



**UFOP**

Universidade Federal  
de Ouro Preto

**Universidade Federal de Ouro Preto  
Instituto de Ciências Exatas e Aplicadas  
Departamento de Computação e Sistemas**

**Modelagem da Curiosidade Social para  
Diversificação e Personalização de  
Recomendações de Músicas Online**

**Luiza Borges Polita**

**TRABALHO DE  
CONCLUSÃO DE CURSO**

**ORIENTAÇÃO:**  
Alexandre Magno de Sousa

**Agosto, 2025  
João Monlevade–MG**

**Luiza Borges Polita**

**Modelagem da Curiosidade Social para  
Diversificação e Personalização de  
Recomendações de Músicas Online**

Orientador: Alexandre Magno de Sousa

Monografia apresentada ao curso de Engenharia da Computação do Instituto de Ciências Exatas e Aplicadas, da Universidade Federal de Ouro Preto, como requisito parcial para aprovação na Disciplina “Trabalho de Conclusão de Curso II”.

**Universidade Federal de Ouro Preto**

**João Monlevade**

**Agosto de 2025**



## FOLHA DE APROVAÇÃO

**Luiza Borges Polita**

### **Modelagem da Curiosidade Social para Diversificação e Personalização de Recomendações de Músicas Online**

Monografia apresentada ao Curso de Engenharia de Computação da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Engenharia de Computação.

Aprovada em 08 de setembro de 2025.

#### Membros da banca

Prof. Dr. Alexandre Magno de Sousa - Orientador (Universidade Federal de Ouro Preto)  
Prof. Dr. Eduardo da Silva Ribeiro (Universidade Federal de Ouro Preto)  
Prof. Dr. Elton Máximo Cardoso (Universidade Federal de Ouro Preto)

Alexandre Magno de Sousa, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 22/09/2025.



Documento assinado eletronicamente por **Alexandre Magno de Sousa, PROFESSOR DE MAGISTERIO SUPERIOR**, em 25/01/2026, às 17:51, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [http://sei.ufop.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **1046956** e o código CRC **1E607CEA**.

# Resumo

Este trabalho investiga como incorporar curiosidade social para diversificar recomendações em consumo de música online, integrando históricos individuais do LFM-1b e a rede de amizades do Last.fm via `user-id` e avaliando em partição temporal (4 semanas de treino seguidas de 4 semanas de teste). Foi proposto um método que combina variáveis colativas (novidade, incerteza, conflito e complexidade) com sinais sociais (influências direta e indireta e novidade social) e reordena candidatos de modo a aproximar o estímulo do usuário de sua zona ótima. O protocolo compara duas estratégias: um ranqueador por frequência e um ranqueador orientado por curiosidade, medidos por precisão (MAP@K) e diversidade entre usuários (IUS@K). Em quatro execuções independentes, o ranqueador por frequência superou o de curiosidade em precisão (MAP), sendo em média  $1,37\times$  melhor no Top-10,  $1,42\times$  no Top-20 e  $1,50\times$  no Top-50 (com variações por run de  $0,90\text{--}2,03\times$ ,  $0,94\text{--}2,08\times$  e  $1,02\text{--}2,13\times$ , respectivamente). Em contrapartida, o ranqueador por curiosidade superou o de frequência em diversidade entre usuários (IUS), alcançando ganhos médios de  $0,69\times$  no Top-10,  $1,95\times$  no Top-20 e  $3,25\times$  no Top-50 (faixas  $0,20\text{--}0,93\times$ ,  $1,73\text{--}2,14\times$  e  $2,89\text{--}3,55\times$ , respectivamente), indicando listas menos homogêneas e maior descoberta conforme o corte do ranking aumenta. Esses resultados evidenciam o trade-off entre precisão imediata e exploração, mostrando que sinais de curiosidade social podem aumentar diversidade com perda controlada de MAP. Como contribuição prática, o estudo entrega um processo reprodutível de fusão histórico-social e um método de ranqueamento por estímulo; como próximos passos, propõe-se um sistema híbrido que una colaborativo, texto e curiosidade social para melhorar equilíbrio entre relevância e descoberta em cenários de teste temporal.

**Palavras-chave:** sistemas de recomendação; curiosidade social; variáveis colativas; teoria da informação; consumo de música online; diversidade de recomendações.

# Abstract

This study investigates how to incorporate social curiosity to diversify recommendations in online music consumption by integrating individual histories from LFM-1b and the Last.fm friendship network via `user-id`, and evaluating under a temporal split (4 training weeks followed by 4 testing weeks). We propose a method that combines collative variables (novelty, uncertainty, conflict, and complexity) with social signals (direct and indirect influences and social novelty) and reorders candidates to bring the user’s stimulus closer to their optimal zone. The protocol compares two strategies—a frequency-based ranker and a curiosity-oriented ranker—measured by precision (MAP@K) and inter-user diversity (IUS@K). Across four independent runs, the frequency ranker outperformed curiosity in precision (MAP), being on average  $1.37\times$  better at Top-10,  $1.42\times$  at Top-20, and  $1.50\times$  at Top-50 (with per-run ranges of  $0.90\text{--}2.03\times$ ,  $0.94\text{--}2.08\times$ , and  $1.02\text{--}2.13\times$ , respectively). Conversely, the curiosity ranker outperformed frequency in inter-user diversity (IUS), achieving average gains of  $0.69\times$  at Top-10,  $1.95\times$  at Top-20, and  $3.25\times$  at Top-50 (ranges  $0.20\text{--}0.93\times$ ,  $1.73\text{--}2.14\times$ , and  $2.89\text{--}3.55\times$ , respectively), indicating less homogeneous lists and greater discovery as the ranking cutoff increases. These results highlight the trade-off between immediate precision and exploration, showing that social curiosity signals can increase diversity with a controlled loss in MAP. As a practical contribution, the study delivers a reproducible user–social fusion process and a stimulus-based re-ranking method; as future work, we propose a hybrid system that combines collaborative filtering, text, and social curiosity to improve the balance between relevance and discovery under temporal testing.

**Keywords:** recommender systems; social curiosity; collative variables; information theory; online music consumption; recommendation diversity.

# Lista de ilustrações

Figura 1 – Curva de Wundt: relação em U invertido entre o potencial de excitação e a resposta hedônica. . . . .	19
Figura 2 – Distribuição acumulada do número de conexões por usuário no dataset Last.fm Social Connections. . . . .	33
Figura 3 – Configuração dos Experimentos. . . . .	34
Figura 4 – Separação de Treino e Teste. . . . .	34
Figura 5 – Função de distribuição acumulada (CDF) da quantidade de eventos por semana. . . . .	35
Figura 6 – Função de distribuição acumulada (CDF) da quantidade de eventos por mês. . . . .	36
Figura 7 – Exemplo da sequência de músicas consideradas para o cálculo dos estímulos. . . . .	39
Figura 8 – Heatmap de Correlação entre métricas. . . . .	46

# Lista de tabelas

Tabela 1 – Comparativo de modelos de curiosidade e limitações. . . . .	26
Tabela 2 – Descrição das principais <i>features</i> do histórico de acessos dos usuários no dataset LFM-1b. . . . .	30
Tabela 3 – Estatísticas do conjunto Last.fm Social Connections. . . . .	31
Tabela 4 – Dados estatísticos sobre usuários e conteúdo. . . . .	32
Tabela 5 – Métricas selecionadas após a análise de redundância (definições conforme as equações do modelo). . . . .	47
Tabela 6 – Descrição das variáveis. . . . .	47
Tabela 7 – Histórico de músicas do usuário. . . . .	48
Tabela 8 – Média das Validações Run 1 (número de usuários por categoria de Beta). . . . .	53
Tabela 9 – Média das Validações Run 2 (número de usuário por categoria de Beta). . . . .	53
Tabela 10 – Média das Validações Run 3 (número de usuários por categoria de Beta). . . . .	53
Tabela 11 – Média das Validações Run 4 (número de usuários por categoria de Beta). . . . .	54
Tabela 12 – Média da métrica MAP por run para o método de Curiosidade. . . . .	54
Tabela 13 – Média da métrica MAP por run para o método de Frequência. . . . .	54
Tabela 14 – MAP — quantas vezes o ranking de Frequência é melhor que o de Curiosidade. . . . .	55
Tabela 15 – Média, desvio padrão e intervalo de confiança de 95% da métrica MAP para os métodos Curiosidade e Frequência. . . . .	55
Tabela 18 – IUS: quantas vezes o ranking de Curiosidade é melhor que o de Frequência. . . . .	56
Tabela 16 – Média da métrica IUS por run para o método de Curiosidade. . . . .	56
Tabela 17 – Média da métrica IUS por run para o método de Frequência. . . . .	56
Tabela 19 – Média, desvio padrão e intervalo de confiança de 95% (IC95%) da métrica IUS para os métodos Curiosidade e Frequência. . . . .	57
Tabela 20 – Ganho da Frequência em relação à Curiosidade na métrica MAP. . . . .	61
Tabela 21 – Ganho da Curiosidade em relação à Frequência na métrica IUS. . . . .	61
Tabela 22 – Resultados completos da métrica MAP para o método Curiosidade (por Run e recomendação) . . . . .	69
Tabela 23 – Resultados completos da métrica MAP para o método Frequência (por Run e recomendação) . . . . .	70
Tabela 24 – Resultados completos da métrica IUS para o método Curiosidade (por Run e recomendação) . . . . .	71
Tabela 25 – Resultados completos da métrica IUS para o método Frequência (por Run e recomendação) . . . . .	72

# Lista de abreviaturas e siglas

API	Interface de Programação de Aplicações
B2C	Business to Consumer
CDF	Função de distribuição acumulada
CV	Collative Variables
IC	Intervalo de Confiança
IC95%	Intervalo de confiança de 95%
IUS	Inter-User Similarity
LFM-1B	Conjunto de dados de eventos de escuta do Last.fm
MAP	Mean Average Precision
SI	Stimulus Indicator (modelo de Fu & Niu, 2023)
UAM	Matriz usuário–artista (User–Artist Matrix)

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>9</b>
1.1	Motivação e Justificativa	10
1.2	Definição do Problema	11
1.3	Objetivos Geral e Específicos	12
1.4	Resultados e Contribuições	12
1.5	Estrutura e Organização da Monografia	14
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>15</b>
2.1	Curiosidade	15
2.1.1	Variáveis Colativas	16
2.1.2	Curva de Wundt	17
2.1.3	Curiosidade Social	19
2.2	Modelos de Curiosidade na Computação	20
2.3	Sistemas de Recomendação	22
2.4	Trabalhos relacionados	24
2.5	Considerações Finais	27
<b>3</b>	<b>METODOLOGIA</b>	<b>28</b>
3.1	Coleção de Dados	28
3.1.1	LFM-1B	29
3.1.2	Last.fm Social Connections	30
3.2	Integração do Histórico de Acessos do Usuário com as Conexões de Amizade	31
3.3	Configuração dos Experimentos	33
3.3.1	Seleção Aleatória de Dias	34
3.3.2	Dias válidos	34
3.3.3	Número mínimo de músicas	35
3.4	Premissas da Curiosidade	36
3.5	Métricas de curiosidade social	39
3.5.1	Curiosidade social direta	40
3.5.2	Curiosidade social indireta	41
3.5.3	Métricas de grupo	41
3.6	Variáveis colativas tradicionais	42
3.6.1	Novidade	43
3.6.2	Incerteza	43
3.6.3	Conflito	44

3.6.4	Complexidade . . . . .	44
3.7	Processo de seleção de métricas de curiosidade . . . . .	45
3.8	Método de Avaliação de Ranking . . . . .	46
3.9	Métrica Inter-User Similarity (IUS) . . . . .	49
3.10	Considerações Finais . . . . .	50
4	<b>RESULTADOS . . . . .</b>	<b>52</b>
4.1	Considerações Iniciais . . . . .	52
4.2	Validações das Curvas . . . . .	53
4.3	Resultados da Métrica MAP . . . . .	54
4.4	Resultados da Métrica IUS . . . . .	56
4.5	Considerações Finais . . . . .	57
5	<b>CONCLUSÃO . . . . .</b>	<b>59</b>
5.1	Limitações do Trabalho . . . . .	61
5.2	Principais Contribuições . . . . .	62
5.3	Trabalhos Futuros . . . . .	63
	<b>REFERÊNCIAS . . . . .</b>	<b>64</b>
	<b>APÊNDICES . . . . .</b>	<b>68</b>
	<b>APÊNDICE A – RESULTADOS DAS MÉTRICAS . . . . .</b>	<b>69</b>
A.1	Métrica MAP por Run . . . . .	69
A.2	Métrica IUS por Run . . . . .	71

# 1 Introdução

Sistemas de recomendação tornaram-se parte central de plataformas digitais, auxiliando usuários a “navegar grandes espaços de itens” e a tomar decisões com menos esforço cognitivo (JANNACH et al., 2010). Na última década, redes neurais profundas avançaram o estado da arte, com arquiteturas “híbridas” e abordagens baseadas em sequências elevando a qualidade preditiva em diferentes domínios (GRIDACH, 2020; MOREIRA; FERREIRA; CUNHA, 2018). Ainda assim, o cenário online impõe desafios recorrentes, como *sparsity* de dados, perfis de usuário pouco densos e rápida evolução do catálogo, o que exige modelos robustos e sensíveis ao contexto (GRIDACH, 2020).

Além da acurácia, cresce o interesse por utilidades de descoberta (diversidade, novidade e serendipidade) (JANNACH et al., 2010) pois recomendações excessivamente “ótimas” podem reduzir exploração e valor de longo prazo. Nessa área, a curiosidade aparece como uma solução para busca por informação: de um lado, a psicologia clássica descreve variáveis colativas (*novelty, uncertainty, conflict, complexity*) que modulam interesse e de outro, surgem formulações computacionais que incorporam tais sinais ao processo de recomendação (WU; MIAO, 2013). Outro método estudado na área da curiosidade é a Curva de Wundt que sugere uma relação em “U invertido” entre intensidade do estímulo e resposta hedônica, na qual níveis moderados tendem a maximizar engajamento (WUNDT, 1874).

Especificamente em Sistemas de Recomendação, mecanismos *curiosity-driven* foram explorados para equilibrar relevância e “novidade”, personalizando o quanto de novidade é útil “dependendo da curiosidade do usuário” (ZHOU et al., 2020). Abordagens recentes propõem um **stimulus-evoked curiosity mechanism** para ajustar o ranking por meio de sinais de estímulo e resposta, preservando precisão enquanto promovem descoberta (XU et al., 2021). Em ambientes musicais e de mídia social, estudos apontam que a curiosidade é mensurável no tempo e se relaciona com padrões de consumo e propagação de informação (SOUSA; ALMEIDA; FIGUEIREDO, 2019).

Este trabalho investiga a curiosidade social como utilidade de descoberta em recomendações musicais juntamente com as variáveis colativas (*novelty, uncertainty, conflict e complexity*) para descobrir recomendações que elevem interesse e engajamento (SOUSA; ALMEIDA; FIGUEIREDO, 2022). A proposta usa recomendação guiada pela curiosidade e regula a novidade por perfil para atingir o estímulo ideal da Curva de Wundt (ZHAO; LEE, 2016).

Com base nesses fundamentos, definimos um conjunto de métricas “estímulo” e “resposta” de curiosidade social no ranking (por exemplo, medidas relacionadas a incerteza e

novidade informacional) integrando-as ao processo de recomendação. O objetivo é investigar se mecanismos de curiosidade por estímulo podem “preservar precisão” ao mesmo tempo em que elevam “diversidade” e descoberta em listas musicais, aproximando-se de um ponto de equilíbrio entre exploração e aproveitamento.

Por fim, o trabalho adota metodologias conhecidas da área de Sistemas de Recomendação, com avaliação baseada em métricas de acurácia e utilidades de descoberta. Assim, o estudo contribui para a linha que incorpora “curiosidade” como princípio de projeto em recomendações.

## 1.1 Motivação e Justificativa

A curiosidade tem sido constantemente reconhecida como um fator importante que influencia o comportamento humano em todos os estágios do ciclo de vida (BERLYNE, 1960), tanto em aspectos positivos como negativos. Na área da psicologia, Berlyne propõe que a curiosidade é provocada por diferentes tipos de estímulos e que a mesma pode ser mantida ou reduzida pela exposição contínua a esses estímulos (LOEWENSTEIN, 1994). Esse autor também afirma que a curiosidade pode ser mensurada por meio de um conjunto de quatro características que foram denominadas variáveis colativas: novidade, incerteza, conflito e complexidade. De acordo com Berlyne essas variáveis colativas estão diretamente relacionadas à intensidade do estímulo (SHRESTHA et al., 2020).

O desenvolvimento de pesquisas em curiosidade é uma forma de construção de conhecimentos que também podem ser utilizados fora do campo teórico da academia, principalmente em ambientes corporativos (SHANKAR, 2018). Logo, o resultado dessas pesquisas sobre curiosidade têm implicações econômicas diretas, tais como: (a) no *marketing*; (b) em áreas com aplicação da inteligência artificial em serviços online e robótica; (c) no mercado de trabalho; e também (d) no capital humano (BUYALSKAYA; CAMERER, 2020). Além disso, com a rápida disseminação e avanços tecnológicos dos serviços de vendas online na Internet, o comércio eletrônico está crescendo cada vez mais. Nos últimos anos, surgiram vários sites populares de comércio eletrônico, dentre eles, podem ser citados a Amazon, o eBay, o TaoBao (na China) no campo *Business to Consumer* (B2C), o Kompass, o Thomasnet e Made-in-Chin 1 (também da China) no campo *Business to Business*. Em sistemas de e-commerce, é natural realizar a busca por clientes que expressam preferências similares a outros simplesmente para aumentar o engajamento com novos produtos (ZHANG; SHENG-HUA, 2007). No entanto, à medida que a sobrecarga de informações sobre produtos se torna cada vez mais acentuada, é muito difícil encontrar os produtos que os usuários realmente precisam (DONG et al., 2013). Diante desse cenário, surgem os sistemas de recomendação e a ideia por trás desses sistemas é recomendar itens que possam alcançar mais o perfil dos usuários.

