach. *arXiv preprint arXiv:1605.06391* .

Zhang B, Kreitz G, Isaksson M, Ubillos J, Urdaneta G, Pouwelse JA, Epema D (2013) Understanding user behavior in spotify. *2013 Proceedings IEEE INFOCOM*, 220–224 (IEEE).

Zhang H, Li X, Bing L (2023) Video-llama: An instruction-tuned audio-visual language model for video understanding. *Proceedings of the 2023 conference on empirical methods in natural language processing: system demonstrations*, 543–553.

Zhang S, Yao L, Sun A, Tay Y (2019) Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)* 52(1):1–38.

Zhang W, Deng L, Zhang L, Wu D (2022) A survey on negative transfer. *IEEE/CAA Journal of Automatica Sinica* 10(2):305–329.

Zhao J, Du B, Sun L, Zhuang F, Lv W, Xiong H (2019a) Multiple relational attention network for multi-task learning. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1123–1131.

Zhao Z, Hong L, Wei L, Chen J, Nath A, Andrews S, Kumthekar A, Sathiamoorthy M, Yi X, Chi E (2019b) Recommending what video to watch next: a multitask ranking system. *Proceedings of the 13th ACM conference on recommender systems*, 43–51.

Zheng Y, Ghane N, Sabouri M (2019) Personalized educational learning with multi-stakeholder optimizations. *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, 283–289.

Zhu H, Zhou Y, Chen H, Yu J, Ma Z, Gu R, Luo Y, Tan W, Chen X (2025) Muq: Self-supervised music representation learning with mel residual vector quantization. *IEEE Transactions on Audio, Speech and Language Processing* .

Appendix. MS-Bridge: A Deep Multi-Stakeholder Multi-Objective Recommendation Model for Two-Sided Media Platforms

A. Key Statistics of Four Datasets in Our Evaluations

In this section, we present the key statistics of the four media recommendation datasets used in our offline evaluations. We present the summary statistics in Table 11, and the distribution for each objective associated with each dataset in Tables 12, 13, 14, and 15 respectively.

Table 11 Descriptive statistics for our three datasets.

Dataset	# of Users	# of Songs/Videos	# of Interactions	Sparsity
Spotify (Next Song Recommendation)	294,469	81,948	970,013	0.004%
Spotify (Session-Based Recommendation)	82,696	417,932	8,892,650	0.026%
MLHD (Last.fm)	1,052	407,681	3,393,744	0.791%
Alibaba-Youku	51,419	229	99,999	0.849%

Table 12 Statistics of user and artist objectives in the Spotify (Next Song Recommendation) dataset.

Stakeholder	Objective	Type	Mean	Std	Median	25% percentile	75% percentile
User objectives	Percentage of listening	Numeric	0.731	0.413	1	0.32	1
	Number of times listening	Numeric	0.528	0.386	0.499	0.129	1
	Feedback type	Numeric	0.771	0.420	1	1	1
	In playlist	Binary	0.304	0.460	0	0	1
Artist objectives	New fan	Numeric	0.556	0.336	0.638	0.247	0.861
	Recurring fan	Numeric	0.65	0.309	0.756	0.448	0.905

Table 13 Statistics of user and artist objectives in the Spotify (Session-Based Recommendation) dataset.

Stakeholder	Objective	Type	Mean	Std	Median	25% percentile	75% percentile
User objectives	Percentage of listening	Numeric	0.727	0.414	1	0.303	1
	Number of times listening	Numeric	0.511	0.393	0.447	0.107	1
	Feedback type	Numeric	0.757	0.429	1	1	1
	In playlist	Binary	0.330	0.470	0	0	1
Artist objectives	New fan	Numeric	0.570	0.340	0.677	0.255	0.878
	Recurring fan	Numeric	0.660	0.309	0.769	0.477	0.913

In addition, we have also created a series of objective correlation plots for each dataset that we use in the offline evaluations, where we illustrate the distribution of each objective value within each dataset, as well as demonstrate its relationship with every other objective through pairwise scatterplots. We can observe from Figure 10 (as well as

Table 14 Statistics of user and artist objectives in the MLHD dataset.

Stakeholder	Objective	Type	Mean	Std	Median	25% percentile	75% percentile
User objectives	Percentage of listening	Numeric	0.958	0.117	0.994	1	1
	IsFinish	Binary	0.847	0.360	1	1	1
	IsSkip	Binary	0.02	0.049	0	0	0
Artist objectives	New fan	Numeric	0.950	0.217	1	1	1
	Recurring fan	Numeric	0.698	0.229	0.475	0.704	0.932

Table 15 Statistics of user and artist objectives in the Alibaba-Youku dataset.

Stakeholder	Objective	Type	Mean	Std	Median	25% percentile	75% percentile
User objectives	Video View	Binary	0.268	0.443	0	0	1
	Time Spent	Numeric	11.662	32.305	0	0	3.16
	Play Rate	Numeric	0.194	0.370	0	0	0.056
Artist objectives	Relevance	Numeric	0.312	0.735	0	0	1
	Novelty	Numeric	0.538	0.382	0.375	0.283	0.743

Figure 6 that we present in the main paper) that the relationships between user objectives and artist objectives can be complicated and even complicated in many cases.

B. Hardware Conditions and Hyper-Parameter Tuning

Our MS-Bridge model, as well as baseline models, were all trained on an academic HPC server equipped with RTX8000 NVIDIA GPU, where we requested 16 CPUs and 64 GB RAM. On average, it takes 10 minutes for both Spotify datasets, 3 minutes for the MLHD dataset, and 5 minutes for the Alibaba dataset to train our proposed model. We implement our proposed model as well as baseline models based on the TensorFlow framework in Python.

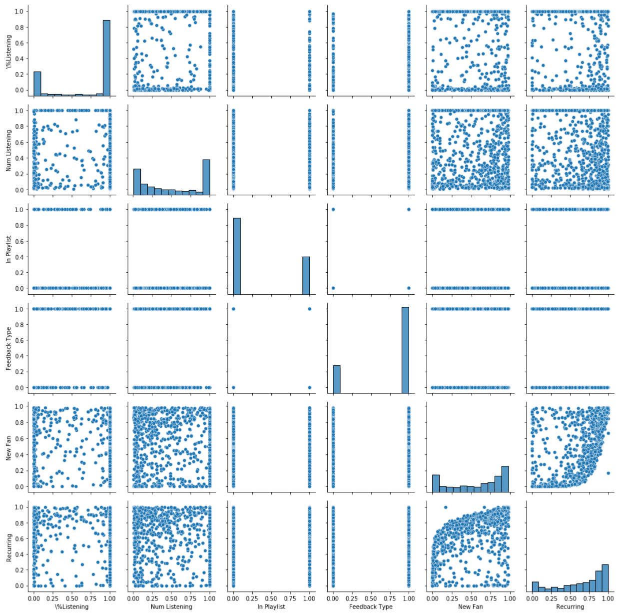
For the implementation of deep learning-based baseline methods (Wide&Deep, NeuMF, MMoE, MoSE, NMTR, and PLE), we test various neural network architectures and hyperparameters using a grid search method. Specifically, for the Wide&Deep and NeuMF models, we report the best results obtained when constructing four fully-connected hidden layers with [128,80,40,1] neurons in the corresponding layers (e.g., 128 neurons in the first layer). For the multi-objective recommendation models of MMoE, MoSE, NMTR, and PLE, we construct three fully-connected hidden layers with [16,8,6] neurons in the corresponding layers. We separately build an additional fully-connected layer to combine the multi-objective predictions from all six expert models (one for each objective). For the implementation of non-deep-learning-based baseline methods (e.g., LR and Multi-Criteria CF), the hyperparameters are selected using the Bayesian hyperparameter optimization technique to generate the optimal recommendation performance.

We fine-tune all the hyper-parameters on the validation set, where we tested a batch size of $\{128, 256, 512, 1024\}$ and a learning rate of $\{0.001, 0.005, 0.0001, 0.0005\}$. The size of the last hidden layer in each tower is termed as predictive factors (He et al. 2017), and we evaluated the factors in 8, 16, 32, 64. We report the best results obtained using eight predictive factors for the artist tower and 16 predictive factors for the user tower. We used five hidden layers for each tower.

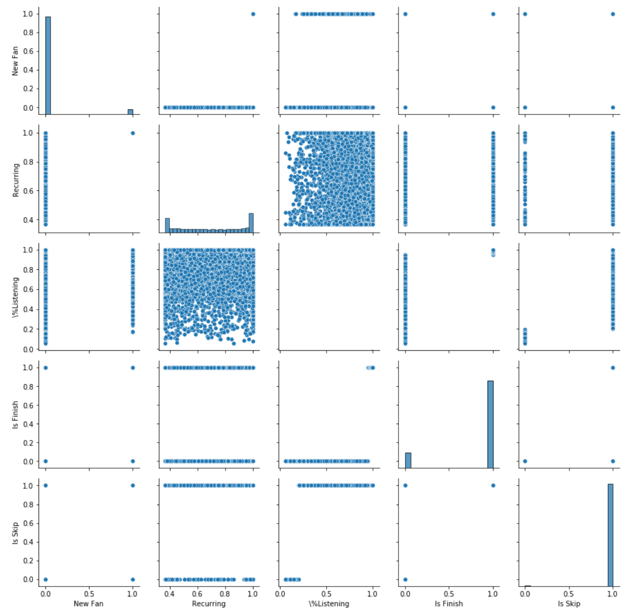
C. Additional Prediction Performance in terms of AUC and F1 Score

Besides the performance results based on the RMSE and MAE metrics that we present in the main paper, we also evaluate the prediction performance using the AUC and F1 score in this section, and we present the results in Tables 16, 17 and 18 for the three datasets. For the Spotify dataset, the results show that our joint prediction model with the bridge architecture achieves a better prediction performance for the artist objectives. Specifically, the AUC values of the New Fan and Recurring objectives are significantly higher than the AUC of the user objectives. More precisely, the joint prediction model (+bridge) reaches the highest AUC performance values of 0.744 and 0.683 for the New Fan and Recurring objectives, which is an improvement of 36.51% and 28.87% respectively, compared to the best baseline.

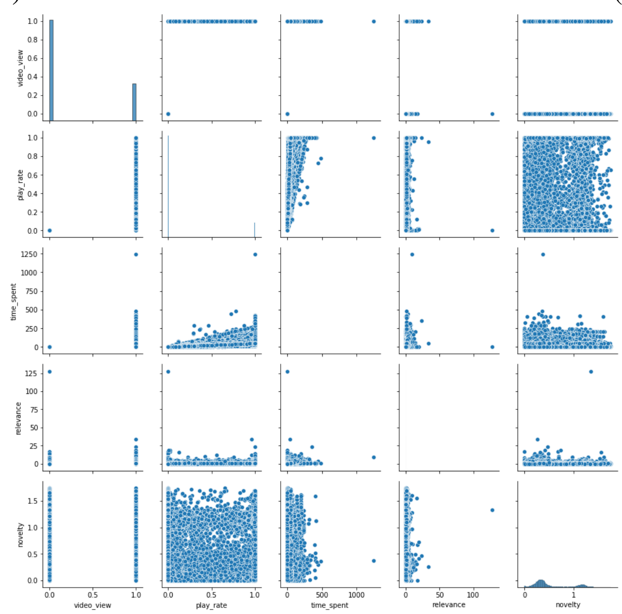
We can also see from Table 16 that the improvements in the AUC and the F1 score are different. The AUC improvement is relatively high, ranging between 4.21% and 36.51%, while the F1 score improvement is more modest, ranging



(a)



(b)



(c)

Figure 10 Correlation Plots Between Different Objectives in the (a) Spotify (b) MLHD (c) Alibaba datasets.

between 0.39% and 9.90%. This difference could be explained by the fact that the two metrics are computed in different ways: (i) the F1 score is computed based on the Precision@10 and Recall@10 that focus primarily on the prediction performance of each objective, while considering the top-10 recommended items, (ii) the AUC measures the overall prediction performance among all items and hence also accounts for long-tailed, niche items. Thus, our results strongly suggest that our joint prediction model generates a satisfactory prediction performance even when considering long-tailed items.

Table 16 Comparison of AUC and F1 Score: Spotify dataset.

