Latent Factor Models vs. Collaborative Filtering in Large-Scale Music Recommendation Efficiency

Assignee Research

June 2, 2026

Abstract

This report synthesises findings from 10 peer-reviewed papers addressing the following research question: How does the inference efficiency of latent factor models scale with dataset size in music recommendation compared to collaborative filtering methods when evaluated on metrics like throughput and. Music recommender systems have become central parts of popular streaming platforms such as Last.fm, Pandora, or Spotify to help users find music that fits their preferences. These systems learn from the past listening events of users to recommend music a user will likely listen. 9 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Modeling Artist Preferences of Users with Different Music Consumption Patterns for Fair Music Recommendations. Research question: How does the inference efficiency of latent factor models scale with dataset size in music recommendation compared to collaborative filtering methods when evaluated on metrics like throughput and memory footprint on large-scale benchmarks like Spotify or KKBox?.

2 Methodology

Systematic literature search across multiple databases yielded 10 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.3/10.

3 Results

10 papers retrieved. 9 claims extracted; 0 independently verified. Quality review score: 3.3/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The dataset consists of three user groups: LowMS, MedMS, and HighMS, each with 1,000 distinct users.	×	0.02
The LowMS group has 82,417 distinct artists, 6,915,352 listening events, an average of 239 artists listened to per user,	×	0.05
The MedMS group has 86,249 distinct artists, 7,900,726 listening events, an average of 496 artists listened to per user,	×	0.05
The HighMS group has 92,690 distinct artists, 8,251,022 listening events, an average of 1,194 artists listened to per us	×	0.05
The Base-Level Learning (BLL) equation from the cognitive architecture ACT-R is used to model music listening habits.	×	0.07
The BLL equation accounts for the time-dependent decay of item exposure in human memory.	×	0.03
The BLL equation quantifies the usefulness of a piece of information based on how frequently and how recently it was acc	×	0.04
The BLL equation models time-dependent decay using a power-law distribution.	×	0.03
The BLL equation has been utilized in previous works to recommend tags in social bookmarking systems and to recommend ha	×	0.03

References

- http://arxiv.org/abs/2603.29727v2
- http://arxiv.org/abs/1907.09781v1
- http://arxiv.org/abs/2507.19375v1