Atualmente, com o objetivo das plataformas de serviços digitais de capturar cada vez mais a atenção de usuários, os sistemas de recomendação online têm grande influência na economia (JANNACH et al., 2010). Portanto, é por essa razão que observa-se um aumento constante na condução de pesquisas que buscam desenvolver novas abordagens para antecipar com maior precisão os itens aos quais os usuários dedicarão maior atenção e envolvimento. Nesse ponto, a curiosidade humana tem um papel fundamental no surgimento de novas propostas de tecnologia para melhorar os serviços de recomendação (JANNACH et al., 2010).

Existem diferentes modelos de curiosidade computacionais como os mencionados no trabalho de (WU; MIAO, 2013). Esses modelos podem ser classificados de acordo com o tipo de variável colativa utilizada, tais como os que são baseados em novidade, surpresa, mudança, incerteza, conflito e complexidade. A novidade, por exemplo, é inversamente proporcional à frequência, recência e dissimilaridade com que o usuário tem acessado itens em seu histórico de acessos. Por outro lado, a surpresa pode ser interpretada como a diferença entre a expectativa do usuário e aquilo que realmente acontece quando ele finalmente recebe o estímulo de curiosidade. A mudança faz com que uma prioridade seja dada na direção do comportamento exploratório da curiosidade do indivíduo. Por sua vez, a incerteza está relacionada ao quão clara é a resposta de um indivíduo ao estímulo recebido. O conflito, por definição, está diretamente ligado à ocorrência de duas ou mais respostas incompatíveis que um usuário pode ter quando ele recebe estímulos de curiosidade. Por fim, a complexidade diz respeito à capacidade que o indivíduo tem de compreender as informações recebidas por intermédio de um estímulo. Conforme especificado em (WU; MIAO, 2013), existem várias áreas da computação que aplicam diferentes modelos de curiosidade: (i) *Machine Learning*, (ii) robótica, (iii) criatividade artificial, (iv) jogos, (v) educação e (vi) sistemas de informação online.

## 1.2 Definição do Problema

Atualmente, os sistemas de recomendação utilizam das similaridades entre usuários e itens para recomendar itens aos usuários, porém, considerar apenas esse aspecto pode fazer com que o usuário perca o interesse devido à exposição contínua que esse tipo de recomendação pode trazer. Para tanto, a hipótese de trabalho é: de maneira a curiosidade social é eficaz para diversificar e personalizar as recomendações de músicas?

Levando em consideração o aspecto da curiosidade na recomendação, cada usuário pode ser classificado de acordo com os níveis de estímulos de curiosidade que ele recebe. Por exemplo, um usuário que possui menos curiosidade pode receber itens mais semelhantes aos já acessados, em contrapartida, um usuário mais curioso possui a tendência de aceitar itens diferentes do que ele costuma consumir e, por isso, receberá recomendações mais

diversificadas. Isso pode fazer com que o usuário seja tratado de forma mais individual e personalizada, conseqüentemente o usuário pode dar mais atenção aos itens recomendados.

Diante desse contexto, também pode ser adicionado ao modelo de recomendação, além das variáveis colativas tradicionais, a curiosidade social que mensura a quantidade de estímulo que vem de conteúdos que suas conexões sociais consomem. Portanto, a curiosidade social afeta o usuário de forma positiva levando-o a se interessar por itens que suas conexões sociais têm interesse. No trabalho de (SOUSA; ALMEIDA; FIGUEIREDO, 2022), é proposto uma forma de se quantificar a curiosidade social como um estímulo que leva o usuário a tomar decisões que afetam seu comportamento na plataforma alvo. Além disso, esse trabalho também mostra como os usuários são caracterizados em termos de estímulos da curiosidade social mostrando a importância da influência social.

Nesse cenário, é importante destacar que a incorporação de modelos de curiosidade em algoritmos de recomendação melhoram a personalização dos itens recomendados bem como sua diversidade (ZHOU et al., 2020; XU et al., 2021). Isso evidencia a importância da necessidade de uma investigação mais aprofundada e mais ampla sobre a interação entre as variáveis colativas tradicionais e a curiosidade social na recomendação (SOUSA; ALMEIDA; FIGUEIREDO, 2022).

### 1.3 Objetivos Geral e Específicos

O objetivo geral deste trabalho é criar um modelo que incorpora a curiosidade social para aprimorar a diversidade das recomendações de músicas. Para tanto, uma plataforma de consumo de música online será utilizada para este projeto, a saber o Last.fm. Dessa forma, pretende-se investigar a compreensão de como as interações sociais influenciam a curiosidade dos usuários no consumo de músicas online. Para alcançar o objetivo geral, os seguintes objetivos específicos são definidos:

1. Criar um modelo de curiosidade social para a plataforma de músicas online do Last.fm;
2. Avaliar o algoritmo do modelo de curiosidade baseando-se na dimensão e no aspecto da diversidade.

### 1.4 Resultados e Contribuições

Os resultados obtidos neste trabalho permitiram compreender de maneira mais aprofundada o papel da curiosidade social em sistemas de recomendação musical. De modo geral, os experimentos mostraram que a inclusão da dimensão social realizada pelo *merge*

entre o histórico de acessos do usuário-alvo e de seus amigos contribui para enriquecer a diversidade das recomendações e oferecer listas mais personalizadas.

Em termos quantitativos, foi possível comparar dois métodos de recomendação distintos: um baseado na **frequência** de consumo e outro orientado pela **curiosidade**. As análises com a métrica *Mean Average Precision* (MAP) evidenciaram que o método de frequência tende a obter valores mais altos de precisão imediata, por privilegiar itens populares ou já conhecidos pelos usuários. Por outro lado, os resultados com a métrica *Inter-User Similarity* (IUS) demonstraram que o método de curiosidade apresenta maior diversidade e menor sobreposição entre listas de diferentes usuários, reforçando seu potencial para recomendações mais exploratórias e menos homogêneas.

Na métrica MAP (maior é melhor), as médias agregadas foram: frequência = 0,4058 (Top-10), 0,4104 (Top-20) e 0,4117 (Top-50); curiosidade = 0,1723 (Top-10), 0,1704 (Top-20) e 0,1649 (Top-50). Os intervalos de confiança de 95% não se sobrepueram entre os métodos nesses cortes, indicando superioridade consistente da estratégia por frequência. Considerando quatro runs, a frequência superou a curiosidade em 4/4 casos no MAP.

Na métrica IUS (menor é melhor), a curiosidade apresentou valores médios em torno de 0,0002, enquanto a frequência alcançou até cerca de 0,0009, com intervalos de confiança de 95% também sem sobreposição. Em quatro runs, a curiosidade superou a frequência em 4/4 casos na IUS, com efeito mais pronunciado para listas maiores (Top-50).

Esses resultados numéricos confirmam o padrão de trade-off: frequência otimiza precisão de ranking (MAP), ao passo que curiosidade melhora diversidade e personalização (IUS), com superioridade estatística consistente e repetibilidade ao longo das execuções.

Entre as principais contribuições deste trabalho, destacam-se:

- A adaptação do modelo de curiosidade social ao contexto do Last.fm, integrando tanto as premissas individuais quanto os aspectos sociais no consumo musical;
- A implementação de um processo de integração entre históricos de usuários e de seus amigos, permitindo analisar de forma prática a influência social no comportamento de escuta;
- A avaliação empírica comparativa entre recomendações baseadas em frequência e recomendações orientadas pela curiosidade, utilizando métricas complementares (MAP e IUS);
- A demonstração de que a curiosidade, ainda que menos eficaz em termos de precisão imediata, contribui significativamente para aumentar a diversidade e personalização das recomendações, apontando caminhos para sistemas que priorizem engajamento a longo prazo.

Assim, este trabalho avança no entendimento de como a curiosidade social pode ser formalizada e aplicada a sistemas de recomendação, oferecendo resultados que podem servir de base para pesquisas futuras em recomendação personalizada e orientada à descoberta.

## 1.5 Estrutura e Organização da Monografia

Esta monografia está organizada em cinco capítulos, além dos apêndices. O Capítulo 1 (Introdução) apresenta a motivação e a justificativa do estudo, a definição do problema, os objetivos geral e específicos, bem como os principais resultados e contribuições esperados, encerrando com uma visão geral sobre a estrutura do trabalho. O Capítulo 2 (Revisão da Literatura) reúne os fundamentos teóricos que embasam o estudo, discutindo os conceitos de curiosidade e suas variáveis colativas (novidade, incerteza, conflito e complexidade), a curva de Wundt e a noção de curiosidade social; também são apresentados modelos de curiosidade aplicados à computação, os principais sistemas de recomendação e os trabalhos relacionados, seguidos de considerações que conectam a revisão à proposta deste trabalho. O Capítulo 3 (Metodologia) detalha os conjuntos de dados utilizados, incluindo LFM-1B e Last.fm Social Connections; explica a integração entre histórico de acessos dos usuários e de suas conexões de amizade; descreve a configuração dos experimentos (seleção aleatória de dias, definição de dias válidos e número mínimo de músicas); apresenta as premissas do modelo de curiosidade, as métricas de curiosidade social e as métricas das variáveis colativas tradicionais, o processo de seleção de métricas e o método de avaliação de ranking; por fim, introduz a métrica Inter-User Similarity (IUS) e traz considerações finais da metodologia. O Capítulo 4 (Resultados) apresenta e discute os experimentos realizados, iniciando pelas considerações iniciais, passando pela validação das curvas e, em seguida, pelos resultados das métricas MAP e IUS; evidencia-se a comparação entre recomendação por frequência e por curiosidade, destacando diferenças em precisão e diversidade, e encerrando com considerações gerais sobre os achados. O Capítulo 5 (Conclusão) sintetiza os principais resultados alcançados e discute as limitações do trabalho, como o uso de um conjunto de dados antigo e a premissa de influência social homogênea; em seguida, apresentam-se as contribuições desta pesquisa e sugestões de trabalhos futuros que podem expandir e aprofundar a abordagem proposta.

Ao final, incluem-se os Apêndices, que trazem informações complementares, incluindo resultados completos das métricas MAP e IUS por runs.

## 2 Revisão da Literatura

A revisão da literatura tem por objetivo estabelecer os fundamentos conceituais e empíricos que sustentam este trabalho, situando-o no estado da arte de curiosidade (humana e computacional) e de sistemas de recomendação. É iniciada da noção clássica de curiosidade enquanto mecanismo motivacional ligado à busca por novidade, redução de incerteza e resolução de conflito, discutindo sua evolução histórica das leituras filosófico-morais às formulações psicológicas e informacionais mais recentes. Em seguida, examinamos como esses conceitos foram operacionalizados em modelos computacionais e incorporados a técnicas de recomendação, com ênfase em métricas derivadas da teoria da informação e em abordagens que exploram sinais sociais.

Estruturalmente, este capítulo está organizado da seguinte forma. Na Seção 2.1, é apresentado os conceitos de curiosidade, detalhando as *variáveis colativas* (novidade, incerteza, conflito e complexidade) e a Curva de Wundt, que descreve a relação não linear entre estímulo e resposta. Na sequência, a Seção 2.2 trata de modelos de curiosidade na computação e categoriza cada um conforme o tipo de estímulo e o princípio do potencial intermediário de excitação. A Seção 2.3 trata de Sistemas de Recomendação, revisando métodos, taxonomias e implicações práticas. A Seção 2.4 apresenta os trabalhos relacionados onde são apresentados estudos representativos, suas contribuições e limitações, preparando o terreno para as escolhas metodológicas do capítulo seguinte. Por fim, a Seção 2.5 apresenta as considerações finais do capítulo.

### 2.1 Curiosidade

A curiosidade é uma característica que pode ser encontrada em diferentes seres vivos e o estudo dela ao longo dos anos tem se tornado cada vez mais importante, pois o entendimento dela pode trazer benefícios a vários campos diferentes como: (a) no *marketing*; (b) em áreas com aplicação da inteligência artificial em serviços online e robótica; (c) no mercado de trabalho; e também (d) no capital humano (BUYALSKAYA; CAMERER, 2020).

Nos dias atuais, com o crescimento de serviços distribuídos por meio da internet, a curiosidade desempenha um papel cada vez mais importante na economia da era da informação, especialmente tendo em conta que grande parte da tecnologia moderna é voltada para atrair, manter e redirecionar a atenção do usuário. A curiosidade também é inquestionavelmente uma das forças dominantes que impulsionam muitas fontes de crescimento econômico, como a descoberta científica e o nível de escolaridade (WOJTOWICZ; LOEWENSTEIN, 2020).

No início a curiosidade foi definida por narrativas religiosas do cristianismo e da mitologia pagã. Narrativas da Queda do Homem retratam Eva como a mulher curiosa que ao ultrapassar os limites dados por Deus condena a humanidade. Representações da curiosidade como a Torre de Babel, a Caixa de Pandora, as Sereias ou os Pilares de Hércules constituem metáforas através das quais vários conceitos de curiosidade eram transmitidos (BROEDERS, 2021).

Logo depois, a curiosidade foi estudada no campo da filosofia, onde ela é vista como uma virtude que auxilia o progresso da sociedade. Daston observou que a história da curiosidade tem relação com outras ideias modernas relacionadas ao conhecimento e ao desejo de conhecimento, como “maravilha” ou “terror”, mas também “interesse, estranheza, sutileza, segredo e raridade” (DASTON, 1998).

Portanto, as primeiras discussões sobre a curiosidade, anteriores ao surgimento do campo da psicologia, foram conduzidas por filósofos e pensadores religiosos e fundamentadas na questão do status moral da curiosidade, e não em seus fundamentos psicológicos (LOEWENSTEIN, 1994).

Com o passar dos anos a curiosidade chegou ao campo da psicologia onde Berlyne propõe que ela é provocada por diferentes tipos de estímulos e que a mesma pode ser mantida ou reduzida pela exposição contínua a esses estímulos (LOEWENSTEIN, 1994).

Esse autor também afirma que a curiosidade pode ser mensurada por meio de um conjunto de quatro características que foram denominadas variáveis colativas: novidade, incerteza, conflito e complexidade. De acordo com Berlyne essas variáveis colativas estão diretamente relacionadas à intensidade do estímulo (BERLYNE, 1960; SHRESTHA et al., 2020).

### 2.1.1 Variáveis Colativas

A força de um estímulo é quantificada pelo grau de estímulo, do inglês *stimulus degree*, que é definido por um número de fatores chamados variáveis colativas (ZHAO; LEE, 2016). Segundo Berlyne (BERLYNE, 1960) existem quatro colativas: novidade, incerteza, conflito e complexidade. A avaliação de cada variável implica em uma análise de semelhanças e diferenças entre elementos de um padrão de estímulo (WU; MIAO, 2013).

A novidade como o nome já diz, refere-se a algo novo. Esse novo é dado sobre o tempo que o estímulo aconteceu. Para medir a novidade podemos usar três critérios: a novidade está inversamente relacionada com (1) a frequência com que os estímulos foram experimentados anteriormente, (2) a recente experiência dos estímulos e (3) a semelhança entre os estímulos a experiências anteriores. Novidade é frequentemente acompanhada por outras propriedades e cada uma delas tem diferentes influências nos comportamentos exploratórios. Berlyne denomina elas de variáveis suplementares a novidade: mudança,

surpresa e incongruência (BERLYNE, 1960).

A incerteza aparece quando um organismo tem dificuldade de selecionar uma resposta a um estímulo. Berlyne (BERLYNE, 1960) propôs medir a incerteza causada por um estímulo da seguinte forma: (1) elaborar uma lista de estímulos que podem ocorrer (como resposta ao estímulo em questão), (2) separar elas em classes e (3) atribuir uma probabilidade a cada classe. A probabilidade  $p_i$  de cada classe denota a força competitiva de cada resposta possível, e a entropia de Shannon  $H = -\sum_{i=1}^n (p_i \times \log_2(p_i))$  denota o grau de incerteza.

O conflito ocorre quando um estímulo levanta dois ou mais respostas incompatíveis em um organismo. Segundo Berlyne (BERLYNE, 1960) o conflito está positivamente relacionado com (1) a proximidade da igualdade nas forças das respostas concorrentes, (2) a absoluta pontos fortes de respostas concorrentes, (3) o número de respostas concorrentes e (4) o graus de incompatibilidade entre respostas concorrentes.

Já a complexidade refere-se a variedade ou diversidade em um padrão de estímulo. Ela pode ser mensurada por três propriedades que são (1) o número de elementos distinguíveis em um estímulo, (2) a diferença entre esses elementos, e (3) o grau em que vários elementos são percebidos e respondidos como uma unidade (WU; MIAO, 2013).

### 2.1.2 Curva de Wundt

Wundt é uma curva utilizada na psicologia para modelar a curiosidade humana. A curva de Wundt se baseia em uma forma de U inversa que relaciona a estimulação de um item (sd) e a resposta do organismo a esse item (BERLYNE, 1960) (ZHAO; LEE, 2016). Baseado na teoria de Wundt, Berlyne criou uma teoria de “potencial intermediário de excitação”, onde uma estimulação pequena resulta em tédio e muito estímulo resulta em ansiedade e apenas a estimulação intermediária resulta em curiosidade (BERLYNE, 1960; WU; MIAO, 2013).

A **curva de Wundt** descreve a relação entre a *intensidade do estímulo* e a *resposta hedônica* do indivíduo (prazer/avaliação do estímulo), assumindo um perfil em *U invertido*. Em níveis muito baixos de estímulo, o organismo tende à **relaxação/tédio**; em níveis muito altos, emerge **ansiedade/sobrecarga**; entre esses extremos, um intervalo *intermediário* promove **curiosidade** e comportamentos exploratórios (BERLYNE, 1960; ZHAO; LEE, 2016; SILVIA, 2008).

No eixo horizontal, considera-se a intensidade global do estímulo (quão novo, complexo, incerto ou conflituoso ele é); no eixo vertical, a resposta subjetiva (prazer/curiosidade). A forma em U invertido indica três regiões operacionais: (i) **tédio** (subestimulação), (ii) **música ótima** (curiosidade máxima) e (iii) **ansiedade** (superestimulação). Os *limiares* que separam essas regiões *não são fixos*: variam **entre indivíduos**

e **ao longo do tempo** para um mesmo indivíduo, influenciados por experiência prévia, contexto e traços pessoais (BERLYNE, 1960; SILVIA, 2008).