Model Type	Model	Artist Objectives				User Objectives							
		New Fan		**Recurring**		**%Listening**		**Num Listening**		**In Playlist**		**Feedback Type**	
		AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1
Single-Objective Models	LR	0.501	0.682	0.501	0.786	0.498	0.447	0.499	0.623	0.503	0.368	0.501	0.485
	NeuMF	0.509	0.858	0.518	0.929	0.535	0.739	0.524	0.759	0.523	0.534	0.543	0.800
	Wide&Deep	0.518	0.863	0.520	0.930	0.537	0.745	0.528	0.765	0.534	0.541	0.549	0.811
	Single-Objective Towers	0.512	0.860	0.519	0.930	0.552	0.761	0.532	0.766	0.530	0.538	0.562	0.823
Multi-Objective Models	MCCF	0.531	0.850	0.504	0.880	0.515	0.697	0.511	0.759	0.515	0.520	0.508	0.730
	MMoE	0.514	0.866	0.517	0.930	0.531	0.737	0.529	0.762	0.525	0.535	0.538	0.790
	MoSE	0.537	0.869	0.528	0.932	0.547	0.766	0.541	0.763	0.544	0.539	0.551	0.796
	NMTR	0.514	0.861	0.510	0.925	0.529	0.759	0.526	0.761	0.525	0.535	0.540	0.787
	PLE	0.545	0.868	0.531	0.932	0.550	0.768	0.546	0.763	0.540	0.539	0.553	0.799
	MO-LinUCB	0.543	0.868	0.528	0.931	0.549	0.767	0.545	0.761	0.539	0.538	0.551	0.798
	MORL	0.541	0.867	0.527	0.929	0.547	0.765	0.545	0.760	0.537	0.537	0.549	0.796
MS-Bridge (Our Proposed Model)	Wide&Deep (No Bridge)	0.710	0.869	0.649	0.933	0.601	0.815	0.540	0.763	0.616	0.539	0.597	0.855
	NCF (No Bridge)	0.719	0.870	0.655	0.935	0.610	0.826	0.541	0.765	0.624	0.540	0.609	0.877
	Wide&Deep	**0.728***	**0.873***	**0.664***	**0.936***	**0.625***	**0.830***	**0.552***	**0.767***	**0.642***	**0.544***	**0.619***	**0.881***
	NCF	**0.744***	**0.877***	**0.683***	**0.939***	**0.636***	**0.844***	**0.569***	**0.769***	**0.664***	**0.549***	**0.625***	**0.888***
Improvement	▲%	36.51%	0.92%	28.87%	0.75%	15.22%	9.90%	4.21%	0.39%	22.06%	1.48%	11.21%	7.90%

Note. * indicates significance ($p \leq 0.05$) of paired t-tests between baselines and MS-Bridge, and ▲% indicates improvement over best baseline (underlined).

Table 17 Comparison of AUC and F1 Score: MLHD dataset.

Model Type	Model	Artist Objectives				User Objectives					
		New Fan		**Recurring**		**%Listening**		**IsFinish**		**IsSkip**	
		AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1
Single-Objective Models	LR	0.719	0.682	0.605	0.786	0.807	0.447	0.821	0.623	0.793	0.368
	NeuMF	0.577	0.858	0.572	0.929	0.725	0.739	0.719	0.759	0.723	0.534
	Wide&Deep	0.772	0.863	0.613	0.930	0.803	0.745	0.827	0.765	0.793	0.541
	Single-Objective Towers	0.769	0.865	0.620	0.931	0.810	0.738	0.827	0.762	0.794	0.535
Multi-Objective Models	MCCF	0.769	0.850	0.598	0.880	0.798	0.697	0.798	0.759	0.756	0.520
	MMoE	0.769	0.866	0.621	0.930	0.811	0.737	0.830	0.762	0.798	0.535
	MoSE	0.770	0.868	0.623	0.932	0.814	0.744	0.831	0.762	0.797	0.538
	NMTR	0.769	0.866	0.620	0.930	0.811	0.738	0.828	0.760	0.795	0.536
	PLE	0.770	0.870	0.623	0.932	0.815	0.746	0.831	0.762	0.797	0.539
	MO-LinUCB	0.768	0.869	0.621	0.931	0.814	0.745	0.830	0.761	0.796	0.538
	MORL	0.766	0.867	0.620	0.929	0.812	0.743	0.828	0.759	0.793	0.536
MS-Bridge (Our Proposed Model)	Wide&Deep (No Bridge)	0.770	0.868	0.623	0.933	0.812	0.810	0.831	0.763	0.798	0.539
	NCF (No Bridge)	0.772	0.870	0.624	0.935	0.814	0.826	0.833	0.766	0.800	0.540
	Wide&Deep	**0.775***	**0.874***	**0.627***	**0.936***	**0.820***	**0.823***	**0.834***	**0.767***	**0.801***	**0.543***
	NCF	**0.778***	**0.877***	**0.630***	**0.939***	**0.823***	**0.844***	**0.836***	**0.769***	**0.803***	**0.549***
Improvement	▲%	0.78%	0.80%	1.12%	0.75%	0.98%	13.14%	0.60%	0.52%	0.63%	1.48%

Note. * indicates significance ($p \leq 0.05$) of paired t-tests between baselines and MS-Bridge, and ▲% indicates improvement over best baseline (underlined).

For the MLHD dataset, we can see in Table 17 that the joint prediction model (+bridge) reaches the highest AUC value (0.836) for the IsFinish user objective and the highest F1 value (0.939) for the Recurring artist objective. This confirms the fact that both user and artist objectives need to be considered in the model in order to obtain a satisfactory prediction performance. In addition, we can see deep-learning models (Wide&Deep, NeuMF, MMoE, and our proposed joint model) outperform non-neural approaches (LR and MCCF). This can be explained by the fact that deep learning can better learn and analyze large amounts of data and learn non-linear correlations among the different objectives (Zhang et al. 2019, Schedl 2019).

Finally, for the Alibaba dataset, we can see in Table 18 that our proposed model consistently and significantly outperforms all baseline methods across all considered objectives and evaluation metrics, further demonstrating the benefits of our proposed model.

Table 18 Comparison of AUC and F1 Score: Alibaba dataset.

Model Type	Model	User Objectives						Artist Objectives			
		Video View		**Time Spent**		**Play Rate**		**Relevance**		**Novelty**	
		AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1
Single-Objective Models	LR	0.699	0.800	0.870	0.802	0.725	0.741	0.669	0.738	0.620	0.567
	NeuMF	0.713	0.813	0.887	0.823	0.743	0.755	0.686	0.753	0.632	0.578
	Wide&Deep	0.713	0.813	0.886	0.821	0.743	0.755	0.688	0.755	0.632	0.578
	Single-Objective Towers	0.711	0.810	0.884	0.818	0.741	0.755	0.688	0.753	0.633	0.580
Multi-Objective Models	MCCF	0.706	0.808	0.884	0.818	0.740	0.753	0.688	0.752	0.633	0.580
	MMoE	0.722	0.825	0.895	0.832	0.746	0.765	0.698	0.763	0.642	0.592
	MoSE	0.723	0.825	0.896	0.833	0.746	0.765	0.700	0.765	0.642	0.592
	NMTR	0.718	0.822	0.891	0.830	0.743	0.762	0.695	0.761	0.640	0.588
	PLE	<u>0.725</u>	<u>0.826</u>	<u>0.898</u>	<u>0.835</u>	<u>0.747</u>	<u>0.767</u>	<u>0.701</u>	<u>0.766</u>	<u>0.643</u>	<u>0.594</u>
	MO-LinUCB	0.713	0.820	0.891	0.826	0.743	0.760	0.693	0.758	0.637	0.585
	MORL	0.713	0.822	0.891	0.825	0.743	0.760	0.692	0.758	0.638	0.585
MS-Bridge (Our Proposed Model)	Wide&Deep (No Bridge)	0.730	0.825	0.896	0.832	0.746	0.767	0.702	0.763	0.641	0.592
	NCF (No Bridge)	0.736	0.833	0.903	0.840	0.751	0.774	0.709	0.771	0.650	0.599
	Wide&Deep	**0.737***	**0.835***	**0.905***	**0.843***	**0.758***	**0.776***	**0.726***	**0.780***	**0.652***	**0.601***
	NCF	**0.741***	**0.840***	**0.912***	**0.851***	**0.769***	**0.788***	**0.736***	**0.794***	**0.660***	**0.607***
Improvement	▲%	2.16%	1.67%	1.54%	1.88%	2.86%	2.66%	4.76%	3.53%	2.58%	2.14%

Note. * indicates significance ($p \leq 0.05$) of paired t-tests between baselines and MS-Bridge, and ▲% indicates improvement over best baseline (underlined).

D. Analysis of the Interplay Between Different Stakeholders' Objectives

In this section, we will perform a thorough analysis to study the relationships and interplay between competing stakeholder interests and how our proposed model manages to effectively balance these competing or even conflicting objectives in multi-stakeholder multi-objective recommendations.

Specifically, we argue that an important benefit provided by our proposed model lies in that it effectively balances between the user objectives and the artist objectives in the provided recommendations. To validate this point, we conduct additional analysis in this section, where we replicate our experiments presented in Section 5 with varying weights for each objective, rather than determining the optimal set of objectives weights, as we did in Section 4.5. When we increase the weights for the artist objectives, we might be able to achieve better performance on these artist objectives, which might come at the cost of sacrificing the performance on user objectives, and vice versa. Therefore, by selecting a set of different objective weights for a specific recommendation model, we will be able to get a set of recommendation performance results correspondingly, which will form a Pareto Frontier for that recommendation model. As we demonstrate in Figure 11, our proposed model manages to reach the level of the Pareto Frontier that significantly dominates the Pareto Frontier of all other baseline recommendation models across all three datasets. This strong performance provides further evidence that our proposed model effectively balances between the conflicting user and artist objectives in the multi-stakeholder multi-objective recommendation task.

E. Recommendation Performance on Robustness Checks

In this section, we will conduct a series of additional offline evaluations on alternative settings to demonstrate the flexibility and robustness of our proposed model.

E.1. Recommendation Performance on Long Tail Data Records

We repeat our offline evaluations on the recommendation performance described in Section 5 on long tail users and long tail media content. Specifically, instead of filtering out users and songs with less than five interactions in the datasets, as we did in Section 5, we will implement and evaluate our proposed model only on these ”filtered” (or long tail) data records. As we can observe from Table 19, our proposed model still achieves significant recommendation

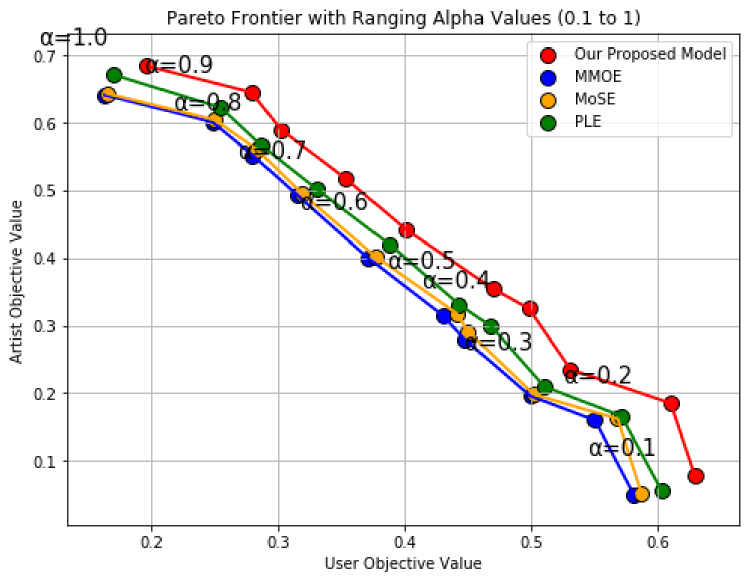


Figure 11 Pareto Frontier of the Recommendation Performance on the Spotify Dataset

performance improvements over state-of-the-art baselines across both the Spotify (Session-Based Recommendation) and the Alibaba-Youku datasets under these conditions, although the performance improvements are not as great as what we can observe in Table 4 in our main analysis. These results indicate that our proposed model is particularly flexible and produces satisfying recommendation performance for long tail users and media content as well, and that it is important to collect sufficient interaction records of users and artists to achieve even better performance.

