

Multi-objective Learning to Rank by Model Distillation

✎ 著者	Jie Tang (Airbnb)
📅 発行年	2024
🏠 学会名	KDD
🔑 キーワード	Multi-objective optimization Search Ranking learning to rank model distillation
🔗 URL	https://arxiv.org/abs/2407.07181

概要

- オンラインマーケットプレイスにおける検索ランキングは、最重要目的（CVR）だけでなく、副次的な目的（返品、評価、問い合わせ、長期成長）も考慮する必要がある
- この研究ではモデル蒸留を用いたマルチ目的ランキング最適化を提案

研究背景

- オンラインマーケットプレイスでは、短期的な成果（CVR）だけでなく、キャンセルや問い合わせ、既存品を過度に優先することによる長期的成長の阻害などを考慮した、多様な指標の最適化が重要
- 既存アプローチには以下の課題がある
 - モデル融合：目的ごとに別モデルを学習し重み付きでスコアを合成するが、目的間の干渉により最適化が困難
 - スカラーリゼーション：目的関数を加重和で結合し単一の損失として最適化するが、重み調整が困難かつデータ不均衡に弱い
 - マルチタスク学習：目的をタスクとして統合し共有ネットワークで学習するが、目的間に相関がない場合に性能が低下
- また、非微分可能なビジネス目標（例：新規リスティングの優遇）を損失関数に組み込めないという制約もあった

論文の肝

- マルチ目的ランキングをモデル蒸留として定式化した
 - 本来は推論コスト削減のために用いられてきたモデル蒸留だが、その損失関数の構造がスカラリゼーション（加重和最適化）と類似していることに着目
 - これにより、多目的最適化における既存の課題を解決：
 - 不均衡データ問題の緩和
 - パラメータチューニングの簡略化
 - 非微分可能なビジネス目標の取り込み

提案手法

コアアイデア

1. 目的ごとに「教師モデル」を個別に学習

各目的（CVR、キャンセル率など）に対して、その目的だけを最適化するモデルを個別に学習

→ 各目的に最適化されたスコアが得られ、目的間の不均衡にも対応可能

2. 「ソフトラベル」を生成

各教師モデルの出力スコアを重み付き平均して、ソフトラベルを作成：

$$\text{softlabel} = \sum_k \omega_k \cdot f_k^*(X)$$

→ 複数の目的を統合した理想的なランキングスコアを表現

3. 「ハードラベル」と「ソフトラベル」の両方で生徒モデルを学習

生徒モデル $f(\theta, X)$ の損失関数は次のように構成される：

$$\text{Loss} = \alpha \cdot \text{CrossEntropy}(f, \text{hardlabel}) + (1 - \alpha) \cdot \text{CrossEntropy}(f, \text{softlabel})$$

- **hardlabel**：主目的（例：CVR）の正解ラベル
- **softlabel**：複数の目的をまとめた教師モデルの知識
- **α** ：ハードとソフトのバランス（実験では $\alpha=0.2$ が最適）

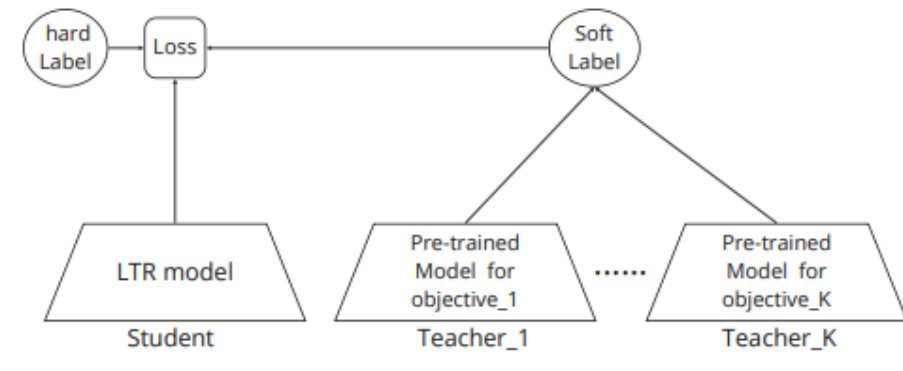


Figure 1: Model training graph

4. 最終的に生徒モデル1本だけをデプロイ

- 学習時に使った教師モデル群は推論時には使わない
- これにより、推論が軽量かつ高速になり、システム運用がシンプルに

拡張

1. 非微分可能なビジネス目標への対応

ソフトラベルに手動でスコア補正を加えることで、例えば「新規リスティングを上位に出す」といったルールもモデルに反映可能

2. 自己蒸留による運用の簡略化

再学習時には前回の生徒モデルを教師とすることで、教師モデル群を再構築せずに済み、再現性と効率性が向上

実験

実験設定

- データ
 - 約3.6億件の生徒モデル用データ
 - 約5億件の教師モデル用データ
 - ラベル：クリック・予約・キャンセル・拒否 など
- 評価指標
 - オフライン：NDCG
 - オンライン：CVR（予約率）＋副目的の変動

- **比較対象（ベースライン）**

- Airbnbで既に導入されていたマルチタスク学習モデル（複数の目的損失を加重平均し、推論時にも目的ごとにスコアを融合）

性能評価

指標	提案手法の改善
NDCG（オフライン）	+1.1%
CVR（オンライン）	+0.37%（ $p=0.02$ ）
推論レイテンシ	-1.6%
学習データ量	1.4億件削減
学習パラメータ数	重み α の1つに簡略化

- **自己蒸留による性能維持**

- 教師モデルを使わずに前回の生徒モデルを教師として学習
- NDCGやCVRがほぼ同等 → 再学習コスト削減

- **再学習時の安定性**

- 初期化を変えて複数回学習し、検索結果の差異を評価
- ベースラインに比べて出力のブレが大幅に減少

- **非微分目標の取り込み**

- スコア補正を推論時 vs 学習時の2パターンで比較
- 学習時にソフトラベルを補正する方が全体性能を保ちやすい

結論

- モデル蒸留の枠組みを活用し、多目的最適化における実運用上の課題（チューニング・不均衡・非微分性）を解決
- オンライン実験でも、CVR・副目的・レイテンシの全てにおいて改善を確認
- モデル運用・再訓練の効率性にも寄与する、実用的な多目的ランキング手法

所感

- 複数目的のランキング最適化という発想自体が新鮮だった
- 目的が相反・非相関なケースにも対応できる点で、モデル蒸留という手段を使うのは非常に理にかなっていると感じた

- ランキングモデルを作成するにあたって，このような複数目的を最適化するモデルがあることを頭に入れて，実装する機会があればやってみたい