Na tradição de Berlyne, a intensidade do estímulo resulta da combinação de *variáveis colativas* como **novidade** (recência, frequência e dissimilaridade), **incerteza** (dificuldade de decidir a resposta), **conflito** (forças de respostas incompatíveis) e **complexidade** (diversidade estrutural). Cada uma dessas dimensões pode empurrar o estímulo para fora da zona ótima (por falta ou excesso), alterando a posição do ponto de máximo da curva para cada usuário e contexto (BERLYNE, 1960).

Em sistemas de recomendação, mapear itens para a escala de estimulação e *calibrar* a entrega próxima ao nível ótimo favorece exploração sem sacrificar totalmente a relevância. Trabalhos operacionais estimam a curva com diferentes parametrizações (p. ex., distribuições *Beta* ou *Normal*, ou ajustes não paramétricos via histograma/sigmóide) para transformar métricas de novidade/afins em *escores de curiosidade* usados na seleção e ordenação de itens (ZHAO; LEE, 2016; XU et al., 2021; BUYALSKAYA; CAMERER, 2020; ABBAS; NIU, 2019). Em linhas gerais, a curva explica o *trade-off* observado entre **precisão** e **diversidade**: aumentar a estimulação além do confortável tende a ampliar descoberta, mas pode reduzir acertos imediatos.

A curva de Wundt provê um *modelo unificador* para entender como múltiplas características do estímulo se combinam e, de forma *não linear*, modulam a curiosidade. Seu ponto central é que a curiosidade máxima ocorre em estímulo moderado, com limiares dependentes da pessoa e do contexto; por isso, estimativas *personalizadas* da curva são preferíveis em aplicações reais.

Na Figura 1, o eixo horizontal representa o *potencial de excitação* (isto é, quão estimulante é o item), e o eixo vertical representa a avaliação hedônica (*pleasure*). Conforme ilustrado na Figura 1, distinguem-se três zonas: (i) **tédio** (à esquerda), em que a novidade/complexidade é baixa e o item mal engaja o usuário; (ii) **música ótima** (*optimal point*), região central em que a estimulação é suficiente para gerar curiosidade e exploração sem causar desconforto; e (iii) **ansiedade** (à direita), quando o excesso de novidade, incerteza, conflito ou complexidade afasta o usuário (BERLYNE, 1960). Os marcadores na Figura 1 exemplificam, respectivamente, um item no *rising slope* (ainda ganhando interesse) e outro no *falling slope* (já excessivo para aquele usuário/contexto).

Para operacionalizar a Figura 1 em recomendações, a literatura estima a curva por modelos paramétricos (por exemplo, distribuições *Beta* ou *Normal*) ou por formas não paramétricas (ajustes por histograma e funções sigmóides), mapeando métricas de novidade/afins em *escores de curiosidade* que guiam a ordenação dos itens (ZHAO; LEE, 2016; XU et al., 2021; SHRESTHA et al., 2020; ABBAS; NIU, 2019). Em termos práticos, mirar a região próxima ao *ponto ótimo* explica o conhecido *trade-off* entre **exploração/diversidade** e **precisão**: aumentar a estimulação além do confortável tende

a ampliar descoberta, mas pode reduzir o acerto imediato.

A curva de Wundt oferece um unificador para entender por que o mesmo item pode ser *insuficiente* ou *exagerado demais* para pessoas distintas (ou para a mesma pessoa em momentos diferentes). Em aplicações reais, estimativas *personalizadas* da curva são desejáveis, pois alinham a recomendação à música de estímulo que maximiza curiosidade e engajamento do usuário.

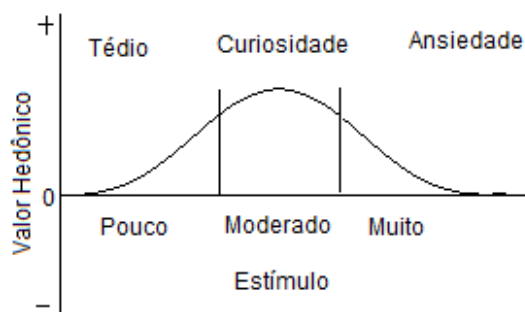


Figura 1 – Curva de Wundt: relação em U invertido entre o potencial de excitação e a resposta hedônica.

### 2.1.3 Curiosidade Social

A curiosidade tem sido considerada um traço fundamental da cognição humana, responsável por motivar processos de aprendizagem, tomada de decisão e desenvolvimento saudável (BERLYNE, 1960; LOEWENSTEIN, 1994). Tradicionalmente, o estudo da curiosidade tem se apoiado em variáveis colativas como novidade, complexidade, conflito e incerteza que descrevem os estímulos externos responsáveis por desencadear esse comportamento (BERLYNE, 1960; SILVIA, 2008). No entanto, em ambientes de interação social, um componente adicional emerge como fator determinante: a influência social. Esse aspecto tem sido tratado na literatura recente como *curiosidade social* (RENNER, 2006; RENNER; SCHÜTZ, 2012).

A curiosidade social pode ser definida como o interesse em obter novas informações sobre os outros, como pensam, agem ou se comportam, motivando condutas exploratórias (RENNER, 2006). Essa dimensão distingue-se de outras por estar diretamente associada às interações interpessoais e à observação do comportamento alheio. Estudos recentes sugerem ainda duas formas distintas: a curiosidade social aberta (*overt*), caracterizada por interações explícitas e relacionadas a competências interpessoais, e a curiosidade social encoberta (*covert*), associada a comportamentos mais discretos, como fofocas ou observações indiretas, podendo estar relacionada a desfechos menos saudáveis, como ansiedade social (RENNER; SCHÜTZ, 2012).

No contexto das redes sociais digitais, a curiosidade social adquire relevância especial, uma vez que grande parte das interações é mediada por plataformas que amplificam tanto os estímulos sociais quanto os canais de difusão de informação. De acordo com (XU et al., 2018), incorporar a influência social a modelos de curiosidade permite explicar melhor padrões de consumo e compartilhamento de informação, indo além das dimensões clássicas propostas por Berlyne (BERLYNE, 1960). Assim, a curiosidade social pode ser entendida como uma variável colativa adicional que, em conjunto com novidade, incerteza, conflito e complexidade, contribui para explicar os fatores que estimulam a curiosidade em ambientes digitais.

Especificamente no WhatsApp, plataforma com mais de dois bilhões de usuários, a dinâmica de grupos cria um ambiente fértil para a manifestação da curiosidade social (SOUSA; ALMEIDA; FIGUEIREDO, 2022). Nesses espaços, a decisão de um usuário compartilhar uma mensagem pode ser influenciada não apenas pelo conteúdo em si, mas também pelo comportamento dos demais membros do grupo. Dessa forma, a curiosidade social emerge como força motriz por trás da disseminação de informações, contribuindo para compreender tanto padrões individuais de participação quanto fenômenos coletivos de circulação de conteúdo.

Portanto, ao incorporar a curiosidade social como dimensão analítica, amplia-se a compreensão sobre como a interação social impulsiona a curiosidade e molda o comportamento dos usuários em plataformas digitais de comunicação em grupo (SOUSA; ALMEIDA; FIGUEIREDO, 2022; XU et al., 2018). Esse entendimento é crucial para avançar em modelos explicativos sobre consumo e propagação de informações, oferecendo subsídios teóricos e práticos para a análise de sistemas de recomendação, estudos de difusão e dinâmicas sociais em redes online (SOUSA; ALMEIDA; FIGUEIREDO, 2019; SOUSA; ALMEIDA; FIGUEIREDO, 2022).

## 2.2 Modelos de Curiosidade na Computação

Na computação os modelos implementados de curiosidade podem ser divididos nas seguintes categorias: novidade, surpresa, incerteza, conflito, mudança e complexidade. Dentro de cada categoria, os modelos são novamente classificados em duas subcategorias, conforme o princípio do potencial de excitação intermediário seja seguido nos modelos. A primeira subcategoria de modelos segue este princípio e suporta tanto AoB quanto AoA na seleção de estímulos, enquanto a segunda subcategoria suporta apenas AoB (WU; MIAO, 2013).

Para sanar a descontinuidade entre o parágrafo e a listagem, apresenta-se a seguir uma síntese descritiva de cada categoria e de suas variantes, indicando, quando pertinente, exemplos e interpretações usuais na literatura.

- **Baseado em Novidade:** a avaliação da novidade depende do tempo e envolve uma comparação entre os estímulos atuais e as experiências anteriores, portanto os modelos baseados em novidade podem ser divididos em baseado em similaridade e baseado em similaridade e tempo (WU; MIAO, 2013). Merrick e Maher aplicaram este modelo de curiosidade para controlar o comportamento de personagens não-jogadores (NPCs) e robôs reconfiguráveis (MERRICK; MAHER, 2006; MAHER; MERRICK; SAUNDERS, 2008; MERRICK; HUNTINGTON, 2008);
- **Baseados em Surpresa:** existem duas interpretações de surpresa na literatura sobre curiosidade computacional. A primeira interpreta a surpresa como a diferença entre uma expectativa e o resultado real (WU; MIAO, 2013). O erro de previsão se encaixa bem com essa interpretação e é amplamente utilizado em modelos de curiosidade para quantificar o nível de surpresa (SCHMIDHUBER, 1991; BARTO; SINGH; CHENTANEZ, 2004; SCHMIDHUBER, 1999; UĞUR; ŞAHIN; ÇAKMAK, 2007). A segunda interpretação descreve a surpresa como o grau de não esperar algo. Storck, Hochreiter e Schmidhuber (1995) modelaram esse tipo de surpresa usando o ganho de informação antes e depois de uma observação, enquanto Macedo e Cardoso (2001) propuseram outra medida usando improbabilidade;
- **Baseados em incerteza:** segundo Berlyne (BERLYNE, 1960) a incerteza surge quando não há uma resposta clara em relação aos estímulos. A entropia na teoria da informação foi proposta para medir o grau de incerteza contaminado em um estímulo. Porém, quando um estímulo é difícil de classificar, dificilmente a entropia pode ser usada (WU; MIAO, 2013);
- **Baseados em conflito:** especificamente, Wu e Miao (2013) definiram o conflito como a incompatibilidade entre a compreensão do aprendiz humano e o conhecimento especialista incorporado no mundo virtual. Se uma relação do conceito  $C_i$  para o conceito  $C_j$  no conhecimento do usuário tiver um rótulo diferente da relação de  $C_i$  para  $C_j$  no conhecimento do mundo, então o entendimento do usuário dessa relação estará em conflito com o do especialista. Então, basicamente, a medição de conflitos ocorre quando você compara o que os usuários sabem e o que o mundo sabe e depois é calculada contando o número de pontos de conhecimento com conflitante em ambas as bases de conhecimento. Este modelo de curiosidade faz um esforço baseado no conflito para dirigir a atenção do agente de aprendizagem para os objetos virtuais que contêm mais informações em conflito com o conhecimento do usuário, a fim de impulsionar a curiosidade do usuário e melhorar sua experiência de aprendizagem;
- **Baseados em mudança:** os modelos baseados em mudanças dirigem a atenção do sistema para as mudanças no ambiente, ao invés de apenas medir quanto uma mudança ou em que grau é efetivamente um estímulo. Esses modelos não medem

a quantidade de estimulação ou curiosidade que acompanha uma mudança, mas oferecem mecanismos para redirecionar a atenção para mudanças em amostras futuras. O foco de atenção é definido por uma expressão Gaussiana que rege a geração de exemplos de treinamento, direcionando a atenção para as áreas onde recentemente apareceram novos exemplos. A propagação da Gaussiana está relacionada com o desempenho (taxa de sucesso) do mapeamento, de forma que o modelo é capaz de encaminhar a atenção com sucesso para as mudanças, levando em consideração fatores de sucesso, *timesteps*, assim como os coeficientes relacionados ao tédio e falha (WU; MIAO, 2013). O sistema exhibe um movimento em direção a regiões onde o mapeamento já existente não está fazendo previsões e um movimento afastado de regiões onde nenhuma melhoria pode ser alcançada (WU; MIAO, 2013); Do ponto de vista do aprendizado de máquina, este modelo pode ser visto como um caso especial de modelos baseados em preditores, sob a suposição de que um preditor sempre prevê que a próxima observação é igual à observação atual, significando falta de mudança. Quando ocorre mudança, o preditor comete um erro e atualiza seu conhecimento sobre o mundo (WU; MIAO, 2013);

- **Baseados em complexidade:** o modelo baseado em complexidade interpreta a complexidade como a dificuldade de um objeto virtual (estímulo) para o aprendiz. Quanto mais pontos de conhecimento contidos em um objeto virtual que para um aprendiz já não são novos, mais complexo (ou difícil) é para o aprendiz. Portanto, a complexidade de um objeto virtual é uma quantidade inteira que pode ser medida, baseada no número de pontos de conhecimento (relacionados ao objeto virtual) e no conhecimento de mundo que ainda não foram aprendidos pelo aprendiz. Quando o aprendiz é detectado em um estado de estar perdido, o agente calculará um valor de curiosidade por cada objeto virtual na vizinhança do aprendiz no mundo virtual, e fará o aprendiz explorá-lo com o maior valor de curiosidade (WU; MIAO, 2013).

## 2.3 Sistemas de Recomendação

Nos anos 90 os estudos sobre sistemas de recomendação foram iniciados, eles surgiram com o intuito de aproveitar as opiniões de milhões de usuários da Internet para filtrar conteúdo útil e interessante (JANNACH et al., 2010). Com o passar dos anos, ideias mais concretas sobre filtragem colaborativa começaram a ser desenvolvidas, levando ao surgimento dos primeiros sistemas de recomendação. Esses sistemas possuíam técnicas semelhantes nos quais seus algoritmos identificavam usuários com históricos de uso semelhantes e combinavam suas avaliações de itens para criar listas organizadas por ordem de preferência. Durante a virada do milênio, a pesquisa em sistemas de recomendação ganhou impulso significativo e atraiu profissionais de diversas disciplinas, como inteligência artificial, recuperação de informação, mineração de dados, segurança, *marketing* e negócios

(JANNACH et al., 2010). Com o aumento de dados e avanços em técnicas de aprendizado de máquina, as empresas começaram a investir em pesquisas e competições para impulsionar o desenvolvimento desses sistemas, resultando em melhorias contínuas nas técnicas de geração de recomendações (JANNACH et al., 2010).

Hoje, os sistemas de recomendação são integrados em inúmeras aplicações e plataformas em que os usuários são apresentados a um vasto leque de opções. Os mecanismos de busca utilizam sistemas de recomendação para filtrar documentos relevantes com base nas palavras-chave pesquisadas. As plataformas de e-commerce recomendam produtos com base nos comportamentos de usuários semelhantes. Além disso, os serviços de *streaming* aproveitam o histórico de uso do usuário para sugerir itens a serem consumidos (ISINKAYE; FOLAJIMI; OJOKOH, 2015). Esses exemplos ilustram como os sistemas de recomendação se tornaram uma parte integrante da experiência do usuário em ambientes digitais.

Em geral, os métodos de recomendações são classificados em duas categorias: *memory-based* e *model-based*. Em *memory-based* os métodos são definidos por operar matrizes de avaliações de usuários para os itens e utilizam as avaliações antes do processo de classificação dos itens, além disso, calculam a distância entre os usuários e os itens utilizando métricas de similaridade com base em suas proporções (BOBADILLA et al., 2013). Já no caso dos métodos baseados em modelos as informações do sistema de recomendação são utilizadas para criar um modelo que gera as recomendações, além disso esses modelos utilizam as avaliações para aprender um modelo preditivo (RICCI; ROKACH; SHAPIRA, 2011).

Segundo Bobadilla (BOBADILLA et al., 2013) a base para geração de recomendações pode ser dividida em: tipo de dados disponíveis, algoritmos de filtragem, técnicas como algoritmo de vizinho mais próximo, esparsidade da base de dados, escalabilidade e desempenho do sistema e objetivo da recomendação.

Entre os principais algoritmos de filtragem destacam-se o Filtragem Baseada em Conteúdo, Filtragem Colaborativa, Filtragem Híbrida, Filtragem Demográfica e Sistemas Baseados em Conhecimento (BOBADILLA et al., 2013). A Filtragem Baseada em Conteúdo é uma técnica que aprende a recomendar itens que são semelhantes àqueles que o usuário consumiu e gostou no passado (BOBADILLA et al., 2013). Já a Filtragem Colaborativa, identifica padrões de preferências entre usuários e recomenda itens com base nas escolhas de outros usuários semelhantes (BOBADILLA et al., 2013). Os sistemas híbridos, por sua vez, combinam múltiplas abordagens, como a filtragem colaborativa e baseada em conteúdo, para aprimorar a qualidade das recomendações. Os sistemas demográficos utilizam informações como idade, sexo e localização para prever as preferências dos usuários. Além disso, os sistemas baseados em conhecimento utilizam uma base de conhecimento para criar recomendações com base nas características do usuário e do item (BOBADILLA et al., 2013).

## 2.4 Trabalhos relacionados

Sistemas de recomendação são ferramentas essenciais na era da informação, visando ajudar usuários a descobrir itens relevantes em meio a vastas quantidades de dados. Esses sistemas aplicam técnicas de aprendizado de máquina para prever e sugerir itens personalizados com base nos interesses e histórico do usuário, enriquecendo a experiência de descoberta e tornando a seleção de itens mais eficiente e relevante (JANNACH et al., 2010). Seja para filmes, músicas, produtos ou conteúdo digital, os sistemas de recomendação desempenham um papel importante na personalização e otimização da interação do usuário. Na busca por melhorar a diversidade das recomendações desses sistemas, vários estudos já foram realizados. Assim, a partir desses trabalhos, é possível identificar vantagens bem como as limitações de cada uma das abordagens utilizadas.

Observa-se que as métricas de novidade, complexidade e incerteza são estímulos chave para despertar a curiosidade de usuários (SOUSA; ALMEIDA; FIGUEIREDO, 2019). A análise da Curva de Wundt (i.e., perfil de curiosidade), bem como as métricas relacionadas à frequência, recência e dissimilaridade (que formam a novidade), entre outras, contribuem para a compreensão mais abrangente de como diferentes aspectos da curiosidade podem influenciar na maneira como os usuários interagem com as recomendações (ZHAO; LEE, 2016). Além disso, a consideração da curiosidade social traz uma dimensão ainda mais rica para a análise de recomendações e a inclusão de métricas que abordam a curiosidade no contexto de interações sociais, como exemplificado pelo estudo de Sousa, Almeida e Figueiredo (2022), expandiu o escopo da pesquisa ao considerar como as relações interpessoais influenciam nas preferências e interações dos usuários.