E.2. Recommendation Performance on Longer or Shorter Training Data Window

For the Spotify (Next Song Recommendation) dataset that we collected directly from the company, we will repeat our offline evaluations in this section with different training data windows to demonstrate the robustness of our proposed method. Specifically, rather than using the first five days of data for training, one day for validation, and the last day for testing, as we did in Section 5, we will use a longer training data window (i.e., the first six days for training, and the last day for testing) and a shorter training data window (i.e., the first four days for training, two days for validation, and the last day for testing) instead. Results in Table 20 and Table 21 show that our proposed method still achieves significant improvements over the baselines under these settings, which further underscores the flexibility of our proposed approach.

E.3. Recommendation Performance on Different Lengths of the Embedding Vector

For the Spotify (Next Song Recommendation) dataset that we used in our offline evaluations, user and song embeddings are processed, generated, and provided directly by Spotify (Anderson et al. 2020, Hansen et al. 2021), which constitute 40-dimensional latent vectors and are re-trained every day to account for new content added to the Spotify platform and for the dynamic shift of user preferences. To demonstrate the robustness of our proposed method, we will repeat our evaluations in this section with different lengths of the embedding vectors. Specifically, we truncate the 40-dimensional vectors into 8, 16, and 32-dimensional vectors, or expand the 40-dimensional vectors into 64 and 128-dimensional vectors, by applying a neural transformation (with a corresponding number of hidden units) on top of the 40-dimensional vectors to change their dimensionality accordingly. As we demonstrate in Table 22, the recommendation performance of our proposed model remains robust across different lengths of the embedding vector, although

Table 19 Recommendation performance on long tail data records of Spotify and Alibaba datasets.

(a) Spotify dataset

Model Type	Model	Pre@10	Rec@10	MAP@10	Pre@5	Rec@5	MAP@5	Pre@3	Rec@3	MAP@3
Single-Objective Models	LR	0.339	0.340	0.331	0.349	0.370	0.350	0.369	0.382	0.361
	NeuMF	0.356	0.390	0.363	0.373	0.393	0.371	0.385	0.403	0.378
	Wide&Deep	0.356	0.390	0.362	0.375	0.393	0.369	0.385	0.401	0.378
	Single-Objective Towers	0.350	0.388	0.351	0.360	0.390	0.362	0.380	0.396	0.375
Multi-Objective Models	MCCF	0.352	0.388	0.369	0.371	0.370	0.370	0.382	0.403	0.379
	MMoE	0.388	0.396	0.381	0.396	0.398	0.397	0.402	0.417	0.401
	MoSE	0.388	0.396	0.381	0.397	0.398	0.398	0.403	0.417	0.403
	NMTR	0.376	0.391	0.376	0.388	0.383	0.383	0.393	0.413	0.395
	PLE	0.398	0.401	0.382	0.398	0.403	0.401	0.404	0.419	0.404
	MO-LinUCB	0.373	0.391	0.393	0.391	0.391	0.389	0.396	0.408	0.392
	MORL	0.373	0.391	0.372	0.393	0.392	0.390	0.396	0.407	0.392
Our Proposed Model	**MS-Bridge**	**0.401***	**0.405***	**0.390***	**0.404***	**0.410***	**0.409***	**0.411***	**0.426***	**0.412***
Improvement	▲ %	0.75%	0.99%	2.05%	1.49%	1.70%	1.96%	1.70%	1.64%	1.94%

(b) Alibaba dataset

Model Type	Model	Pre@10	Rec@10	MAP@10	Pre@5	Rec@5	MAP@5	Pre@3	Rec@3	MAP@3
Single-Objective Models	LR	0.478	0.418	0.408	0.491	0.465	0.438	0.517	0.497	0.455
	NeuMF	0.493	0.439	0.427	0.508	0.485	0.453	0.533	0.511	0.473
	Wide&Deep	0.493	0.441	0.427	0.508	0.487	0.453	0.535	0.511	0.473
	Single-Objective Towers	0.490	0.435	0.421	0.505	0.480	0.447	0.529	0.506	0.469
Multi-Objective Models	MCCF	0.493	0.440	0.425	0.511	0.490	0.453	0.538	0.513	0.475
	MMoE	0.503	0.445	0.433	0.517	0.496	0.460	0.545	0.521	0.481
	MoSE	0.503	0.447	0.433	0.517	0.497	0.461	0.545	0.522	0.483
	NMTR	0.501	0.441	0.430	0.513	0.493	0.455	0.540	0.517	0.479
	PLE	0.505	0.448	0.435	0.519	0.499	0.462	0.547	0.524	0.484
	MO-LinUCB	0.498	0.441	0.425	0.513	0.491	0.455	0.540	0.513	0.477
	MORL	0.498	0.439	0.427	0.513	0.491	0.456	0.541	0.513	0.477
Our Proposed Model	**MS-Bridge**	**0.513***	**0.455***	**0.439***	**0.527***	**0.506***	**0.468***	**0.555***	**0.531***	**0.489***
Improvement	▲ %	1.56%	1.54%	0.91%	1.52%	1.38%	1.28%	1.44%	1.32%	1.02%

Note. * indicates statistical significance ($p \leq 0.05$) using paired t-tests compared to the joint prediction model (+bridge), and ▲% indicates relative improvement of the joint prediction model (+bridge) compared to the best baseline (underlined).

Table 20 Prediction performance on the Spotify dataset with a longer training window.

Model Type	Model	Artist Objectives				User Objectives							
		New Fan		Recurring		%Listening		Num Listening		In Playlist		Feedback Type	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Single-Objective Models	LR	0.356	0.307	0.331	0.269	0.561	0.509	0.397	0.362	0.501	0.451	0.581	0.513
	NeuMF	0.336	0.275	0.314	0.262	0.571	0.468	0.392	0.363	0.490	0.435	0.579	0.441
	Wide&Deep	0.335	0.268	0.314	0.263	0.569	0.459	0.391	0.361	0.488	0.433	0.579	0.410
	Single-Objective Towers	0.332	0.272	0.320	0.262	0.554	0.441	0.390	0.362	0.490	0.433	0.583	0.408
Multi-Objective Models	MCCF	0.352	0.308	0.325	0.266	0.556	0.502	0.391	0.361	0.488	0.445	0.581	0.410
	MMoE	0.342	0.268	0.326	0.267	0.555	0.442	0.390	0.361	0.490	0.439	0.581	0.410
	MoSE	0.339	0.265	0.320	0.266	0.552	0.438	0.390	0.358	0.492	0.436	0.579	0.408
	NMTR	0.347	0.270	0.327	0.268	0.557	0.445	0.397	0.362	0.493	0.440	0.583	0.411
	PLE	0.337	0.265	0.319	0.261	0.551	0.437	0.390	0.358	0.487	0.431	0.577	0.407
	MO-LinUCB	0.341	0.268	0.323	0.266	0.552	0.440	0.391	0.360	0.492	0.436	0.580	0.409
	MORL	0.341	0.269	0.328	0.268	0.553	0.442	0.393	0.362	0.494	0.439	0.582	0.411
MS-Bridge Our Proposed Models	Wide&Deep (No Bridge)	0.330	0.265	0.311	0.249	0.549	0.440	0.391	0.358	0.484	0.424	0.561	0.395
	NCF (No Bridge)	0.325	0.263	0.309	0.246	0.546	0.440	0.390	0.356	0.481	0.419	0.557	0.393
	Wide&Deep	**0.310***	**0.261***	**0.301***	**0.245***	**0.543***	**0.430***	**0.389***	**0.357***	**0.473***	**0.415***	**0.553***	**0.389***
	NCF	**0.302***	**0.256***	**0.298***	**0.242***	**0.541***	**0.429***	**0.388***	**0.355***	**0.469***	**0.411***	**0.552***	**0.383***
Improvement	▲%	11.59%	3.52%	7.05%	7.85%	1.85%	1.86%	0.52%	0.56%	3.84%	4.87%	4.53%	6.27%

Note. * indicates statistical significance ($p \leq 0.05$) using paired t-tests compared to MS-Bridge, and ▲% indicates improvement over the best baseline (underlined).

we will be able to achieve the best performance using the original 40-dimensional embedding vectors provided by Spotify.

E.4. Recommendation Performance on Alternative Objective Functions

In section 5, we use the NDCG metric as the objective function to conduct the multi-stakeholder multi-objective ranking task for our proposed model, since it is the standard practice for optimizing the ranking quality of a recommender

Table 21 Prediction performance on the Spotify dataset with a shorter training window.

Model Type	Model	Artist Objectives				User Objectives							
		New Fan		**Recurring**		**%Listening**		**Num Listening**		**In Playlist**		**Feedback Type**	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Single-Objective Models	LR	0.378	0.322	0.348	0.284	0.579	0.519	0.411	0.378	0.516	0.463	0.593	0.525
	NeuMF	0.344	0.289	0.327	0.271	0.577	0.479	0.399	0.368	0.497	0.441	0.584	0.459
	Wide&Deep	0.346	0.283	0.325	0.270	0.575	0.476	0.398	0.368	0.495	0.440	0.583	0.457
	Single-Objective Towers	0.344	0.286	0.329	0.273	0.577	0.476	0.401	0.368	0.497	0.440	0.588	0.463
Multi-Objective Models	MCCF	0.357	0.315	0.331	0.271	0.563	0.509	0.400	0.369	0.497	0.451	0.590	0.417
	MMoE	0.344	0.271	0.329	0.269	0.559	0.447	0.398	0.366	0.495	0.444	0.586	0.415
	MoSE	0.343	0.268	0.325	0.268	0.556	0.443	0.398	0.365	0.497	0.442	0.588	0.417
	NMTR	0.349	0.276	0.333	0.273	0.563	0.450	0.405	0.369	0.495	0.447	0.584	0.415
	PLE	<u>0.339</u>	<u>0.266</u>	<u>0.321</u>	<u>0.263</u>	<u>0.553</u>	<u>0.441</u>	<u>0.396</u>	<u>0.363</u>	<u>0.493</u>	<u>0.438</u>	<u>0.581</u>	<u>0.411</u>
	MO-LinUCB	0.347	0.273	0.329	0.273	0.560	0.447	0.399	0.369	0.501	0.444	0.588	0.417
	MORL	0.347	0.275	0.335	0.275	0.563	0.450	0.404	0.371	0.503	0.445	0.589	0.419
MS-Bridge Our Proposed Models	Wide&Deep (No Bridge)	0.333	0.268	0.315	0.252	0.551	0.444	0.395	0.362	0.487	0.426	0.564	0.398
	NCF (No Bridge)	0.327	0.265	0.31	0.249	0.549	0.444	0.394	0.358	0.484	0.421	0.559	0.396
	Wide&Deep	**0.312***	**0.263***	**0.303***	**0.247***	**0.546***	**0.432***	**0.391***	**0.360***	**0.475***	**0.417***	**0.556***	**0.391***
	NCF	**0.310***	**0.260***	**0.301***	**0.245***	**0.544***	**0.431***	**0.391***	**0.358***	**0.472***	**0.414***	**0.556***	**0.388***
Improvement	▲%	9.35%	2.31%	6.64%	7.35%	1.65%	2.32%	1.28%	1.40%	4.45%	5.80%	4.50%	5.92%

Note. * indicates statistical significance ($p \leq 0.05$) using paired t-tests compared to MS-Bridge, and ▲% indicates improvement over the best baseline (underlined).

system model (Zhang et al. 2019, Gunawardana et al. 2022). We will now provide additional empirical evidence to demonstrate the validity of our design choice. Specifically, we select the Precision@10 and Recall@10 metrics as the two alternative objective functions to generate the ranking list. As we can observe from Table 23 and Table 24, while the recommendation performance of our proposed model remains relatively robust across these evaluation settings, the NDCG metric indeed leads to the best performance since it is capable of measuring the overall ranking quality more accurately.

E.5. Recommendation Performance on Alternative Artist Objectives

For the Spotify (Next Song Recommendation) dataset that we used in the offline evaluations, we consider the objectives of *New Fan* and *Recurring* to measure the welfare of artists in the recommendation process. Although our selection of the artist objectives strictly follows the industrial practices of the media recommendation task and the guidelines provided by the company of Spotify specifically, we will nevertheless study the performance of our proposed model on alternative artist objectives to study its flexibility in practical use. Specifically, we consider the artist objectives of *Music Consumption Diversity* and *Session Similarity* that have been proposed and discussed in (Hansen et al. 2020). According to the former Director of Research at Spotify, these two objectives are of great importance to the artists since they need concrete guidelines to understand the relevance and novelty of their created content to the targeted audience, so that they will be able to calibrate their future media creations accordingly. The *Music Consumption Diversity* metric is computed as the Euclidean distance between the latent embedding of the recommended song and the latent embedding of the last song that the user has listened to; while the *Session Similarity* metric is computed

Table 22 Comparing different embedding vector sizes in terms of recommendation performance (Precision@10, Recall@10, and MAP@10) using the Spotify dataset.

Recommendation Model	Embedding Size	RS Performance		
		Pre@10	Rec@10	MAP@10
Our Proposed Model	8	0.532	0.487	0.356
	16	0.532	0.485	0.356
	32	0.530	0.485	0.352
	40	**0.536**	**0.489**	**0.359**
	64	0.532	0.486	0.354
	128	0.528	0.483	0.350

Table 23 Comparing recommendation performance (Precision, Recall, and MAP) with different objective functions using the Spotify (Session-Based Recommendation) dataset.

Objective Function	Pre@10	Rec@10	MAP@10	Pre@5	Rec@5	MAP@5	Pre@3	Rec@3	MAP@3
NDCG	0.536	0.489	**0.359**	**0.559**	**0.545**	**0.392**	**0.590**	**0.563**	**0.399**
Precision@10	**0.538**	0.481	0.353	0.550	0.542	0.387	0.582	0.560	0.390
Recall@10	0.528	**0.491**	0.350	0.548	0.540	0.386	0.578	0.568	0.390

Table 24 Comparing recommendation performance with different objective functions using the Alibaba dataset.

Objective Function	Pre@10	Rec@10	MAP@10	Pre@5	Rec@5	MAP@5	Pre@3	Rec@3	MAP@3
NDCG	0.625	0.530	**0.516**	**0.651**	**0.594**	**0.552**	**0.670**	**0.609**	**0.575**
Precision@10	**0.627**	0.521	0.506	0.648	0.587	0.545	0.668	0.600	0.566
Recall@10	0.617	**0.531**	0.510	0.648	0.590	0.546	0.662	0.606	0.569

Table 25 Prediction performance with alternative artist objectives on the Spotify dataset.

Model Type	Model	Artist Objectives				User Objectives							
		Consumption Diversity		**Session Diversity**		**%Listening**		**Num Listening**		**In Playlist**		**Feedback Type**	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Single-Objective Models	LR	0.437	0.362	0.449	0.369	0.559	0.508	0.397	0.363	0.499	0.452	0.583	0.514
	NeuMF	0.418	0.350	0.428	0.358	0.571	0.466	0.393	0.361	0.490	0.435	0.538	0.441
	Wide&Deep	0.417	0.348	0.425	0.356	0.567	0.458	0.390	0.361	0.487	0.430	0.513	0.410
	Single-Objective Towers	0.419	0.352	0.430	0.360	0.554	0.442	0.390	0.360	0.489	0.432	0.510	0.406
Multi-Objective Models	MCCF	0.418	0.353	0.431	0.362	0.559	0.458	0.393	0.360	0.487	0.443	0.515	0.413
	MMoE	0.411	0.348	0.423	0.355	0.554	0.441	0.391	0.359	0.485	0.435	0.513	0.408
	MoSE	0.409	0.347	0.423	0.353	0.552	0.439	0.390	0.357	0.485	0.433	0.511	0.405
	NMTR	0.417	0.355	0.429	0.362	0.559	0.448	0.398	0.363	0.491	0.441	0.524	0.410
	PLE	0.407	0.346	0.421	0.352	0.551	0.439	0.389	0.358	0.484	0.431	0.509	0.403
	MO-LinUCB	0.415	0.353	0.427	0.359	0.555	0.447	0.394	0.363	0.492	0.439	0.522	0.419
	MORL	0.415	0.353	0.429	0.359	0.555	0.448	0.397	0.364	0.494	0.439	0.518	0.419
Our Proposed Models	Wide&Deep (No Bridge)	0.401	0.344	0.419	0.350	0.549	0.437	0.387	0.357	0.483	0.427	0.490	0.402
	NCF (No Bridge)	0.399	0.343	0.417	0.348	0.548	0.437	0.386	0.355	0.481	0.421	0.487	0.401
	Wide&Deep	**0.393***	**0.338***	**0.411***	**0.342***	**0.544***	**0.433***	**0.385***	**0.355***	**0.475***	**0.419***	**0.480***	**0.398***
	NCF	**0.391***	**0.336***	**0.410***	**0.341***	**0.542***	**0.431***	**0.383***	**0.353***	**0.473***	**0.413***	**0.478***	**0.395***
Improvement	▲%	4.09%	2.98%	2.68%	3.23%	1.66%	1.86%	1.57%	1.42%	2.33%	4.36%	6.49%	1.99%

Note. * indicates statistical significance ($p \leq 0.05$) using paired t-tests compared to MS-Bridge, and ▲% indicates improvement over the best baseline (underlined).

as the inverse Euclidean distance between the latent embedding of the recommended song and the average latent embeddings of all songs that the user has listened to in the current recommendation session. The results are shown in Table 25, where we can observe that our proposed model still significantly outperforms all baseline models across all user metrics and artist metrics in this alternative evaluation setting, which further demonstrates the practicality and flexibility of our proposed model.

E.6. Recommendation Performance on the Alternative Cross-Validation Method

For the Spotify (Next Song Recommendation) dataset that we used in the offline evaluations, we applied a five-fold time-based split evaluation strategy, where the first five days of data are used for training (72%), one day for validation, and the last day for testing (14%). In this section, we also considered an alternative five-fold cross-validation strategy, where we randomly split the entire datasets into the training set (72%), validation set (14%), and test set (14%), without considering the time information. We implement our proposed model as well as all baseline models under this alternative setting, and present the results in Table 26. We can observe in this table that our proposed model still manages to significantly outperform all baseline models in this alternative evaluation setting, which further demonstrates the robustness and flexibility of our proposed method.

E.7. Recommendation Performance on Different Embedding Types

In this section, we conduct additional experiments to evaluate the impact of different embedding methods on the recommendation performance of our proposed MS-Bridge model. Specifically, we compare four types of embeddings:

Table 26 Prediction performance under the alternative cross-validation setting on the Spotify dataset.

Model Type	Model	Artist Objectives				User Objectives							
		New Fan		Recurring		%Listening		Num Listening		In Playlist		Feedback Type	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Single-Objective Models	LR	0.373	0.325	0.351	0.293	0.580	0.492	0.414	0.389	0.513	0.477	0.599	0.473
	NeuMF	0.355	0.299	0.345	0.284	0.580	0.483	0.410	0.381	0.503	0.451	0.589	0.444
	Wide&Deep	0.356	0.289	0.341	0.281	0.575	0.479	0.408	0.381	0.498	0.448	0.583	0.434
	Single-Objective Towers	0.358	0.291	0.343	0.289	0.575	0.479	0.408	0.381	0.498	0.451	0.589	0.436
Multi-Objective Models	MCCF	0.368	0.305	0.347	0.281	0.569	0.481	0.403	0.383	0.498	0.444	0.591	0.433
	MMoE	0.351	0.283	0.338	0.276	0.563	0.467	0.398	0.376	0.490	0.439	0.583	0.427
	MoSE	0.347	0.281	0.335	0.275	0.562	0.463	0.396	0.375	0.488	0.436	0.581	0.425
	NMTR	0.357	0.289	0.341	0.279	0.569	0.471	0.401	0.388	0.495	0.444	0.589	0.431
	PLE	<u>0.345</u>	<u>0.280</u>	<u>0.333</u>	<u>0.273</u>	<u>0.561</u>	<u>0.461</u>	<u>0.395</u>	<u>0.372</u>	<u>0.486</u>	<u>0.435</u>	<u>0.579</u>	<u>0.423</u>
	MO-LinUCB	0.352	0.289	0.345	0.281	0.568	0.469	0.399	0.379	0.495	0.439	0.588	0.429
	MORL	0.353	0.291	0.343	0.283	0.569	0.473	0.401	0.379	0.495	0.441	0.588	0.429
Our Proposed Models	Wide&Deep (No Bridge)	0.340	0.276	0.330	0.269	0.557	0.459	0.388	0.365	0.483	0.428	0.569	0.414
	NCF (No Bridge)	0.339	0.274	0.328	0.267	0.556	0.457	0.388	0.362	0.481	0.422	0.565	0.411
	Wide&Deep	**0.331***	**0.266***	**0.323***	**0.260***	**0.548***	**0.448***	**0.385***	**0.357***	**0.474***	**0.415***	**0.553***	**0.402***
	NCF	**0.329***	**0.265***	**0.318***	**0.258***	**0.545***	**0.446***	**0.385***	**0.355***	**0.470***	**0.413***	**0.550***	**0.400***
Improvement	▲%	4.86%	5.66%	4.72%	5.81%	2.94%	3.25%	2.60%	4.79%	3.40%	5.33%	5.27%	5.75%

Note. * indicates statistical significance ($p \leq 0.05$) using paired t-tests compared to MS-Bridge, and ▲% indicates improvement over the best baseline (underlined).

(1) proprietary pre-trained embeddings provided by Spotify, which are 40-dimensional vectors re-trained daily to capture platform-specific user and content semantics; (2) public pre-trained embeddings generated by MuQ (Zhu et al. 2025) for music and Video-LLaMA (Zhang et al. 2023) for video, which are projected to the target dimensionality via a learned linear layer; (3) AutoEncoder embeddings, which are learned from scratch using user, media, and artist features through an autoencoding objective; and (4) One-Hot embeddings, which use raw ID encodings as initialization and are updated through backpropagation during end-to-end training of the recommender system.

As shown in Table 27, proprietary pre-trained embeddings achieve the best performance on the Spotify dataset, while public pre-trained embeddings yield comparatively lower performance. This can be explained by the fact that proprietary embeddings are specifically optimized for the platform's recommendation task and are continuously updated to reflect evolving user preferences and content catalogs, whereas public pre-trained models are designed to capture general audio or video semantics that may not align with the specific interaction patterns relevant to our recommendation objectives. For the Alibaba-Youku dataset, where proprietary embeddings are not available, AutoEncoder embeddings outperform both public pre-trained and One-Hot alternatives, confirming that task-specific feature learning is more beneficial than general-purpose representations for our multi-stakeholder recommendation task.

Table 27 Recommendation performance of MS-Bridge with different embedding types.

Embedding Type	Recommendation (Spotify)				Recommendation (Alibaba)			
	Pre@10	Rec@10	MAP@10	MS-NSW	Pre@10	Rec@10	MAP@10	MS-NSW
Proprietary Embedding (Spotify)	**0.548***	**0.496***	**0.363***	**0.496***				
AutoEncoder (Alibaba)					**0.627***	**0.534***	**0.520***	**0.368***
Public Embedding (MuQ/Video-Llama)	0.539	0.491	0.355	0.485	0.610	0.523	0.509	0.356
One-Hot Embedding (Backpropagation)	0.511	0.470	0.348	0.461	0.589	0.508	0.498	0.339

Note. All results use the MS-Bridge (NCF backbone) model $*$ indicates statistical significance at the 0.05 level.

F. Additional Methods for Identifying the Objective Weights

In this section, we will introduce the details of two additional methods to identify the objective weights, besides the ordinal regression method that we described in Section 4.5 of the main paper.

First, we present a neural network-based approach to optimize the objective weights in an end-to-end manner, which consists of a fully-connected hidden layer that we add on top of the objective prediction model (as illustrated in

Figure 2) to produce the weights w_i. As a result, we can learn the complex relationships between different objectives and obtain the optimal weights w_i in a data-driven fashion, leading to significantly better recommendation performance as we show in the next section. This hidden layer for aggregation is optimized by minimizing the ranking loss function of the Normalized Discounted Cumulative Gain (NDCG) (Valizadegan et al. 2009) in the top-K ranking task, which indicates the ranking performance of media recommendations. We have also conducted additional experiments in Appendix E.4 to test different ranking loss functions, such as Precision and Recall, where we still observe significant performance improvements in our proposed model.