O modelo proposto por Fu e Niu (2023) trata a Curva de Wundt de uma forma distinta, considerando a curiosidade como algo despertado pelo *sweet spot* que seria o ponto exato na Curva de Wundt onde o estímulo está perfeitamente balanceado para o usuário em questão. Além disso, esse artigo também descreve a Zona de Conforto, que seria uma região próxima ao *sweet spot*, que como o próprio nome já indica, refere-se a zona onde o usuário se sente confortável, ou seja, onde terá maiores chances de que sua curiosidade seja despertada. A Zona de Conforto é importante no sentido de que dessa forma a curiosidade pode ser maximizada por esse intervalo, mas o artigo também ressalta que a curiosidade também pode ser gerada à partir de pontos que estão um pouco acima da zona de conforto devido ao fator surpresa. No trabalho de Fu e Niu (2023), o SI representa o quanto um item recomendado serve como um “estímulo” para despertar a curiosidade do usuário, portanto, o artigo propõe usar o nível de surpresa como o fator que define o SI. Isso porque a surpresa capta elementos como novidade, conflito com a expectativa e complexidade, que são considerados desencadeadores da curiosidade (BERLYNE, 1960). Foi utilizado o Gerador de Recomendação que obtém os  $n$  melhores itens do Modelo de Preferência como um *pool* de candidatos. Então, ele re-classifica esses  $n$  itens, favorecendo

os itens cuja distância do limite superior da zona de conforto do usuário é menor. Isso reflete a ideia de que um estímulo um pouco além da zona de conforto do usuário terá maior probabilidade de despertar sua curiosidade.

No entanto, é importante observar que, apesar das contribuições significativas de cada abordagem, também foram identificadas algumas limitações, como a perda de precisão em certos contextos, falta de aplicação da recomendação em diferentes áreas e a falta do uso de métricas como a curiosidade social na maioria das abordagens analisadas (SOUSA; ALMEIDA; FIGUEIREDO, 2019). A contribuição deste trabalho está em uma proposta que inclui a curiosidade social como métrica complementar para um cenário de estudo de recomendações de músicas utilizando o *dataset* do Last.fm. Essa inclusão pode realizar recomendações com maior diversidade considerando não apenas os gostos individuais do usuário, mas principalmente as dinâmicas sociais que moldam as escolhas desses usuários. Porém, é importante ressaltar que o aumento da diversidade nas recomendações resulta em um comprometimento da sua precisão. Segundo Xu et al. (2021) a precisão é inversamente proporcional a diversidade.

Por fim, o estudo da curiosidade e suas implicações nos sistemas de recomendação online é um campo em constante evolução, abrindo portas para futuras investigações e aplicações (ZHOU et al., 2020; XU et al., 2021). As contribuições deste trabalho certamente servirão como base para futuros estudos que buscam não apenas equilibrar a precisão e a diversidade das recomendações, mas também entender melhor o papel da curiosidade social na dinâmica das interações entre os usuários.

A Tabela 1 sintetiza trabalhos que incorporam a curiosidade em sistemas de recomendação, destacando como cada estudo modela esse constructo e quais limitações permanecem. Observa-se a predominância de métricas de novidade (via frequência, recência e dissimilaridade), frequentemente combinadas com surpresa e/ou conflito, enquanto a Curva de Wundt é estimada por abordagens paramétricas (distribuições Beta e Normal) ou não paramétricas (histograma e função sigmoide). Alguns trabalhos introduzem ainda conceitos operacionais como o SeCM (curiosidade estimulada por estímulo), *sweet spot* e zona de conforto, buscando equilibrar estímulo e saturação. Em contrapartida, as limitações convergem para (i) suposições paramétricas restritivas sobre a distribuição dos dados, (ii) cobertura parcial das variáveis colativas (com ênfase na novidade em detrimento de complexidade e incerteza) e, sobretudo, (iii) baixa consideração da curiosidade social. Também são recorrentes restrições de domínio (p. ex., estudos focados em Last.fm ou WhatsApp), desafios de generalização (validação em outros contextos, como música) e cenários de *cold-start*, além do conhecido *trade-off* diversidade–precisão. Em conjunto, esses achados oferecem um panorama das escolhas de modelagem e das lacunas atuais, que motivam a exploração de métricas que integrem a curiosidade social sem comprometer de forma excessiva a precisão das recomendações.

Tabela 1 – Comparativo de modelos de curiosidade e limitações.

Trabalho	Modelo de Curiosidade	Limitações
Zhao e Lee (2016)	Métrica de novidade e a Curva de Wundt é estimada via distribuição Beta.	Assume que os dados seguem uma distribuição Beta e não considera as outras variáveis colativas, inclusive influência social.
Xu et al. (2021)	Métrica de novidade e conflito social. Curva de Wundt estimada via função de histograma e função sigmoide.	Não considera as outras variáveis colativas tradicionais (e.g., complexidade, novidade e incerteza).
Shrestha et al. (2020)	Métrica de curiosidade e a Curva de Wundt é estimada via distribuição Normal.	Assume que os dados seguem uma distribuição Normal e não considera as outras variáveis colativas, inclusive influência social.
Abbas e Niu (2019)	Métrica de novidade e surpresa e a Curva de Wundt é modelada via distribuição Beta.	Assume que os dados seguem uma distribuição Beta e não considera as outras variáveis colativas, inclusive influência social.
Zhao e Lee (2016)	Interação, respostas a estímulos novos e conflitantes	Perda de precisão e não abrange a curiosidade social.
Sousa, Almeida e Figueiredo (2019)	Novidade, Incerteza, Complexidade. Aborda a curiosidade através da incerteza, cria perfis de curiosidade para serem levados em consideração, e é focado no Last.fm.	Perda de precisão e não abrange a curiosidade social.
Sousa, Almeida e Figueiredo (2022)	Curiosidade Social, novidade, incerteza, conflito e complexidade. Aborda a curiosidade com muitas métricas distintas melhorando a diversidade dos itens. Inclui curiosidade social	Perda de precisão e é focado no WhatsApp.
Xu et al. (2021)	Novidade e conflito social como variáveis colativas; estima a Curva de Wundt via histograma e função sigmoide; propõe um mecanismo de curiosidade (SeCM) para reordenar recomendações preservando precisão.	Não abrange a curiosidade social.
Shrestha et al. (2020)	Novidade, Curva de Wundt, Precisão, Curiosidade, Frequência, Recência, Dissimilaridade, CC, SD, MA	Foca apenas na métrica novidade. Trabalho futuro: aplicações em outros tipos de banco de dados para validação do método neles. Ex.: músicas.
Fu e Niu (2023)	Surpresa, Curva de Wundt. Adiciona os conceitos de <i>sweet spot</i> e Zona de Conforto.	Aplicar o framework em um cenário de <i>cold-start</i> , ou seja, sem depender tanto do histórico de acesso do usuário.

## 2.5 Considerações Finais

Esta revisão sistematizou os principais referenciais teóricos e empíricos que embasam o uso de curiosidade em sistemas de recomendação. Foi observado que a literatura da psicologia consolida quatro variáveis colativas de novidade, incerteza, conflito e complexidade, bem como dimensões centrais do estímulo, e que a Curva de Wundt fornece a moldura funcional para capturar o efeito de *sweet spot* entre tédio e ansiedade. No campo computacional, essas ideias foram traduzidas em métricas e algoritmos que operam sobre distribuições, entropias, surpresas e erros de predição, resultando em mecanismos de exploração controlada capazes de aumentar diversidade sem abandonar relevância.

No domínio específico de recomendação, os estudos analisados apontam ganhos de diversidade e engajamento ao introduzir sinais de curiosidade; contudo, registram também *trade-offs* com precisão, sensibilidade a escolhas de janelas temporais e presumem, em muitos casos, distribuições paramétricas que nem sempre se verificam nos dados. Além disso, constatou-se que a dimensão *social* da curiosidade, isto é, a influência exercida por pares e redes, permanece sub-explorada em comparação às métricas individuais, apesar de evidências de que interações sociais modulam significativamente a exposição e a adoção de itens.

Em relação aos trabalhos relacionados, este trabalho posiciona-se como uma extensão dos trabalhos relacionados ao integrar variáveis colativas tradicionais fundamentadas em Berlyne (novidade, incerteza, conflito e complexidade) (BERLYNE, 1960) com variáveis de curiosidade social operacionalizadas sobre traços de interação entre usuários. Entre os estudos revisados, apenas (XU et al., 2018) considera curiosidade social e o faz por meio de uma matriz de similaridade de acessos do usuário; diferentemente, este trabalho adota um modelo completo que combina, de forma explícita, as métricas colativas clássicas com métricas sociais, proporcionando uma base unificada para analisar como sinais individuais e sociais influenciam o processo de recomendação.

Diante desse panorama, este trabalho posiciona-se ao integrar métricas de curiosidade (tradicionais e sociais) a um cenário realista de recomendação musical, explicitando a relação entre estrutura social e sequências temporais de consumo. As lacunas identificadas, tais como dependência de hipóteses fortes sobre distribuições, pouca análise de sensibilidade e cobertura limitada de sinais sociais, informam diretamente as decisões metodológicas do capítulo seguinte, no qual definimos o protocolo experimental, as janelas de análise e o conjunto de métricas selecionadas para avaliar, de forma comparativa, o impacto da curiosidade social sobre a qualidade e a diversidade das recomendações.

## 3 Metodologia

Esta seção apresenta a metodologia adotada para o desenvolvimento do trabalho, detalhando as etapas necessárias para a construção, configuração e avaliação do sistema de recomendação social proposto. Primeiramente, descreve-se o processo de coleta e integração dos dados, que envolveu tanto o histórico de acessos dos usuários quanto suas conexões sociais. Em seguida, são apresentadas as premissas e configurações dos experimentos, onde se definem as métricas de curiosidade social e as variáveis utilizadas.

Na sequência, abordam-se as métricas tradicionais e sociais, assim como o processo de seleção de métricas mais adequadas para o contexto da recomendação. Também são discutidos o método de avaliação de ranking e a métrica de Inter-User Similarity (IUS), fundamentais para mensurar a qualidade das recomendações. Por fim, apresentam-se as considerações finais, que sintetizam as escolhas metodológicas e seu impacto sobre os resultados obtidos.

Este capítulo está organizado da seguinte forma. A Seção 3.1 descreve os conjuntos de dados utilizados, o LFM-1B e o Last.fm Social Connections, detalhando suas características e o processo de integração entre eles. A Seção 3.2 explica como foi realizado o merge entre o histórico de acessos dos usuários e suas conexões de amizade. A Seção 3.3 apresenta a configuração dos experimentos, incluindo a seleção aleatória de dias, a definição de dias válidos e o número mínimo de músicas por usuário. A Seção 3.4 descreve as premissas do modelo de curiosidade adotado. As Seções 3.5 e 3.6 formalizam, respectivamente, as métricas de curiosidade social e as variáveis colativas tradicionais. A Seção 3.7 apresenta o processo de seleção de métricas por análise de redundância. Por fim, as Seções 3.8 e 3.9 descrevem as métricas de avaliação utilizadas, MAP e IUS, e a Seção 3.10 traz as considerações finais do capítulo.

### 3.1 Coleção de Dados

Aqui foi descrita as fontes de dados, o processo de aquisição e os critérios de pré-processamento utilizados para construir a base analítica deste trabalho. Foram utilizados dois conjuntos complementares: (i) os registros de escuta do LFM-1b, que fornecem o histórico comportamental de usuários do Last.fm (SCHEDL, 2016); e (ii) o Last.fm Social Connections, que agrega a dimensão relacional por meio de laços de amizade entre usuários também presentes no LFM-1b (SCHEDL, 2016). A combinação dessas duas camadas é central para investigar como a influência social e a curiosidade social afetam recomendações musicais, bem como para comparar abordagens puramente colaborativas com variantes sensíveis ao grafo social.

### 3.1.1 LFM-1B

O LFM-1B é um conjunto de dados com mais de um bilhão de eventos de escuta gerados por mais de 120 mil usuários do Last.fm (SCHEDL, 2016). Cada evento registra usuário, artista, álbum, música e um carimbo de tempo. No nível de usuário, há dados demográficos anônimos e descritores adicionais do gosto e do comportamento de consumo musical. O conjunto foi criado para pesquisa e avaliação em recuperação de informação musical e sistemas de recomendação e inclui uma matriz pré-computada de contagem de reproduções por usuário–artista, além de *scripts* em linguagem Python para carregar os arquivos e rodar um recomendador colaborativo simples.

A coleta partiu dos 250 principais *tags* da plataforma, obteve artistas e fãs, e então as histórias de escuta de um subconjunto aleatório de usuários. O período de aquisição vai de janeiro de 2013 a agosto de 2014. Os metadados (usuários, artistas, álbuns, músicas e eventos) são distribuídos em arquivos texto UTF-8, e a matriz usuário–artista (UAM) é fornecida como matriz esparsa em formato Matlab/HDF5. Os eventos têm carimbo de tempo em Unix. O pacote completo tem cerca de 8 GB e vem com exemplos de código para importação e estatísticas (SCHEDL, 2016).

As estatísticas principais incluem: 120.322 usuários, 3.190.371 artistas, 15.991.038 álbuns, 32.291.134 músicas e 1.088.161.692 eventos de escuta. Há ainda 61.534.450 pares únicos usuário–artista. Para construir a UAM final e reduzir ruído, os autores descartam usuários com menos de 10 artistas únicos e artistas ouvidos por menos de 10 usuários, o que reduz o número de artistas para 585.095 e torna a matriz resultante ( $120.175 \times 585.095$ ) manejável com cerca de 200 MB (SCHEDL, 2016). A Tabela 2 sintetiza o recorte aplicado ao conjunto de dados original, evidenciando a filtragem necessária para obter uma amostra consistente e representativa.

Tabela 2 – Descrição das principais *features* do histórico de acessos dos usuários no dataset LFM-1b.

Feature	Tipo de dado	Definição	Exemplo
<code>user-id</code>	Inteiro/Anônimo	Identificador único de cada usuário na base, sem informações pessoais.	102345
<code>artist-id</code>	Inteiro	Identificador único do artista ouvido pelo usuário.	58721
<code>album-id</code>	Inteiro	Identificador único do álbum associado à execução.	230145
<code>track-id</code>	Inteiro	Identificador único da música musical.	9457812
<code>timestamp</code>	Inteiro (Unix time)	Momento exato do evento de escuta, registrado em segundos desde 01/01/1970 (UTC).	1372639259
<code>playcount</code>	Inteiro	Número acumulado de execuções de músicas pelo usuário no Last.fm.	1520
<code>country</code>	Texto	País informado pelo usuário.	BR
<code>age</code>	Inteiro	Idade declarada pelo usuário.	23
<code>gender</code>	Texto	Gênero declarado pelo usuário.	male
<code>registered</code>	Data/Horário	Momento em que o usuário se registrou na plataforma.	2011-05-23

### 3.1.2 Last.fm Social Connections

O dataset *Last.fm Social Connections* é derivado do conjunto de dados LFM-1b e foi construído a partir da API oficial do Last.fm (SCHEDL, 2016). Para sua criação, os autores aplicaram uma estratégia de amostragem em largura, partindo de usuários do LFM-1b como nós iniciais e coletando conexões de amizade até dois níveis de distância (amigos e amigos de amigos). Foram mantidas apenas as conexões em que ambos os usuários pertenciam ao dataset original LFM-1b, garantindo consistência na rede.

O conjunto é composto por dois arquivos principais: (i) LFM-1b\_social\_ties.txt, contendo 78.989 conexões entre 11.792 usuários únicos; e (ii) LFM-1b\_users.txt, que armazena informações demográficas e de execução de músicas de cada usuário.

Este dataset foi originalmente utilizado em pesquisas sobre homofilia e predição de links em redes sociais musicais, considerando aspectos como preferências por músicas *mainstream*, novidade e diversidade (SCHEDL, 2016). No contexto deste trabalho, a finalidade do uso desse dataset é voltada para a construção de um sistema de recomendação social, onde se parte do pressuposto de que músicas ouvidas por amigos de um usuário possuem maior probabilidade de serem relevantes e, portanto, recomendadas a ele. Essa abordagem permite avaliar de que forma a influência social impacta o processo de

recomendação musical, em comparação a métodos baseados apenas no histórico individual de escuta.

A Tabela 3 apresenta as principais estatísticas do conjunto *Last.fm Social Connections*, evidenciando sua estrutura em termos de rede social. O dataset contém quase 12 mil usuários interconectados por aproximadamente 79 mil relações de amizade, compondo uma rede suficientemente densa para análises de recomendação social. Nota-se que a média de conexões por usuário é de cerca de 6,70, o que indica que a maioria dos participantes não está isolada, mas inserida em uma teia de relacionamentos que potencialmente influencia suas preferências musicais. Essa característica é fundamental para o estudo da *curiosidade social*, pois permite explorar como os laços de amizade impactam na descoberta e recomendação de músicas.

Tabela 3 – Estatísticas do conjunto Last.fm Social Connections.

Descrição	Quantidade
Usuários	11.792,00
Conexões (arestas)	78.989,00
Média de conexões por usuário	6,70

Após o refinamento dos dados, mantendo apenas usuários com vínculo social (pelo menos uma conexão) e atividade suficiente para a análise obtivemos o resultado final usado nos experimentos. Adotamos como critérios mínimos 100 músicas por semana e 600 por mês, definidos a partir das distribuições observadas no conjunto e usados para garantir volume adequado de interações por usuário. A Tabela 4 resume esse recorte: 6,011 usuários, 162,632 artistas, 1,544,354 álbuns, 4,035,397 músicas e 379,671,300 eventos, com média de 7,97 conexões por usuário. Nota-se que, embora o número total de usuários tenha sido reduzido para pouco mais de seis mil, cada um deles apresenta um volume considerável de interações, o que garante densidade suficiente para a análise de recomendações sociais. Além disso, observa-se que a base refinada ainda preserva grande diversidade de itens, com mais de quatro milhões de músicas distribuídas em mais de um milhão e meio de álbuns e mais de 160 mil artistas distintos. Outro ponto relevante é a média de aproximadamente oito conexões sociais por usuário, indicando a existência de uma rede de relacionamentos suficientemente densa para explorar efeitos de influência social nas recomendações.