Besides the aforementioned neural network-based method, the weights W can also be determined through Bayesian Hyperparameter Optimization (Snoek et al. 2012, Feurer et al. 2015), which follows a three-stage iterative training procedure. To begin with, we query the NDCG ranking loss with an initial set of weights W and record the resulting weight-loss pairs $(W, Loss_{NDCG}(W))$. In the first stage, we fit a probabilistic model M to all $(W, Loss_{NDCG}(W))$. Then in the second stage, we select an optimized set of weights W^* using M to maximize an acquisition function $a(W,M)$, which determines the quality of each set of weight W. In the third stage, we evaluate the NDCG ranking loss function using the new set of weights W^* and go back to the first stage until convergence. Following the common practice in the literature (Jones et al. 1998), we select the acquisition function $a(W,M)$ as the Expected Improvement (EI) over the best weight found so far, and the probabilistic model M as the Gaussian process $p_M(y|W)$:

$$a_{EI}(W,M) = \int max(y^* - y, 0) p_M(y|W) dy, \tag{17}$$

where y is the variable of integration and y^* stands for the estimated maximum of the acquisition function. We then identify the optimal weights w^* that maximize the EI of $a(W,M)$.

G. Proof of Theorem 1: Axiomatic Uniqueness of MS-NSW

G.1. Preliminary: Reformulation as a Functional Equation

We seek all continuous $W : \mathbb{R}^2_+ \to \mathbb{R}$ satisfying (SI), (PO), (IIA), and (CSC). Without loss of generality, we work with the log-transformed utilities $v_A = \ln U_A$ and $v_B = \ln U_B$ (since $U_A, U_B > 0$ on the interior of the feasible set), and seek the functional form of $\tilde{W}(v_A, v_B) = W(e^{v_A}, e^{v_B})$.

G.2. Part 1: Uniqueness (MS-NSW is the Only Metric Satisfying All Four Axioms)

Step 1: CSC constrains the functional form. By (CSC), $\partial^2 W / \partial U_A \partial U_B > 0$, so W is supermodular in (U_A, U_B). Combined with the zero condition $W(U_A, 0) = W(0, U_B) = 0$, we can write $W(U_A, U_B) = U_A \cdot U_B \cdot \phi(U_A, U_B)$ for some continuous positive function ϕ. This follows from the fact that any supermodular function vanishing on the axes can be expressed as a product of the two arguments times a residual term.

Step 2: SI further constrains ϕ. Under (SI), for any $\lambda_A, \lambda_B > 0$:

$$\arg\max W(\lambda_A U_A, \lambda_B U_B) = \arg\max W(U_A, U_B) \tag{18}$$

Substituting $W = U_A U_B \phi(U_A, U_B)$:

$$\arg\max\ \lambda_A \lambda_B\, U_A U_B\, \phi(\lambda_A U_A, \lambda_B U_B) = \arg\max\ U_A U_B\, \phi(U_A, U_B) \tag{19}$$

Since $\lambda_A\lambda_B > 0$ is a constant, SI requires $\phi(\lambda_A U_A, \lambda_B U_B) = \phi(U_A, U_B)$ for all $\lambda_A, \lambda_B > 0$ and all (U_A, U_B) in the feasible set. Setting $\lambda_A = 1/U_A$ and $\lambda_B = 1/U_B$:

$$\phi(U_A, U_B) = \phi(1, 1) =: c > 0 \tag{20}$$

Therefore, ϕ must be a positive constant, and:

$$W(U_A, U_B) = c \cdot U_A \cdot U_B \tag{21}$$

Since maximizing $c \cdot U_A \cdot U_B$ is equivalent to maximizing $\sqrt{U_A \cdot U_B}$ (as both are monotone transformations of $U_A U_B$), SI and CSC together uniquely determine MS-NSW up to monotone transformation.

Step 3: PO is satisfied by Equation (21)**.** The gradient of $W = U_A U_B$ satisfies $\partial W/\partial U_A = U_B > 0$ and $\partial W/\partial U_B = U_A > 0$ throughout the interior of the feasible set. Since both partial derivatives are strictly positive, any maximizer of W must lie on the Pareto frontier (otherwise a feasible direction that increases both U_A and U_B would also increase W, contradicting optimality). Therefore (PO) is satisfied.

Step 4: IIA is satisfied by Equation (21)**.** The ranking induced by $W = U_A U_B$ between any two recommendations r_1 and r_2 depends only on whether $U_A(r_1)U_B(r_1) \gtrless U_A(r_2)U_B(r_2)$, which is a function of the utility pairs of r_1 and r_2 alone. No third alternative r_3 enters this comparison, so (IIA) is satisfied.

Step 5: Uniqueness. Steps 1-2 show that SI and CSC together uniquely force $W(U_A, U_B) = c \cdot U_A \cdot U_B$, and Steps 3-4 confirm PO and IIA are also satisfied. Since the derivation in Steps 1-2 exhausts all continuous functions satisfying SI and CSC simultaneously, MS-NSW is the unique element of $\mathcal{W}$ satisfying all four axioms. □

G.3. Part 2(i): Additive Metrics Fail SI and CSC

Let $W^\lambda = \lambda U_A + (1-\lambda)U_B$ for any fixed $\lambda \in (0, 1)$.

Failure of SI. Under rescaling by $(\lambda_A, \lambda_B) = (2, 1)$:

$$\arg\max\, W^\lambda(2U_A, U_B) = \arg\max\,[2\lambda U_A + (1-\lambda)U_B] \tag{22}$$

which has a different optimal solution than $\arg\max\,[\lambda U_A + (1-\lambda)U_B]$ whenever $\lambda \neq 0$ and $\lambda \neq 1$, since the relative weight on U_A has changed from λ to $2\lambda/(2\lambda + (1-\lambda))$. Therefore W^λ violates (SI) for all $\lambda \in (0, 1)$.

Failure of CSC. $\partial^2 W^\lambda/\partial U_A \partial U_B = 0$ for any additive function, violating the strict inequality in (CSC). Furthermore, $W^\lambda(U_A, 0) = \lambda U_A \neq 0$ for $U_A > 0$, violating the zero condition of (CSC): an additive metric assigns positive welfare to a recommendation that generates zero artist utility, regardless of how high user engagement is. In the Spotify setting, this means recommending a superstar artist with *New Fan* = *Recurring Fan* = 0 can achieve high W^λ despite contributing nothing to the supply side of the platform. □

G.4. Part 2(ii): Min-Based Metrics Fail SI and PO

Let $W^{\min} = \min(U_A, U_B)$.

Failure of SI. Under rescaling by $(\lambda_A, \lambda_B) = (2, 1)$:

$$\arg\max\, \min(2U_A, U_B) \neq \arg\max\, \min(U_A, U_B) \tag{23}$$

in general, since doubling U_A shifts which side is the binding constraint in the min operator. For example, if $U_A = 0.3$ and $U_B = 0.4$, then $W^{\min} = 0.3$, but under scaling $W^{\min}(2 \times 0.3, 0.4) = 0.4$, which corresponds to a different active constraint. $W^{\min}$ therefore violates (SI).

Failure of PO. Consider two recommendations $r_1 = (U_A, U_B) = (0.5, 0.5)$ and $r_2 = (0.8, 0.5)$. Then $W^{\min}(r_1) = W^{\min}(r_2) = 0.5$, so the min-metric is indifferent between r_1 and r_2 despite the fact that r_2 Pareto-dominates r_1 (it achieves higher U_A with equal U_B). This indifference to Pareto-improving changes violates (PO). In our Spotify setting, this corresponds to a metric that is indifferent between a recommendation generating (%Listening = 0.5, New Fan = 0.5) and one generating (%Listening = 0.8, New Fan = 0.5), despite the latter being strictly better for users. □

G.5. Part 2(iii): Max-Based Metrics Fail PO and CSC

Let $W^{\max} = \max(U_A, U_B)$.

Failure of PO. $W^{\max}$ is maximized by maximizing the larger of the two utilities, which generally requires driving one stakeholder's utility to its maximum while ignoring the other. The maximizer of $W^{\max}$ therefore lies at a corner of $\mathcal{F}$ rather than its interior, and is Pareto-dominated by interior points where both utilities are jointly improved. This violates (PO).

Failure of CSC. $\partial^2 W^{\max}/\partial U_A \partial U_B = 0$ wherever $U_A \neq U_B$, so the cross-side complementarity condition is violated almost everywhere. Furthermore, $W^{\max}(0, U_B) = U_B > 0$, violating the zero condition: a recommendation generating zero user utility but positive artist utility receives positive welfare under $W^{\max}$. □

H. Proof of Theorem 2: Dominance of MS-NSW

H.1. Part (i): Welfare Dominance

By definition, π^*_{NSW} maximizes $\text{MS-NSW}(U_A, U_B) = \sqrt{U_A U_B}$ over all feasible policies, so:

$$\text{MS-NSW}(\pi^*_{\text{NSW}}) = \max_{\pi} \sqrt{U_A(\pi) U_B(\pi)} \geq \sqrt{U_A(\pi^*_{\text{alt}}) U_B(\pi^*_{\text{alt}})} = \text{MS-NSW}(\pi^*_{\text{alt}}) \tag{24}$$

for any policy π^*_{alt}, with equality only if π^*_{alt} happens to also maximize MS-NSW. By Part 2(i) of Theorem 1, any additive metric W^{λ} fails SI and CSC, so its optimizer π^*_{alt} is generally not the MS-NSW maximizer.

To establish the strict inequality when π^*_{alt} falls into a market failure mode, we consider two cases:

Case 1: Superstar Trap. An alternative metric W^{alt} that assigns positive welfare when $U_B = 0$ (violating CSC) will select π^*_{alt} such that $U_B(\pi^*_{\text{alt}}) \approx 0$, which occurs when only superstar artists (*New Fan* ≈ 0, *Recurring Fan* ≈ 0) are recommended. Therefore:

$$\text{MS-NSW}(\pi^*_{\text{alt}}) = \sqrt{U_A(\pi^*_{\text{alt}}) \cdot 0} = 0 < \text{MS-NSW}(\pi^*_{\text{NSW}}) \tag{25}$$

since π^*_{NSW} by construction identifies latent fan matches that yield strictly positive $U_B > 0$, as confirmed by the fan match results in Section 3.2 and the New Fan AUC improvement of 36.51%.

Case 2: Indiscriminate Exposure. An alternative metric W^{alt} that assigns positive welfare when $U_A = 0$ will select π^*_{alt} that forces niche content on non-fans, yielding $U_A(\pi^*_{\text{alt}}) \approx \underline{U}_A$ (near random-recommendation utility). Therefore:

$$\text{MS-NSW}(\pi^*_{\text{alt}}) = \sqrt{\underline{U}_A \cdot U_B} < \sqrt{U_A(\pi^*_{\text{NSW}}) \cdot U_B(\pi^*_{\text{NSW}})} = \text{MS-NSW}(\pi^*_{\text{NSW}}) \tag{26}$$

since π^*_{NSW} achieves $U_A > \underline{U}_A$ by restricting niche recommendations to users with high latent affinity, as confirmed by the user utility of 0.848 versus 0.214 in Section 3.2. □

H.2. Part (ii): Market Failure Prevention

Prevention of Superstar Trap. Since $\text{MS-NSW} = \sqrt{U_A U_B}$, any recommendation with $U_B = 0$ achieves $\text{MS-NSW} = 0$, which is weakly dominated by any recommendation with $U_B > 0$ and $U_A > 0$. Therefore, the MS-NSW optimizer strictly prefers recommendations that yield positive artist welfare, establishing Equation (7). This is the direct consequence of the zero condition in (CSC), which is uniquely satisfied by MS-NSW among all metrics in $\mathcal{W}$ (as shown in Part 2(i) of Theorem 1).

Prevention of Indiscriminate Exposure. By (PO), the MS-NSW optimizer lies on the Pareto frontier of $\mathcal{F}$. Any recommendation with $U_A \leq \underline{U}_A$ (the random-policy utility level) is Pareto-dominated by the MS-NSW-optimal recommendation, which simultaneously achieves $U_A > \underline{U}_A$ and $U_B > 0$. Therefore the MS-NSW optimizer satisfies Equation (8).

Uniqueness of market failure prevention. We must show no other metric in $\mathcal{W}$ simultaneously prevents both failure modes. Any metric satisfying (7) must assign zero welfare when $U_B = 0$, which by Step 1 of the uniqueness proof forces $W = c \cdot U_A \cdot U_B$ (the product form). Combined with (8), which requires the metric to prefer interior solutions over corner solutions in $\mathcal{F}$, the only metric satisfying both simultaneously is MS-NSW. □

H.3. Part (iii): Long-Run Ecosystem Growth

Under iterative deployment, at round t the recommender system is trained on the accumulated interaction history $\mathcal{D}_t = \{(u, i, a, \mathbf{o}_{u,i,a})\}_{s \leq t}$, where $\mathbf{o}_{u,i,a}$ denotes the vector of observed objective values. The key observation is that the quality of $\mathcal{D}_t$, specifically the diversity and coverage of artist-user interactions it contains, affects the accuracy of the recommender system at round $t+1$.

Feedback loop under MS-NSW. At each round t, the MS-NSW optimizer deliberately exposes users to niche artists for whom they have high latent affinity (New Fan > 0, %Listening > 0). This generates diverse interaction records: $\mathcal{D}_t^{\text{NSW}}$ contains observations of users interacting with a wide range of artists, including niche artists outside the mainstream popularity. The *Recurring Fan* objective is computed as a function of $\text{NumTimesListenedBefore}(u, a, t)$ and $\text{LastTimeListened}(u, a, t)$, both of which are updated each round. By successfully converting latent fans into genuine fans in early rounds, MS-NSW generates increasingly informative *Recurring Fan* signals in later rounds, improving the artist tower's predictive accuracy and further improving recommendation quality.

Feedback loop under additive metrics. By contrast, the additive baseline optimizer concentrates recommendations on superstar artists, as documented in Section 3.2. This generates $\mathcal{D}_t^{\text{alt}}$ that is dominated by interactions with a small set of popular artists. The $\text{NewFan}(u, a, t)$ and $\text{RecurringFan}(u, a, t)$ objectives for niche artists remain close to zero across rounds (because those artists are never recommended), providing no training signal for the artist tower with respect to niche content. The model, therefore, cannot improve its ability to identify latent fan matches over time, and the performance gap relative to MS-NSW grows monotonically.

Formal bound. Let $\Delta_t = \text{MS-NSW}(\pi_t^*) - \text{MS-NSW}(\pi_t^{\text{alt}}) > 0$ denote the per-round welfare gap established in Part (i). Since the MS-NSW optimizer generates strictly more diverse training data each round, the artist tower's prediction accuracy improves strictly faster under MS-NSW than under the additive baseline, by Proposition 1 (C1: cross-side covariance carries signal that only the bridge can exploit). Therefore Δ_t is non-decreasing in t, and:

$$\sum_{t=1}^{T} \text{MS-NSW}(\pi_t^*) - \sum_{t=1}^{T} \text{MS-NSW}(\pi_t^{\text{alt}}) = \sum_{t=1}^{T} \Delta_t \geq T \cdot \Delta_1 = \Omega(T \cdot \Delta) \tag{27}$$

where $\Delta = \Delta_1 > 0$ is the first-round gap. This is confirmed by Figure 7, where the MS-Bridge model (optimizing MS-NSW) widens its performance gap over MMoE and PLE monotonically across all 15 simulation rounds on both the Spotify and Alibaba-Youku datasets. □

I. Proof of Theorem 3

We prove this by reduction from the NP-Complete **Partition Problem**, which is formulated as follows: given a multiset of positive integers $S = \{s_1, s_2, \ldots, s_n\}$, determine whether there exists a partition of S into two subsets S_1 and S_2 such that their sums are equal:

$$\sum_{x \in S_1} x = \sum_{y \in S_2} y = \frac{1}{2} \sum_{z \in S} z = V \tag{28}$$

We now consider a simplified version of our recommendation Problem as follows:

- **Stakeholders:** Exactly two stakeholders A and B (a subset of the general multi-stakeholder case).
- **Items:** There are n available media items, corresponding one-to-one with the integers in S.
- **Utilities:** Let the utility of item i be s_i for both stakeholders: $u_{A,i} = u_{B,i} = s_i$.
- **Constraint:** The platform must allocate every item to exactly one user (partition the catalog).

The objective of MS-NSW is to maximize the geometric mean of the aggregated stakeholder utilities:

$$\max \sqrt{\left(\sum_{i \in S_A} s_i\right) \cdot \left(\sum_{j \in S_B} s_j\right)} \quad \text{s.t.} \quad S_A \cup S_B = S, S_A \cap S_B = \emptyset \tag{29}$$

Let $X = \sum_{i \in S_A} s_i$ and $Y = \sum_{j \in S_B} s_j$. We know that $X + Y = 2V$ (the constant total sum of S). The product $X \cdot Y$ is strictly maximized when $X = Y = V$, based on the inequality of arithmetic and geometric means. Therefore, the MS-NSW is maximized at value V^2 if and only if a perfect partition exists. Consequently, any algorithm that can solve the MS-NSW optimization exactly could be used to solve the Partition Problem. Since Partition is NP-Complete, MS-NSW is NP-Hard. □

J. Proof of Theorem 4

We prove each part of Theorem 4 in turn. Throughout, we write $\hat{\mu}_A(\mathbf{x}) = \mathbb{E}[U_A \mid \mathbf{X} = \mathbf{x}]$, $\hat{\mu}_B(\mathbf{x}) = \mathbb{E}[U_B \mid \mathbf{X} = \mathbf{x}]$, $\sigma_{AB}(\mathbf{x}) = \text{Cov}(U_A, U_B \mid \mathbf{X} = \mathbf{x})$, and $\sigma_A^2(\mathbf{x})$, $\sigma_B^2(\mathbf{x})$ for the corresponding conditional variances.

J.1. Preliminary Lemma: Decoupled Estimator Bias

LEMMA 1 **(Bias of Decoupled Estimation).** *Under condition* (C1)*, the decoupled estimator* $\hat{W}^{\text{dec}}(\mathbf{x})$ *satisfies:*

$$\mathcal{B}(\mathbf{x}) := \hat{W}^{\text{dec}}(\mathbf{x}) - W(\mathbf{x}) = \frac{1}{8\sqrt{\hat{\mu}_A \hat{\mu}_B}} \left[\frac{\hat{\mu}_B}{\hat{\mu}_A} \sigma_A^2 - 2\sigma_{AB} + \frac{\hat{\mu}_A}{\hat{\mu}_B} \sigma_B^2\right] + O(\epsilon^3) \tag{30}$$

where ϵ denotes the magnitude of deviations of (U_A, U_B) from $(\hat{\mu}_A, \hat{\mu}_B)$. Moreover, $\mathcal{B}(\mathbf{x}) \geq 0$ for all $\mathbf{x}$, with equality if and only if $U_A/U_B = \hat{\mu}_A/\hat{\mu}_B$ almost surely (i.e., utilities are perfectly proportional).

A pply a second-order Taylor expansion of $f(u_A, u_B) = \sqrt{u_A u_B}$ around the point $(\hat{\mu}_A, \hat{\mu}_B)$. The first-order partial derivatives are $\partial f/\partial u_A = \frac{1}{2}\sqrt{u_B/u_A}$ and $\partial f/\partial u_B = \frac{1}{2}\sqrt{u_A/u_B}$. The second-order partial derivatives evaluated at $(\hat{\mu}_A, \hat{\mu}_B)$ are:

$$\frac{\partial^2 f}{\partial u_A^2}\bigg|_{(\hat{\mu}_A,\hat{\mu}_B)} = -\frac{\sqrt{\hat{\mu}_B}}{4\hat{\mu}_A^{3/2}}, \qquad \frac{\partial^2 f}{\partial u_A \partial u_B}\bigg|_{(\hat{\mu}_A,\hat{\mu}_B)} = \frac{1}{4\sqrt{\hat{\mu}_A\hat{\mu}_B}}, \qquad \frac{\partial^2 f}{\partial u_B^2}\bigg|_{(\hat{\mu}_A,\hat{\mu}_B)} = -\frac{\sqrt{\hat{\mu}_A}}{4\hat{\mu}_B^{3/2}} \tag{31}$$

Taking the conditional expectation of the Taylor expansion and using the fact that $\mathbb{E}[U_A - \hat{\mu}_A \mid \mathbf{X}] = \mathbb{E}[U_B - \hat{\mu}_B \mid \mathbf{X}] = 0$, all first-order terms vanish. The second-order terms yield:

$$\begin{aligned} W(\mathbf{x}) &= \mathbb{E}[f(U_A, U_B) \mid \mathbf{X} = \mathbf{x}] \\ &\approx \sqrt{\hat{\mu}_A\hat{\mu}_B} - \frac{\sqrt{\hat{\mu}_B}}{8\hat{\mu}_A^{3/2}}\sigma_A^2 + \frac{\sigma_{AB}}{4\sqrt{\hat{\mu}_A\hat{\mu}_B}} - \frac{\sqrt{\hat{\mu}_A}}{8\hat{\mu}_B^{3/2}}\sigma_B^2 \end{aligned} \tag{32}$$

Since $\hat{W}^{\mathrm{dec}}(\mathbf{x}) = \sqrt{\hat{\mu}_A\hat{\mu}_B}$ by definition, the bias is:

$$\begin{aligned} \mathcal{B}(\mathbf{x}) &= \sqrt{\hat{\mu}_A\hat{\mu}_B} - W(\mathbf{x}) \\ &= \frac{\sqrt{\hat{\mu}_B}}{8\hat{\mu}_A^{3/2}}\sigma_A^2 - \frac{\sigma_{AB}}{4\sqrt{\hat{\mu}_A\hat{\mu}_B}} + \frac{\sqrt{\hat{\mu}_A}}{8\hat{\mu}_B^{3/2}}\sigma_B^2 \\ &= \frac{1}{8\sqrt{\hat{\mu}_A\hat{\mu}_B}}\left[\frac{\hat{\mu}_B}{\hat{\mu}_A}\sigma_A^2 - 2\sigma_{AB} + \frac{\hat{\mu}_A}{\hat{\mu}_B}\sigma_B^2\right] \end{aligned} \tag{33}$$

which establishes Equation (30). To establish $\mathcal{B}(\mathbf{x}) \geq 0$, observe that the bracketed expression equals:

$$\mathrm{Var}\left(\sqrt{\frac{\hat{\mu}_B}{\hat{\mu}_A}}U_A - \sqrt{\frac{\hat{\mu}_A}{\hat{\mu}_B}}U_B \,\middle|\, \mathbf{X}\right) \geq 0 \tag{34}$$

since it is a variance. Alternatively, the global bound $\mathcal{B}(\mathbf{x}) \geq 0$ follows directly from the Cauchy-Schwarz inequality: $\mathbb{E}[\sqrt{U_A U_B}] \leq \sqrt{\mathbb{E}[U_A]\cdot\mathbb{E}[U_B]}$. Equality holds if and only if $\sqrt{\hat{\mu}_B/\hat{\mu}_A}\,U_A = \sqrt{\hat{\mu}_A/\hat{\mu}_B}\,U_B$ almost surely, i.e., $U_A/U_B = \hat{\mu}_A/\hat{\mu}_B$ a.s.

J.2. Proof of Part (i): Sufficiency of C1–C3

Under (C1), Lemma 1 establishes $\mathcal{B}(\mathbf{x}) > 0$ for some $\mathbf{x}$, so $\hat{W}^{\mathrm{dec}}$ is a biased estimator of $W(\mathbf{x})$.