## 3.2 Integração do Histórico de Acessos do Usuário com as Conexões de Amizade

Para que fosse possível explorar a influência social no consumo musical, foi realizada a integração entre dois conjuntos de dados: o **LFM-1B**, que contém o histórico de acessos individuais de usuários do Last.fm, e o **Last.fm Social Connections**, que descreve a

Tabela 4 – Dados estatísticos sobre usuários e conteúdo.

Descrição	Quantidade
Total de Usuários	6.011,00
Total de Artistas	162.632,00
Total de Álbuns	1.544.354,00
Total de Músicas	4.035.397,00
Total de Eventos	379.671.300,00
Média de Conexões por Usuário	7,97

rede de amizades entre esses mesmos usuários. Enquanto o LFM-1B fornece os eventos de escuta (usuário, artista, música, álbum e carimbo de tempo), o Social Connections fornece as arestas que conectam cada usuário aos seus amigos dentro da plataforma.

A partir dessa complementaridade, o merge dos arquivos foi realizado utilizando o `user-id` como chave primária. Inicialmente, foi extraído do LFM-1B apenas os usuários que possuíam pelo menos uma conexão no Social Connections, de modo a garantir a possibilidade de comparação entre históricos individuais e históricos de amigos. Em seguida, para cada usuário  $u$ , foram agregados não apenas os seus eventos de escuta, mas também os eventos de todos os usuários diretamente conectados a  $u$  na rede de amizades. Dessa forma, cada perfil passou a conter tanto o consumo individual quanto o consumo social, o que possibilita analisar como as preferências musicais dos amigos influenciam potenciais recomendações.

A Figura 2 apresenta a função de distribuição acumulada (CDF) do número de conexões por usuário. Observa-se que a maior parte dos usuários possui até 20 conexões, enquanto uma pequena fração concentra valores mais elevados. Esse comportamento segue o padrão típico de redes sociais reais, caracterizadas por uma distribuição de grau altamente desigual, em que poucos usuários possuem muitas conexões.

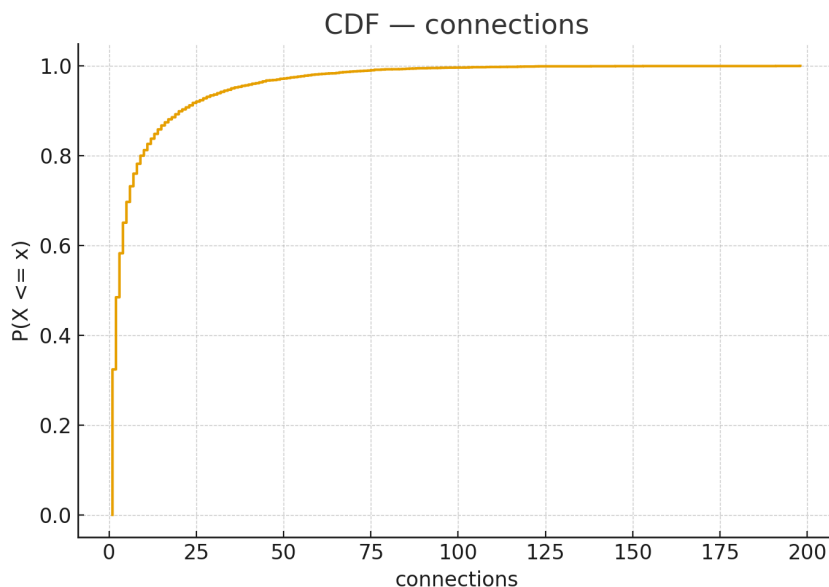


Figura 2 – Distribuição acumulada do número de conexões por usuário no dataset Last.fm Social Connections.

### 3.3 Configuração dos Experimentos

Na Figura 3, observa-se a divisão temporal do conjunto de dados em dois blocos principais: **treino** e **teste**. Cada mês é composto por quatro semanas, sendo que o primeiro mês é utilizado para treino e o segundo para teste. No conjunto de treino, são computadas as variáveis colativas (novidade, incerteza, conflito e complexidade), bem como estimada a *Curva de Wundt* de cada usuário. Essa etapa é fundamental para capturar o perfil de curiosidade individual, que futuramente será utilizado para guiar as recomendações. Já no conjunto de teste, é realizada a avaliação das recomendações, verificando se os itens sugeridos ao usuário correspondem efetivamente ao seu histórico de consumo.

A Figura 4, por sua vez, detalha como são feitas as quatro rodadas de recomendação ao longo do período de teste. Partindo de um mês de treino, cada rodada incorpora progressivamente uma nova semana do mês de teste. Assim, na primeira recomendação apenas a primeira semana do mês de teste é avaliada; na segunda recomendação, as duas primeiras semanas; na terceira, três semanas; e finalmente, na quarta recomendação, todas as quatro semanas do mês de teste. Esse procedimento permite avaliar o desempenho do sistema de forma incremental, observando como as recomendações se adaptam à medida que mais dados reais do usuário são disponibilizados.

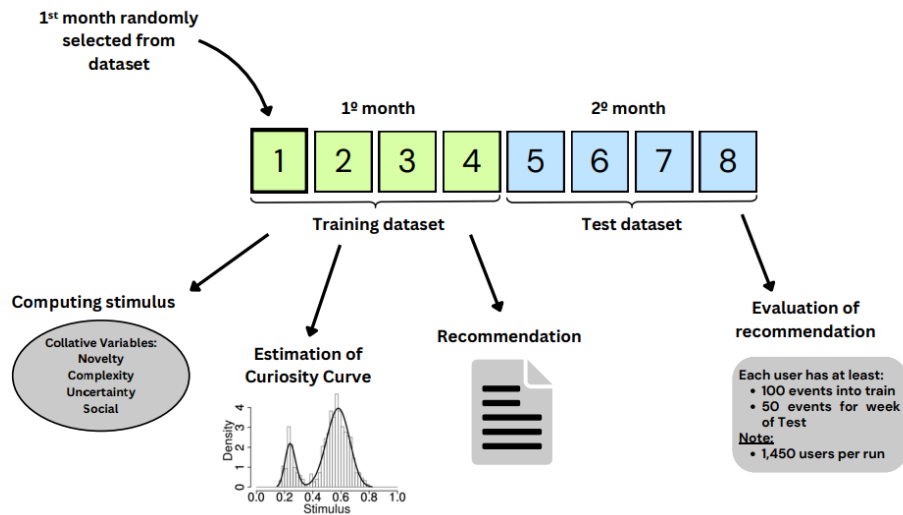


Figura 3 – Configuração dos Experimentos.

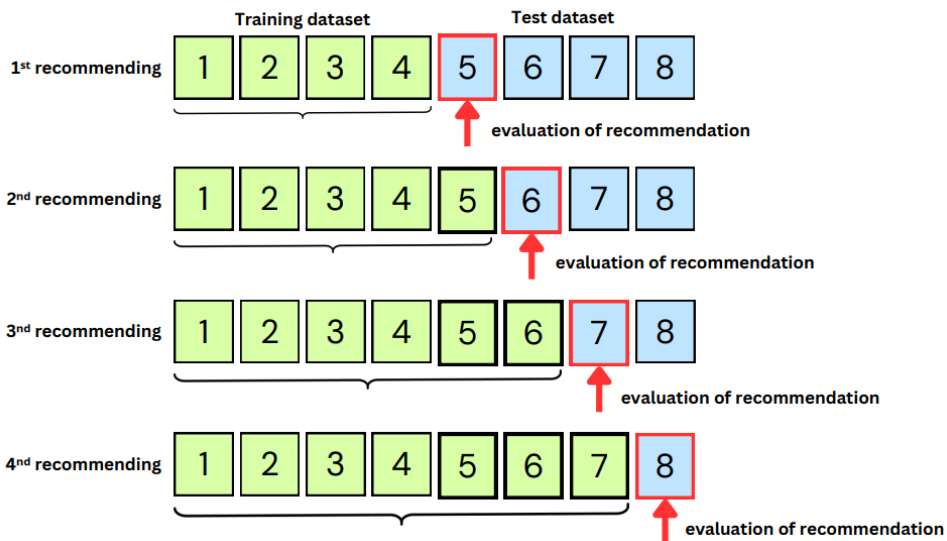


Figura 4 – Separação de Treino e Teste.

### 3.3.1 Seleção Aleatória de Dias

Para cada experimento, são selecionados 4 dias de forma aleatória e uniforme a partir de um conjunto de dias disponíveis. A probabilidade para a seleção de cada dia é de  $1/n$ , onde  $n$  representa o número total de dias válidos no conjunto de dados. Cada *run* do experimento é gerada a partir de uma seleção aleatória desses 4 dias.

### 3.3.2 Dias válidos

O período de tempo considerado para a seleção dos dias aleatórios termina em 5 de novembro, pois esta é a última data possível que ainda permite formar os **56 dias necessários** para a criação das recomendações. Esse número surge da forma como as

semanas vão sendo progressivamente incorporadas ao longo dos experimentos. Em cada rodada, o conjunto de dias válidos é ampliado da seguinte maneira:

- **Primeira recomendação:** 35 dias (5 semanas);
- **Segunda recomendação:** 42 dias (6 semanas);
- **Terceira recomendação:** 49 dias (7 semanas);
- **Quarta recomendação:** 56 dias (8 semanas).

Dessa forma, o intervalo de datas válido para seleção dos dias aleatórios se estende de **1º de janeiro até 5 de novembro**. Esse recorte garante que todas as recomendações possam ser geradas de maneira consistente, respeitando a progressão semanal prevista no experimento.

### 3.3.3 Número mínimo de músicas

A definição de um número mínimo de músicas por usuário é essencial para garantir que o sistema de recomendação tenha dados suficientes para treinar e avaliar cada perfil. Para justificar esse corte, utilizamos a análise das distribuições acumuladas (CDFs) de eventos semanais e mensais de escuta, representadas nas Figuras 5 e 6.

A Figura 5 mostra a CDF da quantidade de músicas ouvidas por semana. Observa-se que cerca de **50% dos usuários escutam pelo menos 100 músicas semanais**. Esse ponto mediano da curva nos permite definir **100 músicas por semana** como o valor mínimo necessário para que um usuário seja considerado válido na fase de treino. Dessa forma, asseguramos que a maioria dos usuários ativos possua dados suficientes para gerar recomendações consistentes.

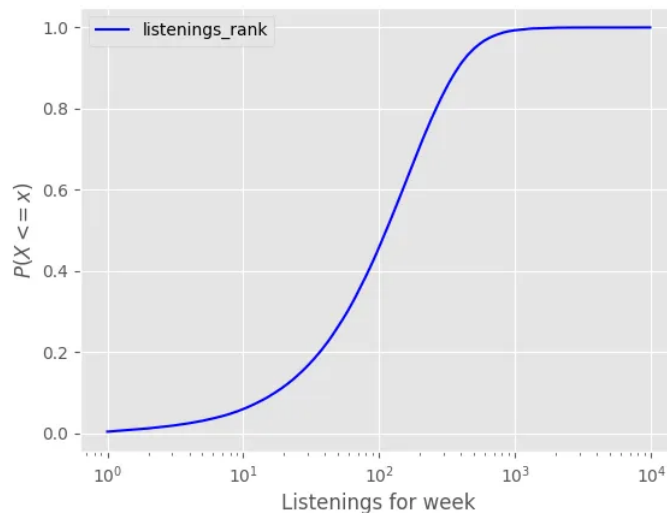


Figura 5 – Função de distribuição acumulada (CDF) da quantidade de eventos por semana.

Já a Figura 6 apresenta a CDF da quantidade de músicas ouvidas por mês. Nota-se que aproximadamente **50% dos usuários escutam pelo menos 600 músicas mensais**. Assim, foi adotado **600 músicas por mês** como o critério mínimo para inclusão de usuários na etapa de teste. Esse valor garante que os perfis utilizados nas avaliações contenham material suficiente para mensurar o desempenho das recomendações.

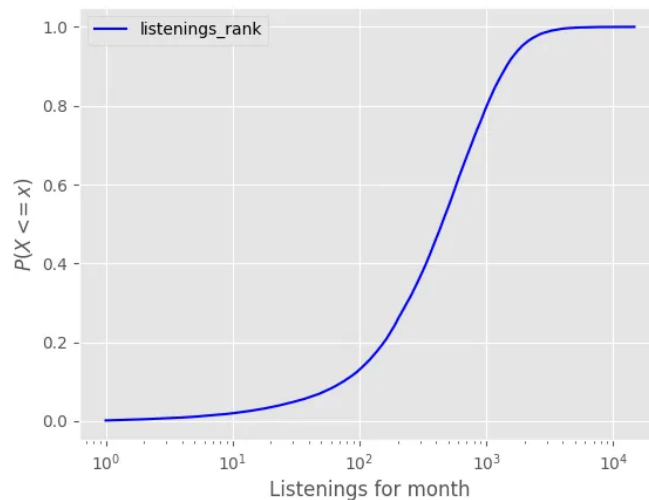


Figura 6 – Função de distribuição acumulada (CDF) da quantidade de eventos por mês.

Portanto, os valores de corte (100 músicas por semana e 600 músicas por mês) foram definidos diretamente a partir das distribuições empíricas do comportamento dos usuários. Essa abordagem torna os experimentos mais robustos, já que evita incluir perfis com atividade insuficiente que poderiam enviesar os resultados (SCHEDL, 2016).

### 3.4 Premissas da Curiosidade

A curiosidade pode ser entendida como um dos principais motores do consumo e da disseminação de informações, funcionando tanto em nível individual quanto coletivo. Do ponto de vista psicológico, autores clássicos como Berlyne (1960) já a descrevem como um mecanismo motivacional associado à busca por novidade e à redução de incertezas. Em paralelo, perspectivas computacionais avançaram no sentido de modelar a curiosidade como critério formal para seleção de itens em sistemas interativos (WU; MIAO, 2013).

No contexto de sistemas de recomendação, a curiosidade assume papel estratégico: ela permite que o modelo incentive a exploração de conteúdos novos sem perder de vista a relevância para o usuário. Isso contrasta com abordagens centradas exclusivamente em popularidade ou frequência, que tendem a reforçar itens comuns e limitar a diversidade. Trabalhos recentes demonstram que a curiosidade, quando incorporada a algoritmos de recomendação, aumenta a diversidade dos resultados e mantém o engajamento dos usuários (ZHOU et al., 2020; ZHAO; LEE, 2016; XU et al., 2021).

Além da dimensão individual, a curiosidade também pode ser observada em âmbito social. Usuários se interessam não apenas por explorar novidades por conta própria, mas também por aquilo que circula em seus grupos de interação. Essa perspectiva de curiosidade social foi consolidada por [Sousa, Almeida e Figueiredo \(2022\)](#), que mostraram como o interesse coletivo por novidades afeta a disseminação de informações em plataformas como o WhatsApp. Assim, as premissas da curiosidade adotadas neste trabalho não se limitam ao indivíduo, mas também abrangem dinâmicas sociais que influenciam padrões de consumo e de recomendação.

O modelo proposto em [Sousa, Almeida e Figueiredo \(2022\)](#) estabelece cinco premissas fundamentais que orientam a derivação das métricas de curiosidade. As três primeiras são premissas básicas e gerais, aplicáveis a qualquer usuário; as duas últimas são premissas específicas da dimensão social, que governam o cálculo das métricas de curiosidade social.

1. **Variação da curiosidade por contexto de grupo:** a curiosidade de um usuário pode ser estimulada de maneira diferente dependendo das pessoas com quem ele interage e das discussões em andamento no grupo ([SOUSA; ALMEIDA; FIGUEIREDO, 2022](#)). Em outras palavras, o estímulo de curiosidade pode variar conforme o grupo no qual o indivíduo participa: um mesmo usuário em grupos distintos pode apresentar padrões de curiosidade completamente diferentes. Portanto, as métricas são calculadas separadamente para cada grupo, considerando apenas as interações ocorridas em cada contexto;
2. **Dinamicidade temporal da curiosidade:** a forma como a curiosidade de um usuário reage a um dado estímulo pode mudar ao longo do tempo ([SOUSA; ALMEIDA; FIGUEIREDO, 2022](#)). Um estímulo que desperta grande curiosidade em determinado momento pode, posteriormente, tornar-se neutro ou até provocar tédio, caso o usuário migre para outra região da Curva de Wundt. Por essa razão, as métricas de curiosidade devem ser recalculadas a cada novo evento de interação do usuário, refletindo a natureza altamente dinâmica e contextual do fenômeno;
3. **Janela de interação como período de ativação:** a curiosidade que motiva a ação de um usuário  $u$  no instante  $t$  possui um período de ativação  $\delta T$ , definido como o intervalo de tempo imediatamente anterior à ação durante o qual os eventos ocorridos no grupo podem contribuir para estimular a curiosidade de  $u$  ([SOUSA; ALMEIDA; FIGUEIREDO, 2022](#)). Esse intervalo é denominado janela de interação e corresponde ao período  $[t - \delta T; t]$ . Eventos anteriores a  $t - \delta T$  são considerados antigos demais para exercer qualquer efeito sobre a curiosidade do usuário no instante  $t$ ;
4. **Estímulo social pela presença de outros usuários:** a curiosidade social de um usuário  $u$  no instante  $t$  é estimulada pelos demais usuários que compartilharam conteúdo no mesmo grupo durante a janela de interação ([SOUSA; ALMEIDA;](#)

FIGUEIREDO, 2022). Embora o conteúdo compartilhado e as características do próprio usuário também possam influenciar a curiosidade, a modelagem da dimensão social concentra-se exclusivamente nos outros usuários ativos na janela, tratando-os como origens potenciais de influência social. As propriedades do conteúdo são exploradas separadamente nas métricas das variáveis colativas tradicionais;

5. **Estimativa da influência social por padrões históricos:** a magnitude com que a curiosidade social de um usuário  $u$  é estimulada por outro usuário  $s$  pode ser estimada por meio de padrões históricos de interação (SOUSA; ALMEIDA; FIGUEIREDO, 2022). Especificamente, a influência de  $s$  sobre  $u$  é proporcional à frequência com que uma ação de  $s$  foi seguida por uma ação de  $u$  dentro de um intervalo máximo de  $\delta T$ , em janelas de interação anteriores à janela corrente. Essa estimativa histórica captura tanto a influência direta — quando  $u$  responde regularmente a  $s$  — quanto a influência indireta — quando  $s$  tende a exercer forte influência sobre múltiplos destinatários, mesmo sem um histórico explícito de resposta de  $u$  em particular.