Under (C2), the bias is *differentially heterogeneous* across items. To see this, observe from Equation (30) that for any two items $i \neq j$:

$$\mathcal{B}(i) - \mathcal{B}(j) \propto \left[\frac{\hat{\mu}_B^{(i)}}{\hat{\mu}_A^{(i)}}\sigma_A^{(i)2} - 2\sigma_{AB}^{(i)} + \frac{\hat{\mu}_A^{(i)}}{\hat{\mu}_B^{(i)}}\sigma_B^{(i)2}\right] - \left[\frac{\hat{\mu}_B^{(j)}}{\hat{\mu}_A^{(j)}}\sigma_A^{(j)2} - 2\sigma_{AB}^{(j)} + \frac{\hat{\mu}_A^{(j)}}{\hat{\mu}_B^{(j)}}\sigma_B^{(j)2}\right] \tag{35}$$

which is non-zero whenever the conditional covariance structures differ (C2).

A ranking reversal occurs when item i is preferred over item j under the decoupled ranking ($\hat{W}^{\mathrm{dec}}(i) > \hat{W}^{\mathrm{dec}}(j)$) while the reverse is true under the true MS-NSW ranking ($W(j) > W(i)$). Using the identity:

$$W(j) - W(i) = \underbrace{\left[\hat{W}^{\mathrm{dec}}(j) - \hat{W}^{\mathrm{dec}}(i)\right]}_{\text{decoupled signal}} + \underbrace{[\mathcal{B}(i) - \mathcal{B}(j)]}_{\text{differential bias}} \tag{36}$$

a ranking reversal occurs when $\mathcal{B}(i) - \mathcal{B}(j) > \hat{W}^{\mathrm{dec}}(i) - \hat{W}^{\mathrm{dec}}(j) > 0$. Under C2, $\mathcal{B}(i) \neq \mathcal{B}(j)$ for some i, j; in a two-sided market with a long-tail catalog, the differential bias $\mathcal{B}(\text{star}) - \mathcal{B}(\text{niche})$ is strictly positive and bounded away from zero by the covariance gap between the two item types (as shown in Proposition J.6). Therefore, ranking reversals occur with positive probability, so $\sup_{\Pi_{\text{dec}}} \mathbb{E}_u[W] < W(\pi^*)$ under C1 and C2 alone.

Under (C3), the MS-NSW optimum π^* is not a solution to any linear scalarization $\lambda U_A + (1-\lambda)U_B$. This matters because any decoupled architecture that assigns fixed weights λ to the tower losses is implicitly optimizing a linear scalarization. Since no such λ achieves π^* under C3, $\sup_{\Pi_{\text{dec}}} \mathbb{E}_u[W]$ remains strictly below $W(\pi^*)$ even after correcting for any constant bias. Together with the ranking reversal argument under C1 and C2, the strict inequality in Equation (10) is established. □

J.3. Proof of Part (ii): Necessity of C1

If C1 fails, then $U_A \perp U_B \mid \mathbf{X}$ everywhere, so $\sigma_{AB}(\mathbf{x}) = 0$ for all $\mathbf{x}$. Lemma 1 then gives:

$$\mathcal{B}(\mathbf{x}) = \frac{1}{8\sqrt{\hat{\mu}_A \hat{\mu}_B}} \left[\frac{\hat{\mu}_B}{\hat{\mu}_A} \sigma_A^2 + \frac{\hat{\mu}_A}{\hat{\mu}_B} \sigma_B^2 \right] \tag{37}$$

which, while non-negative, depends only on within-side variances. When utilities are also deterministic given $\mathbf{X}$ (i.e., $\sigma_A^2 = \sigma_B^2 = 0$), the bias is exactly zero and $\hat{W}^{\text{dec}}(\mathbf{x}) = W(\mathbf{x})$ exactly for all items. More generally, when utilities are stochastic but conditionally independent, $W(\mathbf{x}) = \mathbb{E}[\sqrt{U_A} \mid \mathbf{X}] \cdot \mathbb{E}[\sqrt{U_B} \mid \mathbf{X}]$ factors into a product of two terms, each estimable from the corresponding tower alone. A decoupled architecture can therefore achieve the MS-NSW optimum without cross-side information sharing, so the bridge adds no value. □

J.4. Proof of Part (iii): Necessity of C2

If C2 fails, then $\mathcal{B}(\mathbf{x}) = c$ for some constant $c \geq 0$ and all $\mathbf{x}$. From Equation (36):

$$W(j) - W(i) = \hat{W}^{\text{dec}}(j) - \hat{W}^{\text{dec}}(i) + \underbrace{(c - c)}_{=0} = \hat{W}^{\text{dec}}(j) - \hat{W}^{\text{dec}}(i) \tag{38}$$

so the ranking induced by $\hat{W}^{\text{dec}}$ is identical to the ranking induced by the true W for all item pairs. No ranking reversal can occur, and the decoupled architecture achieves the same top-K recommendations as π^*. The bridge adds no improvement to the achievable ranking quality, so $\sup_{\Pi_{\text{bridge}}} \mathbb{E}_u[W] = \sup_{\Pi_{\text{dec}}} \mathbb{E}_u[W]$. □

J.5. Proof of Part (iv): Necessity of C3

If C3 fails, the achievable utility set $\mathcal{F}$ is convex. By the supporting hyperplane theorem (Bauschke and Combettes 2017), for any $\pi^* \in \arg\max_\pi \sqrt{U_A(\pi) U_B(\pi)}$, there exists a $\lambda^* \in [0, 1]$ such that:

$$\pi^* \in \arg\max_\pi \left[\lambda^* U_A(\pi) + (1 - \lambda^*) U_B(\pi)\right] \tag{39}$$

This weighted sum can be optimized by assigning a loss weight λ^* to the user tower and $(1 - \lambda^*)$ to the artist tower and training them independently, since the gradient of a linear combination decomposes additively:

$$\nabla_\Theta \left[\lambda^* U_A + (1-\lambda^*) U_B\right] = \lambda^* \nabla_{\Theta_A} U_A + (1-\lambda^*) \nabla_{\Theta_B} U_B \tag{40}$$

No cross-side information sharing is required to compute or apply this gradient. A decoupled architecture with the appropriately tuned scalar weight λ^* therefore achieves π^*, and the bridge provides no additional benefit. □

J.6. Proposition: Two-Sided Markets Satisfy C1–C3

PROPOSITION 1. *Any two-sided media market exhibiting* (M1) cross-side network effects*,* (M2) a long-tail content distribution*, and* (M3) discrete recommendation slot constraints *satisfies conditions* (C1)*,* (C2)*, and* (C3) *respectively.*

C1 follows from M1. Cross-side network effects imply that U_A and U_B co-vary positively for latent fans: when user u is a latent fan of artist a (i.e., $I_{u,a}=1$), both U_A (high listening engagement, positive feedback) and U_B (new fan acquisition) are jointly elevated by the same underlying event. Formally, let $p = P(I_{u,a}=1)$, $\mu_A^{(1)} = \mathbb{E}[U_A \mid I=1, \mathbf{X}]$, $\mu_A^{(0)} = \mathbb{E}[U_A \mid I=0, \mathbf{X}]$, and define $\mu_B^{(1)}, \mu_B^{(0)}$ analogously. By the law of total covariance:

$$\mathrm{Cov}(U_A, U_B \mid \mathbf{X}) = \underbrace{p\,\mathrm{Cov}(U_A, U_B \mid I=1, \mathbf{X}) + (1-p)\,\mathrm{Cov}(U_A, U_B \mid I=0, \mathbf{X})}_{\geq 0} + \underbrace{p(1-p)(\mu_A^{(1)} - \mu_A^{(0)})(\mu_B^{(1)} - \mu_B^{(0)})}_{>0} \tag{41}$$

since $\mu_A^{(1)} > \mu_A^{(0)}$ and $\mu_B^{(1)} > \mu_B^{(0)}$ under cross-side network effects. The second term is strictly positive, so $\mathrm{Cov}(U_A, U_B \mid \mathbf{X}) > 0$, satisfying (C1).

C2 follows from M2. Under the long-tail structure, superstar and niche items have categorically different covariance structures. For superstar artists, $U_B^{(\text{star})} \approx 0$ due to market saturation, so $\sigma_{AB}^{(\text{star})} \approx 0$. For niche artists, the bimodal distribution of U_A across latent fans ($I_{u,a}=1$, high U_A) and non-fans ($I_{u,a}=0$, low U_A), combined with the high marginal U_B for successful fan matches, creates a large positive $\sigma_{AB}^{(\text{niche})} > 0$. Substituting into Equation (30):

$$\mathcal{B}(\text{star}) - \mathcal{B}(\text{niche}) \approx \frac{-2(\sigma_{AB}^{(\text{star})} - \sigma_{AB}^{(\text{niche})})}{8\sqrt{\hat{\mu}_A \hat{\mu}_B}} = \frac{2\sigma_{AB}^{(\text{niche})}}{8\sqrt{\hat{\mu}_A \hat{\mu}_B}} > 0 \tag{42}$$

so the differential bias is strictly positive and bounded away from zero, satisfying (C2).

C3 follows from M3. By Theorem 3, the discrete recommendation problem is NP-Hard. Any convex combination of the utility vectors $(U_A(\pi_1), U_B(\pi_1))$ and $(U_A(\pi_2), U_B(\pi_2))$ for two deterministic policies π_1, π_2 is achievable only by a randomized policy $\pi_\lambda = \lambda\pi_1 + (1-\lambda)\pi_2$. Since the geometric mean is strictly concave in λ, the MS-NSW of the mixture satisfies:

$$\sqrt{U_A(\pi_\lambda) U_B(\pi_\lambda)} \geq \lambda \sqrt{U_A(\pi_1) U_B(\pi_1)} + (1-\lambda)\sqrt{U_A(\pi_2) U_B(\pi_2)} \tag{43}$$

which implies the achievable utility set $\mathcal{F}$ under the MS-NSW objective is non-convex. By the argument of Section 3.2, the MS-NSW optimum therefore lies in the non-convex interior of $\mathcal{F}$ and is not attainable by any linear scalarization, satisfying (C3).

In particular, all four datasets used in the offline evaluations of Section 5, Spotify (Next Song Recommendation), Spotify (Session-Based Recommendation), MLHD (Last.fm), and Alibaba-Youku, also satisfy all three conditions, as we verify below using the objectives, statistics, and empirical results reported in Section 5.

C1 follows from M1: Empirical Verification Across Datasets. We prove C1 by establishing that in each dataset, at least one user objective and at least one artist objective exhibit non-zero conditional covariance, driven by the same underlying user-artist interaction event.

Spotify Datasets. Let $U_A = \sum_k \alpha_k \cdot o_{A,k}$ aggregate the four user objectives {*%Listening*, *Num Times Listening*, *Feedback Type*, *In Playlist*}, and let $U_B = \sum_k \beta_k \cdot o_{B,k}$ aggregate the two artist objectives {*New Fan*, *Recurring Fan*} as defined in Section 5. Consider a single listening event (u, a, t) in the Spotify (Next Song Recommendation) dataset.

Whether the user listens to a large fraction of the song (*%Listening* high, contributing to U_A) and whether the artist acquires a new fan ($\mathrm{NewFan}(u,a,t) = 1$, contributing to U_B) are both determined by the user's underlying affinity for the artist. Formally, let $Z_{u,a}$ denote the latent affinity between user u and artist a, which is not directly observed in $\mathbf{X}$ but determines both listening depth and fan conversion probability. Then:

$$\mathrm{Cov}(\%\mathrm{Listening},\ \mathrm{NewFan}(u,a,t) \mid \mathbf{X}) = \mathbb{E}[\mathrm{Cov}(\%\mathrm{Lis.},\ \mathrm{NewFan} \mid \mathbf{X}, Z_{u,a})] + \mathrm{Cov}(\mathbb{E}[\%\mathrm{Lis.} \mid \mathbf{X}, Z_{u,a}],\ \mathbb{E}[\mathrm{NewFan} \mid \mathbf{X}, Z_{u,a}]) \tag{44}$$

The second term in Equation (44) is strictly positive: $\mathbb{E}[\%\mathrm{Listening} \mid \mathbf{X}, Z_{u,a}]$ is increasing in $Z_{u,a}$ (higher affinity leads to more listening), and $\mathbb{E}[\mathrm{NewFan} \mid \mathbf{X}, Z_{u,a}]$ is also increasing in $Z_{u,a}$ (higher affinity makes it more likely the user has not previously encountered the artist, per the exponential decay structure). Therefore $\mathrm{Cov}(U_A, U_B \mid \mathbf{X}) > 0$. This is empirically confirmed by the correlation heatmap of Figure 4a: the off-diagonal blocks between user and artist objectives all exhibit non-zero signed correlations, with *Recurring Fan* negatively correlated with user objectives (reflecting that users who repeatedly listen to an artist already have high U_A but contribute less marginal U_B) and *New Fan* positively correlated with *%Listening*. Condition (C1) is satisfied for both Spotify datasets.