Essas cinco premissas orientam a operacionalização das métricas de curiosidade adotadas neste trabalho. As premissas básicas (1, 2 e 3) fundamentam tanto as variáveis colativas tradicionais quanto as métricas sociais, ao estabelecerem que a curiosidade é dinâmica, contextual e delimitada temporalmente pela janela de interação  $\delta T$ . As premissas de curiosidade social (4 e 5) fundamentam especificamente as métricas de influência direta e indireta, formalizadas na seção seguinte por meio de conceitos da Teoria da Informação (SOUSA; ALMEIDA; FIGUEIREDO, 2022).

A Figura 7 ilustra, no contexto musical, como os estímulos de curiosidade são extraídos do histórico de escutas. No painel (a), cada linha representa uma escuta (usuário, música, tempo). Quando um usuário  $s$  realiza uma escuta no instante  $t$ , abre-se uma janela de interação  $[t, t + \delta T]$  (Premissa 3). Qualquer escuta de um amigo  $u$  que ocorra dentro dessa janela constitui uma resposta potencial ao estímulo social de  $s$  (Premissa 4). A frequência histórica desse padrão  $s \rightarrow u$  é a base para as estimativas de probabilidade condicional  $P(U=u | S=s)$  e marginal  $P(U=u)$ , que alimentam as métricas de curiosidade social (Premissa 5).

O painel (b) mostra, para as escutas-semente de  $s$ , as janelas  $[t, t + \delta T]$  e quais amigos responderam em cada janela. Esse encadeamento define curiosidade social direta quando  $u$  responde a  $s$  no intervalo, e contribui para curiosidade social indireta ao agregar, no tempo, a informação mútua total associada a  $s$ . Em nível de grupo, a distribuição de curiosidade ao longo das janelas define a entropia do grupo, enquanto a coocorrência dirigida  $s \rightarrow u$  define a informação mútua de grupo (Premissa 2). Dessa forma, o estímulo de curiosidade não é um sinal sintético, mas um padrão temporal dirigido extraído do histórico real de escutas (SOUSA; ALMEIDA; FIGUEIREDO, 2022).



Figura 7 – Exemplo da sequência de músicas consideradas para o cálculo dos estímulos.

### 3.5 Métricas de curiosidade social

As métricas de curiosidade social musical são fundamentadas em equações da Teoria da Informação, seguindo a proposta de (BERLYNE, 1960) para a formulação de métricas baseadas em variáveis colativas. Especificamente, essas métricas utilizam o conceito de informação mútua (COVER; THOMAS, 2006), que quantifica a dependência entre duas variáveis aleatórias. A informação mútua expressa a redução da incerteza (ou entropia) associada a uma variável quando a outra é observada (MACKAY, 2005). Nesse sentido, o trabalho de (SOUSA; ALMEIDA; FIGUEIREDO, 2022) oferece a formulação das métricas de curiosidade social, aqui adaptadas para o contexto musical (SOUSA; ALMEIDA; FIGUEIREDO, 2019).

Com o objetivo de descrever o comportamento musical de cada usuário, são utilizadas duas abordagens: curiosidade social direta e curiosidade social indireta. A curiosidade direta ocorre quando um amigo do usuário ouve uma música e, dentro de uma janela temporal específica (denominada  $\delta T$ ), o usuário escuta uma música como resposta, possivelmente motivado pela ação do amigo. Já a curiosidade indireta refere-se à capacidade de certos usuários influenciarem outros, mesmo sem uma frequência histórica explícita de resposta. Ou seja, mesmo que o usuário nunca tenha reagido diretamente a

uma música do amigo, ele pode ser influenciado futuramente.

Para distinguir entre o usuário cuja curiosidade está sendo analisada e os demais usuários que podem estimulá-la, adota-se a nomenclatura de *destino* ( $d$ ) para o primeiro e *origens* ( $o$ ) para os demais. Define-se  $\mathcal{O}_{t_i|d,g}^{\rightarrow}$  como o conjunto de todas as origens para o destino  $d$  que compartilhou conteúdo no instante  $t_{i|d,g}$ , ou seja, o conjunto de todos os outros usuários que compartilharam conteúdo na janela de interações  $[t_{i|d,g} - \delta T; t_{i|d,g}]$ . As probabilidades são estimadas com base nos padrões históricos de interação entre pares de usuários dentro da janela  $\delta T$ , capturados em uma tabela de contingência  $\mathbf{S}_{t,g}^{\leftarrow}$  (SOUSA; ALMEIDA; FIGUEIREDO, 2022).

### 3.5.1 Curiosidade social direta

As primeiras duas métricas são baseadas no conceito de Informação Mútua Pontual (*Pointwise Mutual Information*, PMI), que mede a redução na incerteza do destino  $d$  compartilhar uma mensagem dado o conhecimento de que a origem  $o$  também compartilhou uma mensagem dentro da janela de interação. A PMI é definida como (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$PMI_{t,g}^{\leftarrow}(D = d, O = o) = \begin{cases} \log_2 \left( \frac{P_{t,g}^{\leftarrow}(D = d | O = o)}{P_{t,g}^{\leftarrow}(D = d)} \right), & \text{se } P_{t,g}^{\leftarrow}(D = d) > 0. \\ 0, & \text{caso contrário.} \end{cases} \quad (1)$$

onde  $P_{t,g}^{\leftarrow}(D = d | O = o)$  é a probabilidade de  $d$  escutar uma música após  $o$ , dentro de  $\delta T$ , e  $P_{t,g}^{\leftarrow}(D = d)$  é a probabilidade marginal de  $d$  escutar independentemente de  $o$ , ambas estimadas com base em padrões históricos. Para os fins de estimar a influência social de  $o$  sobre  $d$  como estímulo à curiosidade, são considerados apenas valores positivos de PMI; valores não positivos são truncados em zero, pois indicam ausência de evidência de influência social (SOUSA; ALMEIDA; FIGUEIREDO, 2022). Assim, define-se a influência social de  $o$  sobre  $d$  como:

$$socInf_{t,g}^{\leftarrow}(D = d, O = o) = \max \left( PMI_{t,g}^{\leftarrow}(D = d, O = o), 0 \right) \quad (2)$$

Como a janela de interação corrente pode conter múltiplas origens estimulando a curiosidade do destino  $d$ , agrega-se a informação mútua calculada para cada origem  $o \in \mathcal{O}_{t_i|d,g}^{\rightarrow}$  por meio das funções de média e máximo, construindo duas métricas que capturam a influência direta média e máxima das origens sobre  $d$  (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$avgDirInf(d, t = t_{i|d,g}, g) = \frac{\sum_{o \in \mathcal{O}_{t_{i|d,g}}^{\leftarrow}} socInf_{t,g}^{\leftarrow}(D = d, O = o)}{|\mathcal{O}_{t_{i|d,g}}^{\leftarrow}|} \quad (3)$$

$$maxDirInf(d, t = t_{i|d,g}, g) = \max_{o \in \mathcal{O}_{t_{i|d,g}}^{\leftarrow}} \left( socInf_{t,g}^{\leftarrow}(D = d, O = o) \right) \quad (4)$$

### 3.5.2 Curiosidade social indireta

Além da influência direta, alguns usuários podem ter uma capacidade natural de influenciar outros, mesmo sem interações históricas explícitas com o destino em questão — por exemplo, um usuário recém-chegado ao grupo. Essas origens são identificadas buscando origens  $o$  que tendem a exercer forte influência sobre alguns destinos  $d$ . Para capturar esse efeito, propõem-se métricas baseadas na informação mútua total dos destinos condicionada a uma origem particular (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$MI_{t,g}^{\leftarrow}(D, O = o) = \sum_{d' \in \mathcal{U}_g} P_{t,g}^{\leftarrow}(D = d', O = o) PMI_{t,g}^{\leftarrow}(D = d', O = o) \quad (5)$$

Valores menores de  $MI_{t,g}^{\leftarrow}(D, O = o)$  sugerem influência fraca e difusa de  $o$  sobre qualquer destino. Valores maiores indicam que a influência de  $o$  tende a ser concentrada e forte sobre poucos destinos — exatamente os casos de interesse. Da mesma forma que para a influência direta, valores não positivos são truncados em zero (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$indSocInf_{t,g}^{\leftarrow}(D, O = o) = \max\left(MI_{t,g}^{\leftarrow}(D, O = o), 0\right) \quad (6)$$

Agrega-se novamente por média e máximo sobre todas as origens da janela de interação (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$avgIndInf(d, t = t_{i|d,g}, g) = \frac{\sum_{o \in \mathcal{O}_{t_{i|d,g}}^{\leftarrow}} MI_{t,g}^{\leftarrow}(D, O = o)}{|\mathcal{O}_{t_{i|d,g}}^{\leftarrow}|} \quad (7)$$

$$maxIndInf(d, t = t_{i|d,g}, g) = \max_{o \in \mathcal{O}_{t_{i|d,g}}^{\leftarrow}} \left( MI_{t,g}^{\leftarrow}(D, O = o) \right) \quad (8)$$

### 3.5.3 Métricas de grupo

A informação mútua também pode ser utilizada para quantificar como a influência social direciona a curiosidade de um grupo de usuários, oferecendo uma visão agregada

do estímulo de curiosidade social no ecossistema de um grupo  $g$  até o instante  $t_{i|d,g}$ . Para isso, introduz-se a entropia dos destinos condicionada a uma origem particular  $o$  (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$H_{t,g}^{\leftarrow}(D | O = o) = - \sum_{d' \in \mathcal{U}_g} P_{t,g}^{\leftarrow}(D = d' | O = o) \log_2 \left( P_{t,g}^{\leftarrow}(D = d' | O = o) \right) \quad (9)$$

Agrega-se essa probabilidade condicional sobre todas as origens do grupo para construir uma métrica de nível de grupo (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$groupEntropy(t = t_{i|d,g}, g) = \sum_{o \in \mathcal{U}_g} P_{t,g}^{\leftarrow}(O = o) H_{t,g}^{\leftarrow}(D | O = o) \quad (10)$$

A métrica *groupEntropy* deve ser analisada em conjunto com a entropia dos destinos  $D$  independentemente de qualquer influência social, definida como (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$H_{t,g}^{\leftarrow}(D) = - \sum_{d \in \mathcal{U}_g} P_{t,g}^{\leftarrow}(D = d) \log_2 \left( P_{t,g}^{\leftarrow}(D = d) \right) \quad (11)$$

Com base nas Equações (10) e (11), define-se a informação mútua de nível de grupo entre destinos e origens, que captura a redução na incerteza associada aos destinos devido ao conhecimento da influência social das origens no grupo. Quanto maior a informação mútua, maior a influência das origens sobre os destinos no grupo  $g$  (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$groupMutInf(t = t_{i|d,g}, g) = H_{t,g}^{\leftarrow}(D) - groupEntropy(t = t_{i|d,g}, g) \quad (12)$$

## 3.6 Variáveis colativas tradicionais

Nesta seção são apresentadas as métricas das variáveis colativas tradicionais, adaptadas para o contexto de músicas escutadas em uma plataforma social. As variáveis incluem novidade, incerteza, conflito e complexidade. Estas métricas são fundamentadas em equações da teoria da informação, conforme proposto por (BERLYNE, 1960), e baseiam-se no conceito de entropia de Shannon. A informação de Shannon (ou surpresa) de uma ocorrência específica  $x$  de uma variável aleatória  $X$  é definida como  $h(x) = -\log_2 P(X = x)$ , e a entropia de Shannon é a média ponderada dessa surpresa:  $H(X) = -\sum_x P(X = x) \log_2 P(X = x)$ .

As métricas são calculadas com base no comportamento do usuário curioso e nas características da música (e.g., categoria, artista) dentro da janela de interação corrente, denotada pela notação  $\rightarrow$  (SOUSA; ALMEIDA; FIGUEIREDO, 2022).

### 3.6.1 Novidade

A novidade do usuário considera que a experiência de escutar uma música será menos nova para usuários que frequentemente escutam músicas na janela de interação, de modo que sua curiosidade tenderá a ser menos estimulada. Com base na surpresa de Shannon, a novidade do usuário  $d$  no instante  $t$  é definida como (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$userNovelty(d, t = t_{i|d,g}, g) = \begin{cases} -\log_2\left(P_{t,g}^{\rightarrow}(D = d)\right), & \text{se } P_{t,g}^{\rightarrow}(D = d) > 0 \\ -\log_2\left(1/|\mathcal{U}_{t,g}^{\rightarrow}|\right), & \text{caso contrário} \end{cases} \quad (13)$$

onde  $P_{t,g}^{\rightarrow}(D = d)$  é a probabilidade de  $d$  compartilhar conteúdo no grupo  $g$  durante a janela de interação corrente, definida como  $n_{d|t,g}^{\rightarrow} / \sum_{u \in \mathcal{U}_g} n_{u|t,g}^{\rightarrow}$ , e  $\mathcal{U}_{t,g}^{\rightarrow}$  é o conjunto de usuários distintos que compartilharam conteúdo na janela. O caso  $P_{t,g}^{\rightarrow}(D = d) = 0$  corresponde à surpresa máxima, sendo definida como a distribuição uniforme sobre os destinos.

De forma análoga, a novidade associada à categoria da música (e.g., gênero) é (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$catNovelty(\mathcal{C}_m, t = t_{i|d,g}, g) = \begin{cases} -\log_2\left(\bar{P}_{t,g}^{\rightarrow}(\mathcal{C}_m)\right), & \text{se } \bar{P}_{t,g}^{\rightarrow}(\mathcal{C}_m) > 0 \\ -\log_2\left(1/|\mathcal{C}_{t,g}^{\rightarrow}|\right), & \text{caso contrário} \end{cases} \quad (14)$$

onde  $\bar{P}_{t,g}^{\rightarrow}(\mathcal{C}_m) = \frac{1}{|\mathcal{C}_m|} \sum_{c \in \mathcal{C}_m} P_{t,g}^{\rightarrow}(C = c)$  é a probabilidade média das categorias associadas à mensagem, e  $P_{t,g}^{\rightarrow}(C = c) = n_{c|t,g}^{\rightarrow} / \sum_{k \in \mathcal{C}} n_{k|t,g}^{\rightarrow}$  é a probabilidade de uma categoria específica na janela. Esta métrica captura o quão surpreendente é que a mensagem compartilhada no instante  $t$  possua a categoria  $c$ .

### 3.6.2 Incerteza

A incerteza reflete a diversidade de usuários ou músicas na janela de interação. Quanto maior a diversidade, maior a entropia e, portanto, maior o potencial de curiosidade estimulada (SOUSA; ALMEIDA; FIGUEIREDO, 2022). As métricas de incerteza relacionadas a usuários e categorias são definidas como:

$$userUncertainty(t = t_{i|d,g}, g) = - \sum_{d \in \mathcal{U}_g} P_{t,g}^{\rightarrow}(D = d) \log_2 \left( P_{t,g}^{\rightarrow}(D = d) \right) \quad (15)$$

$$catUncertainty(t = t_{i|d,g}, g) = - \sum_{c \in \mathcal{C}_{t,g}^{\rightarrow}} P_{t,g}^{\rightarrow}(C = c) \log_2 \left( P_{t,g}^{\rightarrow}(C = c) \right) \quad (16)$$

A ideia por trás dessas métricas é que a curiosidade do destino pode ser mais ou menos estimulada pela maior ou menor diversidade nos usuários que compartilham mensagens (e nas categorias dessas mensagens) durante a janela de interação.

### 3.6.3 Conflito

Segundo Berlyne (BERLYNE, 1960), o conflito ocorre quando o mesmo estímulo provoca múltiplas respostas incompatíveis, sendo positivamente relacionado à intensidade das respostas concorrentes. Para operacionalizar esse conceito, os diferentes elementos (categorias/usuários) presentes na janela de interação representam as respostas potencialmente incompatíveis, e a intensidade de cada resposta é capturada pela probabilidade de ocorrência de cada elemento. O conflito é obtido calculando-se a probabilidade média dos usuários (ou categorias) sobre todos os elementos da janela e tomando a surpresa desse resultado (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$userConflict(t = t_{i|d,g}, g) = - \log_2 \left( \frac{1}{|\mathcal{U}_{t,g}^{\rightarrow}|} \sum_{d \in \mathcal{U}_g} P_{t,g}^{\rightarrow}(D = d) \right) \quad (17)$$

$$catConflict(t = t_{i|d,g}, g) = - \log_2 \left( \frac{1}{|\mathcal{C}_{t,g}^{\rightarrow}|} \sum_{c \in \mathcal{C}_{t,g}^{\rightarrow}} P_{t,g}^{\rightarrow}(C = c) \right) \quad (18)$$

### 3.6.4 Complexidade

Como medida alternativa de diversidade, a complexidade é capturada pelas ocorrências únicas de tipos de conteúdo na janela de interação. Ao contrário da incerteza e do conflito, que consideram apenas as categorias presentes na janela, a complexidade captura uma noção de diversidade que leva em conta todas as categorias possíveis (contidas no conjunto  $\mathcal{M}$ ) (SOUSA; ALMEIDA; FIGUEIREDO, 2022):

$$catComplex(t = t_{i|d,g}, g) = - \log_2 \left( \frac{|\mathcal{C}_{t,g}^{\rightarrow}|}{|\mathcal{M}|} \right) \quad (19)$$

Ressalta-se que as métricas definidas nas Equações (13)–(19) são sempre não-negativas e devem ser interpretadas como: quanto maior o valor, maior o estímulo à curiosidade do usuário destino (SOUSA; ALMEIDA; FIGUEIREDO, 2022).

### 3.7 Processo de seleção de métricas de curiosidade

Para identificar quais métricas de curiosidade social capturam aspectos distintos do estímulo de curiosidade, classificamos as métricas como **complementares** ou **redundantes**.

Essa análise é feita por meio do cálculo do coeficiente de correlação de Spearman entre os pares de métricas para cada usuário individualmente. Com quatro variáveis, são obtidos seis coeficientes distintos por usuário. Em seguida, verifica-se o percentual de coeficientes que se encontram nos intervalos  $[-1, -0.5]$  ou  $[0.5, 1.0]$ , indicando correlação moderada ou forte. Tais pares são considerados redundantes. Caso contrário, são classificados como complementares, por apresentarem baixa correlação.