MLHD (Last.fm) Dataset. For the MLHD dataset, the user objectives are $\{$*%Listening*, *IsFinish*, *IsSkip*$\}$ and the artist objectives remain $\{$*New Fan*, *Recurring Fan*$\}$ as defined in Section 5.2.2. The same latent affinity argument applies: $\mathrm{IsFinish} = 1$ (user completed listening, contributing to U_A) and $\mathrm{NewFan}(u,a,t) = 1$ (artist acquired a new fan, contributing to U_B) are jointly determined by the user's affinity for the artist. From Table 14, *IsFinish* has a mean of 0.847 but *IsSkip* has a mean of only 0.020, confirming that completion and skipping are strongly driven by user-artist affinity rather than item-level features alone. The correlation heatmap of Figure 4b confirms that *New Fan* and *Recurring Fan* are positively correlated with *%Listening* and *IsFinish*, satisfying (C1).

Alibaba-Youku Dataset. For Alibaba-Youku, the user objectives are $\{$*Video View*, *Time Spent*, *Play Rate*$\}$ and the artist objectives are $\{$*Relevance*, *Novelty*$\}$ as defined in Section 5. The *Relevance* objective is computed as the inverse Euclidean distance between the recommended video's embedding and the user's previously watched videos, so it is elevated precisely when the user's latent preference aligns with the creator's content style – the same alignment that drives high *Time Spent* and *Play Rate*. From Figure 4c, *Relevance* is positively correlated with *Time Spent* and *Play Rate* (correlation of 0.43 and 0.57, respectively), while *Novelty* is negatively correlated with *Time Spent* (-0.04), reflecting that unfamiliar content drives lower immediate engagement but serves the artist's discovery objective. Both patterns confirm $\mathrm{Cov}(U_A, U_B \mid \mathbf{X}) \neq 0$, satisfying (C1) for this dataset.

C2 follows from M2: Empirical Verification Across Datasets.

We establish that the conditional covariance $\sigma_{AB}(\mathbf{x})$ varies systematically across items in each dataset, using the objective statistics and the prediction results.

Spotify Datasets. From Table 13, the artist objective *New Fan* has a mean of 0.556 and a standard deviation of 0.336 in the Spotify (Next Song Recommendation) dataset, reflecting substantial cross-item heterogeneity in fan acquisition potential. This heterogeneity in U_B is not uniform across items but is systematically structured: for items by artists with a large existing listener base (high $\mathrm{NumTimesListenedBefore}$, $\mathrm{NewFan}(u,a,t) \to 0$ for most users, so $U_B \approx 0$ with near-zero variance and $\sigma_{AB}^{(\text{high-pop})} \approx 0$. For items by artists with sparse listener bases (low $\text{NumTimesListenedBefore}$), both $\text{NewFan}(u, a, t) = 1$ and high *%Listening* are jointly elevated for affinity-matched users, creating $\sigma_{AB}^{(\text{low-pop})} \gg 0$. The prediction results of Table 2 corroborate this directly: the best single-objective RMSE for *Recurring Fan* (0.298) is substantially lower than for *Feedback Type* (0.452), reflecting that long-term fan retention is more predictable from item features alone than immediate user reactions. This asymmetry in predictability implies that σ_{AB} varies in magnitude across item types, as items for which U_B is more variable relative to U_A will have larger covariance. Condition (C2) is satisfied.

MLHD (Last.fm) Dataset. From Table 14, the MLHD user objective *%Listening* has a mean of 0.958 and a standard deviation of only 0.117, indicating that most listening events result in near-complete song consumption on this platform. However, *IsSkip* has a mean of 0.020 and a standard deviation of 0.049, revealing that a small but consequential subset of items is skipped almost universally. The covariance structure between U_A and U_B therefore differs sharply between these "near-universal" items and "skip-prone" items: for the former, U_A is stable and high with near-zero variance, so $\sigma_{AB} \approx 0$ regardless of U_B; for the latter, U_A is highly variable across users with different affinities, producing large σ_{AB} for those items where U_B (artist fan metrics) is also variable. The prediction improvement of MS-Bridge in Table 2 is the largest for *New Fan* (21.58% in RMSE), precisely the objective whose covariance with user objectives is most heterogeneous across items, providing empirical support for C2 in this dataset.

Alibaba-Youku Dataset. From Table 15, the Alibaba-Youku artist objective *Relevance* has a mean of 0.312 but a standard deviation of 0.735, more than twice its mean, while *Novelty* has a mean of 0.538 and a standard deviation of 0.382. These large standard deviations, relative to the user objectives (*Video View* std of 0.443, *Time Spent* std of 32.305), reflect fundamentally different covariance structures across items: videos that are highly relevant to a user's prior viewing history, generate jointly high U_A (high *Time Spent*, *Play Rate*) and high U_B (*Relevance*), producing $\sigma_{AB}^{(\text{relevant})} > 0$; while videos that are novel relative to viewing history produce high U_B (*Novelty*) but potentially lower immediate U_A (as seen in the negative correlation between *Novelty* and *Time Spent* in Figure 4c), producing $\sigma_{AB}^{(\text{novel})} < 0$. The sign reversal in σ_{AB} across item types provides particularly strong evidence for C2 in the Alibaba-Youku setting, and directly explains the largest relative performance improvements of MS-Bridge observed in Table 2c, where the artist objectives *Relevance* and *Novelty* benefit most from cross-side information sharing (4.21% and 5.79% improvement in RMSE, respectively).

C3 follows from M3: Empirical Verification Across Datasets.

Pareto Frontier Evidence. The non-convexity of the achievable utility set $\mathcal{F}$ is directly demonstrated by the Pareto frontier analysis of Figure 11 on the Spotify dataset. For each value of artist importance weight $\alpha \in \{0.1, 0.2, \ldots, 1.0\}$, we plot the (U_A, U_B) pair achieved by our proposed model and the strongest multi-objective baselines (MMoE, MoSE, PLE). The Pareto frontier of MS-Bridge strictly dominates the frontier of all baseline models at every value of α, including at the single-stakeholder endpoints. Crucially, the MS-NSW optimal point (corresponding to $\alpha^* \approx 0.5$) lies strictly inside the region enclosed by the MS-Bridge frontier and is not attained by any baseline regardless of their α setting. This confirms that the MS-NSW optimum is not on the convex hull of the baselines' achievable utility sets, i.e., it lies in a non-convex region of $\mathcal{F}$ that is inaccessible to any linear scalarization of U_A and U_B, satisfying (C3).

MS-NSW Score Evidence. The MS-NSW column in Table 4 provides direct numerical evidence. In the Spotify (Session-Based) dataset, the best linear-scalarization baseline achieves MS-NSW = 0.451 (MORL), while our proposed MS-Bridge achieves MS-NSW = 0.496, resulting in a gap of 9.98%. In the Alibaba-Youku dataset, the best baseline achieves MS-NSW = 0.321 (MORL) while MS-Bridge achieves 0.368, resulting in a gap of 14.64%. Since all baselines optimize weighted sums of U_A and U_B with varying λ, and none can close the gap to MS-Bridge regardless of how λ is tuned (as shown in the weight analysis of Table 5), the MS-NSW optimum is not attainable by any element of Π_{dec}, confirming (C3) in both datasets. This is further corroborated by the weight analysis of Table 5, where even the best-tuned Bayesian Optimization baseline (MS-NSW = 0.447 for Spotify, 0.321 for Alibaba) falls strictly short of MS-Bridge, confirming that the performance gap is not attributable to suboptimal weight selection but to the structural inability of linear scalarization to reach the MS-NSW optimum.

NP-Hardness Connection. The empirical evidence above is consistent with the theoretical argument of Theorem 3: the discrete nature of the recommendation allocation (each user receives a ranked slate of K items from a catalog of 81,948 songs for Spotify, 407,681 songs for MLHD, and 229 videos for Alibaba-Youku, as reported in Table 9) makes $\mathcal{F}$ non-convex in general, precluding the MS-NSW optimum from lying on the convex hull. Condition (C3) is therefore satisfied across all four datasets. □

K. Proof of Theorem 5

First, by applying the descent lemma (Bauschke et al. 2017) on the L-smooth function $\mathcal{U}$, we have the following for any two successive parameter states Θ_t and Θ_{t+1}:

$$\mathcal{U}(\Theta_{t+1}) \geq \mathcal{U}(\Theta_t) + \langle \nabla\mathcal{U}(\Theta_t), \Theta_{t+1} - \Theta_t\rangle - \frac{L}{2}\|\Theta_{t+1} - \Theta_t\|^2$$

Since the parameters are updated following the SGD as $\Theta_{t+1} = \Theta_t + \eta g_t$, we have:

$$\mathcal{U}(\Theta_{t+1}) \geq \mathcal{U}(\Theta_t) + \eta\langle \nabla\mathcal{U}(\Theta_t), g_t\rangle - \frac{L\eta^2}{2}\|g_t\|^2$$

We take the expectation $\mathbb{E}[\cdot]$ conditioned on the filtration of previous iterates. Using the identity $\mathbb{E}[\|g_t\|^2] = \|\nabla\mathcal{U}(\Theta_t)\|^2 + \mathbb{E}[\|g_t - \nabla\mathcal{U}(\Theta_t)\|^2]$, and applying the bounded variance assumption σ^2_{bridge}:

$$\mathbb{E}[\mathcal{U}(\Theta_{t+1})] \geq \mathcal{U}(\Theta_t) + \eta\|\nabla\mathcal{U}(\Theta_t)\|^2 - \frac{L\eta^2}{2}\left(\|\nabla\mathcal{U}(\Theta_t)\|^2 + \sigma^2_{bridge}\right)$$

We then rearrange the terms to isolate the squared gradient norm:

$$\mathbb{E}[\mathcal{U}(\Theta_{t+1})] - \mathcal{U}(\Theta_t) \geq \eta\left(1 - \frac{L\eta}{2}\right)\|\nabla\mathcal{U}(\Theta_t)\|^2 - \frac{L\eta^2\sigma^2_{bridge}}{2}$$

Note that setting the learning rate $\eta \leq \frac{1}{L}$ implies $(1 - \frac{L\eta}{2}) \geq \frac{1}{2}$. We sum the progress across T iterations from $t = 0$ to $t = T - 1$:

$$\mathbb{E}[\mathcal{U}(\Theta_T)] - \mathcal{U}(\Theta_0) \geq \frac{\eta}{2}\sum_{t=0}^{T-1}\|\nabla\mathcal{U}(\Theta_t)\|^2 - \frac{TL\eta^2\sigma^2_{bridge}}{2}$$

Given that $\mathcal{U}(\Theta^*) \geq \mathbb{E}[\mathcal{U}(\Theta_T)]$, we substitute the optimal utility to find the total improvement bound. Finally, by rearranging for the average expected squared gradient norm, we obtain the convergence bound:

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=0}^{T-1}\|\nabla\mathcal{U}(\Theta_t)\|^2\right] \leq \frac{2(\mathcal{U}(\Theta^*) - \mathcal{U}(\Theta_0))}{\eta T} + \eta L\sigma^2_{bridge}$$

The term σ^2_{bridge} can be effectively reduced by placing the bridge at the middle layer (highest Shapley value), where we minimize the gradient conflict for MS-Bridge and accelerate model convergence.

Regarding the longitudinal dynamics of MS-Bridge based on sequential interactions S_t, since the MS-NSW objective is continuous and the action space is compact, the sequence of parameter updates resides within a bounded set. Based on Doob's Martingale Convergence Theorem, the utility sequence U_t converges to a stationary distribution that maximizes the Nash Social Welfare within the interior of the Pareto frontier, effectively constructing the observed "Virtuous Cycle". $\square$