A Figura 8 apresenta uma matriz onde cada célula representa um par de métricas, e seu valor indica o percentual de usuários para os quais esse par apresentou correlação redundante. As últimas quatro linhas da matriz evidenciam que as métricas de curiosidade social são complementares às variáveis colativas tradicionais em mais de 84% dos casos, indicando que capturam dimensões distintas do estímulo de curiosidade.

Além disso, observa-se que as métricas de curiosidade social indireta são complementares às métricas de curiosidade social direta em mais de 59% dos casos, o que sugere que ambas as categorias de métrica capturam informações diferentes entre si.

Para evitar redundância nos resultados finais, foi adotado um critério de seleção entre pares com mais de 70% de correlação redundante: apenas uma métrica do par é mantida. Por exemplo, entre o par (`maxDirCur`, `avgDirCur`), que apresenta 90% dos casos com correlação redundante, opta-se pela métrica `maxDirCur`, por sua facilidade de interpretação.

Dessa forma, as métricas selecionadas para compor os resultados apresentados no próximo capítulo são:

- `maxDirCur`: curiosidade social direta máxima;
- `maxIndCur`: curiosidade social indireta máxima;
- `maxNoveltySocialCur`: curiosidade social de novidade máxima;
- `User Conflict`: conflito entre usuários;
- `User Uncertainty`: incerteza dos usuários;
- `Track Novelty`: novidade das músicas;
- `Genre Complexity`: complexidade de gêneros musicais.

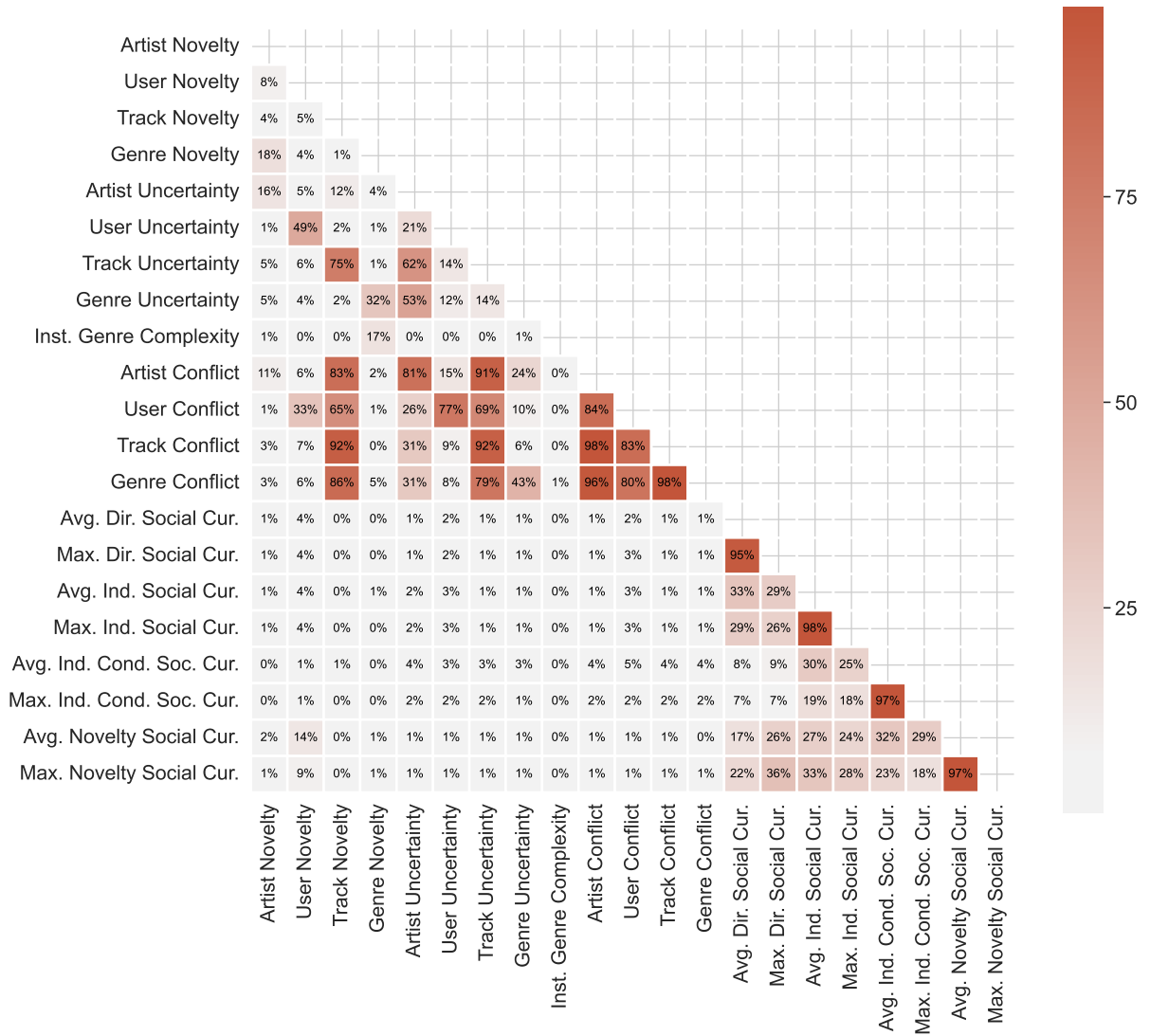


Figura 8 – Heatmap de Correlação entre métricas.

### 3.8 Método de Avaliação de Ranking

MAP calcula a média das melhores recomendações encontradas pela função AP, na tabela 6 é possível visualizar a descrição das variáveis utilizadas para ela (JANNACH et al., 2010). De acordo com essas variáveis será demonstrado o funcionamento da métrica .

A métrica de *Mean Average Precision* (MAP) é definida como (JANNACH et al., 2010):

$$MAP(F) = \frac{1}{|F|} \sum_{f \in F} AP(\text{Rec}(f, ws)). \tag{3.1}$$

Em busca de calcular se as melhores recomendações são recomendadas antes, a função AP faz uma média de P@k para descobrir se em todas as listas de recomendações

Tabela 5 – Métricas selecionadas após a análise de redundância (definições conforme as equações do modelo).

Categoria	Métrica	Definição
Curiosidade social (individual)	$\max\text{DirCur}(u, t)$	Máxima curiosidade social <i>direta</i> do usuário $u$ no instante $t$ : $\max_{s \in S} \{\max(\text{PMI}(u, s), 0)\}$ , onde $S$ é o conjunto de origens (amigos) na janela de interação.
Curiosidade social (individual)	$\max\text{IndCur}(u, t)$	Máxima curiosidade social <i>indireta</i> do usuário $u$ no instante $t$ , baseada na informação mútua condicionada: $\max_{s \in S} \{\max(\text{MI}(U, s), 0)\}$ .
Curiosidade social (gruppal)	$\text{groupMutInf}(t)$	Informação mútua entre destinos e origens na janela corrente; valores maiores indicam mais “estimuladores” sociais de curiosidade no grupo.
Variáveis colativas — Novidade (usuário)	$\text{userNovelty}(u, t)$	Surprisal associado ao usuário $u$ agir na janela; maior quando $u$ age raramente.
Variáveis colativas — Novidade (música)	$\text{trackNovelty}(C, t)$	Surprisal associado à categoria/tipo da música $C$ (p. ex., gênero) na janela; maior quando a categoria é rara.
Variáveis colativas — Incerteza (usuário)	$\text{userUncertainty}(t)$	Incerteza (entropia de Shannon) da distribuição de usuários atuando na janela; maior dispersão implica maior incerteza.
Variáveis colativas — Conflito (usuário)	$\text{userConflict}(t)$	Conflito associado aos usuários, medido por divergências de comportamento/preferências observadas na janela.
Variáveis colativas — Complexidade (gênero)	$\text{genreComplexity}(t)$	Complexidade do conjunto de categorias/gêneros na janela, capturada via entropia; maior diversidade implica maior complexidade.

Tabela 6 – Descrição das variáveis.

Variável	Significado
F	<i>features</i> textuais (título, tags) → id das músicas
f	<i>feature</i> em específico → id da música em análise
ws	peso de cada <i>feature</i> → peso da música
k	número de músicas recomendadas (top-k)
ftags	conjunto de tags (artista, álbum,...) → artista, gênero musical, música

os itens mais relevantes são mostrados primeiro.

A métrica de *Average Precision* (AP) é definida como (JANNACH et al., 2010):

$$AP(\text{Rec}(f, ws)) = \frac{1}{|f_{\text{tags}}|} \sum_{k=1}^{|f_{\text{tags}}|} P@k(\text{Rec}(f, ws), k) \cdot \text{rel}(k). \quad (3.2)$$

A função  $\text{Rec}$  ordena os itens de acordo com o peso da tag ( $ws$ ). Portanto, os temas que aparecem em posições mais altas em  $\text{Rec}$  são os considerados mais promissores para a recomendação (JANNACH et al., 2010):

$$\text{Rec}(f, ws). \quad (3.3)$$

*Precision at k* ( $P@k$ ) tem como objetivo encontrar a porcentagem de recomendações promissoras no conjunto de  $k$ . Ou seja, número de itens relevantes dentro dos recomendados dividido pelo número de itens recomendados no total (JANNACH et al., 2010).

$$P@k(\text{Rec}(f, ws), k) = \frac{|f_{\text{tags}} \cap \text{Rank}(\text{Rec}(f, ws), k)|}{k}. \quad (3.4)$$

A função  $Rel$  é uma função binária que indica se o termo na posição  $k$  é relevante:

$$Rel(k) = \begin{cases} 1, & \text{se é relevante.} \\ 0, & \text{se não é relevante.} \end{cases} \quad (3.5)$$

A função  $Rank$  seleciona os  $k$  termos ordenados antes pela função  $Rec$ :

$$Rank(Rec(f, ws), k). \quad (3.6)$$

Suponha que vamos recomendar músicas para o usuário **user\_id = 123**. Ele já ouviu várias músicas, de diferentes álbuns e artistas. Queremos gerar uma lista das **top-3 músicas recomendadas** para ele (ou seja,  $k = 3$ ), levando em conta:

- Peso de cada música, definido pela métrica de curiosidade;
- Similaridade das músicas com outras que ele já ouviu (baseado em gênero musical).

Como exemplo tem-se o seguinte histórico de músicas para o usuário **user\_id = 123**:

Tabela 7 – Histórico de músicas do usuário.

track_id	album_id	artist_id	ws	gender_id
1	10	100	20	Rock
2	11	101	15	Pop
3	12	100	5	Rock
4	13	102	8	Pop
5	14	103	12	Pop
6	15	104	16	Rap

A função  $Rec$  ordena as músicas pelo peso  $ws$ , ou seja, pelas mais frequentes. A lista de músicas recomendadas para o usuário será ordenada assim:

$$Rec = [musica\_1, musica\_2, musica\_6, musica\_5, musica\_4, musica\_3].$$

A função  $Rank$  seleciona as **top-3** músicas da lista  $Rec$ . Portanto:

$$Rank(Rec(f, ws), 3) = [musica\_1, musica\_2, musica\_6].$$

Suponha que temos um conjunto de tags que definem músicas como relevantes para o usuário (baseadas nos gêneros que ele mais escuta). Nesse caso, vamos definir que o usuário gosta principalmente de **Rock** e **Pop**.

Dado que as músicas recomendadas são [musica\_1, musica\_2, musica\_6], vamos verificar a relevância:

- **musica\_1**: Rock (Relevante);
- **musica\_2**: Pop (Relevante);
- **musica\_6**: Rap (Não Relevante).

A precisão em  $k = 1$  será:

$$P@1 = \frac{\text{Número de itens relevantes}}{\text{Número de itens recomendados}} = \frac{1}{1} = 1.0.$$

A precisão em  $k = 2$  será:

$$P@2 = \frac{\text{Número de itens relevantes}}{\text{Número de itens recomendados}} = \frac{1}{2} = 0.5.$$

A precisão em  $k = 3$  será\*:

$$P@3 = \frac{\text{Número de itens relevantes}}{\text{Número de itens recomendados}} = \frac{2}{3} \approx 0.66.$$

Para calcular a média de precisão **AP**, tomamos o valor de  $P@k$  para cada posição  $k$ :

$$AP = \frac{1}{|f_{tags}|} \sum_{k=1}^{|f_{tags}|} P@k \cdot \text{rel}(k).$$

Neste caso,  $f_{tags}$  seria o total de itens relevantes (3 itens escolhidos para o usuário).

$$AP = \frac{1}{3}(P@1 \cdot 1 + P@2 \cdot 0 + P@3 \cdot 1) = \frac{1}{3}(1 + 0 + 0.66) \approx 0.55.$$

Para calcular o **MAP** (Média da Precisão Média), precisamos da média de **AP** para todos os usuários do sistema. Suponha que tenhamos calculado o **AP** para vários usuários, e para o usuário 123, obtivemos um **AP** de 0.55.

Se tivermos mais usuários e seus respectivos valores de **AP**, o **MAP** seria a média de todos esses valores.

### 3.9 Métrica Inter-User Similarity (IUS)

A Inter-User Similarity (IUS) é uma métrica utilizada para avaliar o grau de personalização em sistemas de recomendação. Seu objetivo é medir o quanto as listas de recomendações entregues a usuários distintos se assemelham entre si (ZHOU et al., 2020). A intuição por trás dessa métrica é que um bom sistema de recomendação deve

ser capaz de produzir listas adaptadas ao perfil individual de cada usuário, evitando que todos recebam conteúdos muito semelhantes.

Formalmente, seja  $L_i$  a lista de itens recomendados ao usuário  $i$ , e  $L_j$  a lista do usuário  $j$ , ambas de tamanho  $K$ . A similaridade entre esses dois usuários é dada por (ZHOU et al., 2020):

$$IUS_{i,j} = \frac{|L_i \cap L_j|}{K}. \quad (3.7)$$

O valor de  $IUS_{i,j}$  varia entre 0 e 1. Valores próximos de 1 indicam que as listas de  $i$  e  $j$  possuem grande sobreposição (baixa personalização), enquanto valores próximos de 0 revelam listas bastante distintas (alta personalização). O valor final de IUS do sistema é calculado como a média da similaridade entre todos os pares de usuários (ZHOU et al., 2020).

Dessa forma, a IUS complementa métricas tradicionais como precisão, recall ou hit rate, pois não avalia apenas a acurácia das recomendações, mas também sua capacidade de diversificação entre usuários. Em particular, valores mais baixos de IUS são desejáveis, pois indicam que o sistema está fornecendo recomendações mais específicas e personalizadas, evitando a homogeneização causada por algoritmos que favorecem apenas itens muito populares (ZHOU et al., 2020).

### 3.10 Considerações Finais

Este capítulo descreveu a metodologia empregada para investigar o impacto da influência social em recomendações musicais. Foi iniciado pela coleta e caracterização dos dados, integrando o histórico de escuta do LFM-1b ao grafo de amigos do Last.fm Social Connections. Essa integração permitiu vincular, para cada usuário, as sequências temporais de eventos (listenings) às conexões de amizade relevantes, possibilitando modelar e mensurar a propagação social de escutas dentro de janelas temporais específicas. Os critérios de filtragem adotados (atividade mínima semanal e mensal) asseguraram densidade suficiente de interações por usuário, reduzindo a esparsidade e fortalecendo a confiabilidade das estimativas de probabilidade.

Em seguida, foi detalhada a configuração experimental baseada em janelas deslizantes ao longo de oito semanas consecutivas: quatro semanas para treino e quatro para teste, com atualização progressiva do conjunto de treino a cada nova semana avaliada. O procedimento de amostragem aleatória de dias, conjuntamente com a definição dos períodos válidos, foi delineado para garantir comparabilidade entre execuções (*runs*) e cobrir diferentes regimes de atividade dos usuários.

Em relação à **modelagem da curiosidade social**, foi formalizado duas famílias de métricas. As métricas de curiosidade social direta que capturam respostas temporais

do usuário às escutas dos amigos dentro de uma janela de interação  $\delta T$ , via PMI e suas variantes agregadas. As métricas de curiosidade social indireta que quantificam o potencial de influência exercido por um usuário sobre o grupo (ou sobre pares) independentemente de histórico explícito de resposta direta, por meio de informação mútua condicionada. Complementarmente, apresentamos as variáveis colativas tradicionais (novidade, incerteza, conflito e complexidade), fundamentadas em entropia e surpresa de Shannon, úteis para decompor dimensões clássicas do estímulo informacional. A análise de correlação (Spearman) guiou o processo de seleção de métricas, reduzindo redundância e retendo apenas as medidas que capturam sinais distintos e relevantes ao objetivo do estudo.

Para a avaliação, adotamos métricas de ordenação (ranking) centradas em relevância, com ênfase na *Average Precision* (AP) por usuário e sua média global (MAP). Essa escolha privilegia não somente a presença de itens corretos nas listas recomendadas, mas a sua posição relativa, refletindo a utilidade prática para o usuário final. Métricas complementares de personalização (e.g., similaridade entre listas de usuários) foram consideradas no desenho metodológico e poderão ser discutidas como análise adicional quando pertinente aos resultados, mantendo o foco principal na eficácia de ranking.

Como pontos positivos da metodologia: (i) o alinhamento entre dados, hipóteses e métricas (integração explícita entre grafo social e histórico temporal), (ii) a preservação de causalidade temporal no protocolo de avaliação e (iii) a curadoria de métricas para evitar sobreposição informacional. Por outro lado, há limitações: (a) dependência de atividade mínima por usuário, que pode excluir perfis de baixa frequência; (b) sensibilidade às escolhas de janela  $\delta T$  e aos cortes temporais; (c) uso de um único domínio (música no Last.fm), o que impõe cautela na generalização para outras plataformas e tipos de item, ainda mais pela data antiga dos dados. Tais limitações são endereçáveis em trabalhos futuros por meio de análise de sensibilidade, validação cruzada de janelas e replicações em domínios distintos.

Em síntese, este capítulo estabeleceu os insumos de dados, o protocolo experimental e o conjunto de métricas necessários para investigar se as escutas dos amigos aumentam a chance de recomendação para o usuário-alvo. No capítulo 4, apresentamos e discutimos os **resultados**, examinando (i) o ganho de desempenho ao incorporar sinais de curiosidade social, (ii) a contribuição relativa das variáveis colativas tradicionais e (iii) a robustez das conclusões sob diferentes configurações experimentais.

## 4 Resultados

Este capítulo apresenta e discute os principais resultados obtidos a partir da aplicação da metodologia descrita no Capítulo 2. O objetivo é analisar o desempenho dos métodos de recomendação baseados em *Curiosidade* e em *Frequência*, considerando tanto métricas clássicas de precisão quanto indicadores de diversidade e personalização. A análise busca responder se a incorporação de métricas de curiosidade é capaz de enriquecer a experiência do usuário, em contraste com abordagens tradicionais que privilegiam exclusivamente a repetição de itens populares.

Além disso, são avaliadas as curvas de validação geradas durante os experimentos, de modo a verificar a consistência dos padrões de uso dos dados em diferentes execuções (*runs*). Em seguida, apresentam-se os resultados da métrica *Mean Average Precision* (MAP), que mensura a qualidade da ordenação das recomendações, e da métrica *Inter-User Similarity* (IUS), voltada para avaliar a diversidade entre as listas geradas para diferentes usuários.

Estruturalmente, este capítulo está organizado da seguinte forma: na Seção 4.1, apresentamos as considerações iniciais do experimento; na Seção 4.2, discutimos as validações das curvas; na Seção 4.3, analisamos os resultados da métrica MAP; na Seção 4.4, apresentamos os resultados da métrica IUS; e, por fim, na Seção 4.5, trazemos as considerações finais que sintetizam os achados e preparam a discussão do próximo capítulo.

### 4.1 Considerações Iniciais

Para análise do método de ranking, foram comparados dois tipos de ranking:

- **Ranking por Curiosidade:** busca novidades, preservando diversidade entre pessoas;
- **Ranking por Frequência:** tende a privilegiar itens mais frequentes/populares no histórico.

Em cada recomendação, há quatro recomendações de datas diferentes (“one”, “two”, “three”, “four”) e quatro execuções independentes (**Run 1** a **Run 4**). Todas elas foram avaliadas em **Top-10**, **Top-20** e **Top-50**:

1. Para cada método (Curiosidade e Frequência) e para cada corte Top- $k$  (Top-10, Top-20, Top-50), coletamos os quatro valores por run;
2. Calculamos a média  $\bar{x}$  e o desvio padrão amostral;
3. Construimos o IC95%;

4. Reportamos, para cada combinação (método, Top- $k$ ), os valores de Média, Desvio padrão e IC95%.

## 4.2 Validações das Curvas

A etapa de validação das curvas tem como objetivo analisar o comportamento dos usuários em diferentes curvas Betas, verificando a consistência dos resultados obtidos em múltiplas execuções (Runs). Para isso, foram consideradas quatro rodadas independentes de validação, onde se calculou a média de usuários e suas respectivas porcentagens em cada categoria de Beta. Essa análise permite avaliar a estabilidade dos resultados e identificar a predominância de certos padrões de utilização entre os usuários.

Tabela 8 – Média das Validações Run 1 (número de usuários por categoria de Beta).

Betas	Usuários
1 Beta	2,25
2 Beta	64,50
3 Beta	206,25
Usuários não válidos	5,50

Tabela 9 – Média das Validações Run 2 (número de usuário por categoria de Beta).

Betas	Usuários
1 Beta	7,00
2 Beta	379,25
3 Beta	1240,25
Usuários não válidos	41,25

Tabela 10 – Média das Validações Run 3 (número de usuários por categoria de Beta).

Betas	Usuários
1 Beta	5,00
2 Beta	320,25
3 Beta	1036,00
Usuários não válidos	35,50

Os resultados obtidos nas quatro rodadas de validação das Tabelas 8, 9, 10, 11 evidenciam uma predominância da categoria 3 Beta, que concentra a maior parte dos usuários em todas as execuções, indicando que este é o cenário mais representativo para os dados analisados. A categoria 2 Beta também apresenta valores consistentes e relevantes, atuando como segunda maior concentração de usuários válidos. Já a categoria 1 Beta aparece em menor escala, com números significativamente reduzidos em comparação às demais.

Tabela 11 – Média das Validações Run 4 (número de usuários por categoria de Beta).

Betas	Usuários
1 Beta	11,25
2 Beta	407,00
3 Beta	1241,75
Usuários não válidos	38,50

Por fim, observa-se que os casos classificados como usuários não válidos são minoritários, mas ainda assim presentes em todas as execuções, servindo como indicador de inconsistências ou registros não aproveitáveis. De maneira geral, as médias obtidas sugerem estabilidade nos resultados entre as execuções, reforçando a confiabilidade das curvas validadas.

### 4.3 Resultados da Métrica MAP

A métrica de *Mean Average Precision* (MAP) avalia se os itens mais relevantes para o usuário aparecem entre as primeiras posições do ranking (JANNACH et al., 2010). Em geral, sistemas baseados em frequência apresentam vantagem nessa métrica, pois eles privilegiam músicas e artistas já familiares ao usuário. No entanto, como será discutido a seguir, esse comportamento não necessariamente reflete maior personalização ou qualidade a longo prazo.

A Tabela 12 apresenta os valores de MAP obtidos para o método de **Curiosidade** e a Tabela 13 mostra os valores para o método de **Frequência**.

Tabela 12 – Média da métrica MAP por run para o método de Curiosidade.

Run	Top-10	Top-20	Top-50
Run 1	0,1737	0,1711	0,1643
Run 2	0,1698	0,1686	0,1629
Run 3	0,1802	0,1783	0,1738
Run 4	0,1656	0,1636	0,1588

Tabela 13 – Média da métrica MAP por run para o método de Frequência.

Run	Top-10	Top-20	Top-50
Run 1	0,3296	0,3317	0,3327
Run 2	0,4479	0,4502	0,4528
Run 3	0,3440	0,3552	0,3640
Run 4	0,5016	0,5045	0,4971

Tabela 14 – MAP — quantas vezes o ranking de Frequência é melhor que o de Curiosidade.

Run	Top-10	Top-20	Top-50
Run 1	0,90	0,94	1,02
Run 2	1,64	1,67	1,78
Run 3	0,91	0,99	1,09
Run 4	2,03	2,08	2,13

Tabela 15 – Média, desvio padrão e intervalo de confiança de 95% da métrica MAP para os métodos Curiosidade e Frequência.

Método / Top- $k$	Média	Desvio-padrão	IC95%
<b>Curiosidade</b>			
Top-10	0,1723	0,1644	(0,0917, 0,2529)
Top-20	0,1704	0,1617	(0,0912, 0,2496)
Top-50	0,1649	0,1565	(0,0883, 0,2415)
<b>Frequência</b>			
Top-10	0,4058	0,3277	(0,2453, 0,5663)
Top-20	0,4104	0,3321	(0,2477, 0,5731)
Top-50	0,4117	0,3335	(0,2482, 0,5752)

Observa-se que o método baseado em frequência alcança valores significativamente mais altos de MAP, especialmente em Run 1 e Run 2, onde as recomendações se concentram em músicas já conhecidas ou muito próximas do perfil histórico do usuário. Isso confirma que a métrica favorece estratégias conservadoras, focadas em maximizar precisão imediata.

Por outro lado, o método de curiosidade apresenta valores mais baixos de MAP, com exceção de alguns casos. Essa oscilação indica que a curiosidade pode, em determinadas execuções, alcançar níveis competitivos de precisão, mas seu foco principal não é maximizar a repetição de itens já esperados.

É importante ressaltar que a métrica MAP, por definição, não captura aspectos de diversidade nem de diferenciação entre usuários, avaliando apenas a ordenação de itens relevantes no ranking individual (SHRESTHA et al., 2020). Portanto, um sistema pode obter MAP alto simplesmente repetindo recomendações populares e homogêneas, o que pode limitar a descoberta de novos conteúdos.

Nesse sentido, o método da curiosidade, mesmo com MAP mais baixo, deve ser analisado em conjunto com outras métricas, como a Inter-User Similarity (IUS). Enquanto o MAP privilegia recomendações óbvias, a IUS permite verificar se as listas geradas são de fato distintas entre usuários. Como será discutido na próxima seção, o modelo de curiosidade apresenta melhor desempenho em termos de personalização e diversidade, o que mostra seu potencial para enriquecer a experiência do usuário a longo prazo.

Tabela 18 – IUS: quantas vezes o ranking de Curiosidade é melhor que o de Frequência.

Run	Top-10	Top-20	Top-50
Run 1	0,20	2,08	3,55
Run 2	0,74	1,73	3,23
Run 3	0,89	1,85	2,89
Run 4	0,93	2,14	3,34

## 4.4 Resultados da Métrica IUS

A métrica Inter-User Similarity (IUS) é utilizada para avaliar o grau de sobreposição entre listas de recomendação geradas para diferentes usuários. Quanto maior o valor de IUS, maior é a similaridade entre as listas recomendadas (ZHOU et al., 2020), indicando que o sistema tende a repetir itens independentemente do perfil do usuário. Por outro lado, valores mais baixos de IUS representam maior diversidade entre as listas, sugerindo que o método de recomendação é capaz de personalizar melhor os resultados.

Neste trabalho, comparamos os métodos de Curiosidade e Frequência a partir da métrica IUS em diferentes cortes de listas (Top-10, Top-20 e Top-50). O objetivo é verificar qual abordagem consegue preservar maior diversidade e personalização, como será apresentado nas próximas tabelas e análises.

Tabela 16 – Média da métrica IUS por run para o método de Curiosidade.

Run	Top-10	Top-20	Top-50
Run 1	0.000301	0.000186	0.000176
Run 2	0.000238	0.000209	0.000203
Run 3	0.000237	0.000227	0.000226
Run 4	0.000258	0.000230	0.000226

Tabela 17 – Média da métrica IUS por run para o método de Frequência.

Run	Top-10	Top-20	Top-50
Run 1	0.000360	0.000573	0.000800
Run 2	0.000413	0.000571	0.000859
Run 3	0.000447	0.000647	0.000879
Run 4	0.000498	0.000722	0.000980

Tabela 19 – Média, desvio padrão e intervalo de confiança de 95% (IC95%) da métrica IUS para os métodos Curiosidade e Frequência.

Método / Top- $k$	Média	Desvio Padrão	IC95%
<b>Curiosidade</b>			
Top-10	0.000259	0.000030	[0.000211 ; 0.000306]
Top-20	0.000213	0.000020	[0.000181 ; 0.000245]
Top-50	0.000208	0.000024	[0.000170 ; 0.000246]
<b>Frequência</b>			
Top-10	0.000430	0.000058	[0.000337 ; 0.000522]
Top-20	0.000628	0.000072	[0.000514 ; 0.000743]
Top-50	0.000880	0.000075	[0.000760 ; 0.000999]

A Tabela 19 apresenta a comparação entre os métodos Curiosidade e Frequência para a métrica IUS. Observa-se que o método de Curiosidade apresenta valores menores em todos os cenários avaliados. No Top-10, a Curiosidade alcançou em média 0.000259, contra 0.000430 do método de Frequência, indicando aproximadamente 40% menos sobreposição entre listas de usuários.

Essa diferença aumenta à medida que o tamanho da lista aumenta. No Top-20, a Curiosidade obteve 0.000213, enquanto a Frequência chegou a 0.000628, ou seja, uma redução de cerca de 66%. Já no Top-50, o contraste é ainda maior: 0.000208 para Curiosidade contra 0.000880 para Frequência, o que representa uma queda de aproximadamente 76% na similaridade entre recomendações.

Outro ponto observado é que a métrica IUS para o método de Frequência cresce de forma acentuada com o aumento de  $k$  (o que indica repetição de recomendações e menor diversidade entre usuários), a Curiosidade se mantém praticamente constante em torno de 0.0002, mesmo em listas mais longas. Essa característica mostra sua capacidade de preservar a personalização e a diferenciação entre usuários, independentemente do tamanho da recomendação.

Portanto, os resultados mostram que a abordagem baseada em Curiosidade é superior em termos de personalização, entregando recomendações mais distintas e relevantes. O método de Frequência, por outro lado, tende a repetir conteúdos e a perder diversidade à medida que o Top- $k$  aumenta, o que compromete a experiência personalizada. Dessa forma, a Curiosidade consolida-se como a estratégia mais eficaz para manter listas de recomendação realmente personalizadas.

## 4.5 Considerações Finais

Os resultados apresentados ao longo deste capítulo permitem destacar diferenças claras entre os métodos de recomendação analisados. O método baseado em **Frequência**

apresentou desempenho superior na métrica MAP, confirmando sua eficácia em maximizar precisão imediata por meio da recomendação de itens já conhecidos ou populares. Contudo, esse comportamento também revelou limitações, pois tende a reduzir a diversidade das recomendações e gerar listas homogêneas entre usuários.

Por outro lado, o método de **Curiosidade** obteve valores de MAP mais modestos, mas apresentou resultados significativamente melhores na métrica IUS, indicando maior diversidade e personalização. A consistência dos valores de IUS em diferentes tamanhos de listas (Top-10, Top-20 e Top-50) evidencia a capacidade dessa abordagem de gerar recomendações diferenciadas entre os usuários, reduzindo a sobreposição de itens e favorecendo a descoberta de novos conteúdos.

As análises das curvas de validação reforçaram a estabilidade dos experimentos, demonstrando padrões consistentes entre execuções independentes. Isso confere robustez às conclusões sobre o comportamento dos métodos avaliados.

Em síntese, pode-se concluir que o uso de métricas de curiosidade, embora não maximize a precisão em curto prazo, contribui para a construção de sistemas de recomendação mais ricos, personalizados e socialmente relevantes. Esse equilíbrio entre precisão e diversidade será retomado no próximo capítulo, onde são discutidas as implicações dos resultados e delineamos perspectivas para trabalhos futuros.

## 5 Conclusão

O presente trabalho teve como objetivo geral investigar a viabilidade e o impacto da curiosidade social na geração de recomendações musicais, com ênfase na promoção de diversidade e descoberta. Em contraposição a abordagens centradas apenas em popularidade e frequência. A seguir, descrevemos como cada objetivo específico foi atingido e como eles se articulam para a consecução do objetivo geral.

Para execução do objetivo de criar um modelo de curiosidade social para a plataforma de músicas online do Last.fm foi realizada uma fundamentação teórica a partir da literatura em psicologia e ciência da computação, incorporando variáveis colativas (novidade, incerteza, conflito e complexidade) e a dimensão social da curiosidade (SOUSA; ALMEIDA; FIGUEIREDO, 2022). A modelagem foi operacionalizada por meio da curva de Wundt, que captura a relação entre estimulação intermediária e maior propensão à curiosidade. Além disso, o trabalho adaptou as premissas de curiosidade social descritas no artigo do WhatsApp para o contexto musical do Last.fm.

Para avaliar o desempenho das recomendações com métricas de ranking e diversidade: as recomendações foram avaliadas com base na métrica *Mean Average Precision* (MAP), que mensura precisão de ranking, e na métrica *Inter-User Similarity* (IUS), que captura a personalização e diversidade das listas. A avaliação foi realizada em quatro *runs* independentes, considerando diferentes cortes (Top-10, Top-20, Top-50).

Comparar recomendações orientadas à curiosidade e à frequência: a comparação empírica demonstrou um *trade-off* claro entre precisão e diversidade. Os métodos baseados em frequência obtiveram valores mais altos de MAP (por exemplo, até 0,50 no Top-10 em algumas execuções), enquanto a curiosidade apresentou valores mais baixos (em torno de 0,17), o que confirma a tendência de que recomendações conservadoras maximizam precisão imediata. Por outro lado, a curiosidade se destacou na métrica IUS, alcançando valores consistentemente menores (em torno de 0,0002), enquanto a frequência alcançou valores mais altos (até 0,0009), indicando maior repetição e menor diferenciação entre usuários. Esse resultado evidencia que a curiosidade gera listas mais personalizadas e menos homogêneas.

De forma geral, os resultados numéricos confirmam as hipóteses iniciais: a curiosidade social é eficaz para ampliar diversidade e personalização, ainda que à custa de menor precisão em termos de ranking. Esse comportamento é coerente com a literatura, que aponta o dilema entre relevância e descoberta em sistemas de recomendação (ZHOU et al., 2020; XU et al., 2021). Assim, o objetivo geral de investigar a aplicação da curiosidade social em recomendações musicais foi alcançado, ainda que a construção de um sistema

híbrido completo permaneça como perspectiva de trabalhos futuros.

A metodologia adotada combinou: (i) preparação e separação dos dados do Last.fm, com seleção aleatória de janelas para gerar execuções independentes (*runs*); (ii) separação de treino e teste; (iii) definição de pré-requisitos mínimos de atividades de usuários; (iv) avaliação por *MAP* (precisão de ranking) e por *Inter-User Similarity* (IUS), métrica que capta personalização/diversidade entre listas; e (v) sumarização estatística por média, desvio-padrão e intervalo de confiança de 95% (IC95%), seguindo procedimento com distribuição *t* para múltiplas *runs*.

Os experimentos compararam um método baseado em *curiosidade social* com uma estratégia por *frequência*, avaliando simultaneamente *precisão* (MAP) e *descoberta/personalização* (IUS). As validações indicaram consistência entre as execuções: casos de usuários não válidos foram minoritários em todas as *runs* e as médias permaneceram estáveis, reforçando a confiabilidade das curvas e dos resultados sumarizados.

A precisão MAP, conforme esperado, a estratégia por frequência obteve os maiores valores de MAP em todos os cortes de *ranking*. Nas estatísticas agregadas, a frequência apresentou médias de 0,4058 (Top-10), 0,4104 (Top-20) e 0,4117 (Top-50), enquanto o método de curiosidade obteve 0,1723 (Top-10), 0,1704 (Top-20) e 0,1649 (Top-50). Os intervalos de confiança de 95% confirmam essa diferença, indicando vantagem consistente da abordagem por frequência na precisão imediata.

Em contrapartida, na Diversidade/Personalização (IUS), a curiosidade se destacou na métrica IUS ao produzir listas mais diferenciadas entre usuários: obteve valores consistentemente menores (em torno de 0,0002), ao passo que a frequência alcançou valores mais altos (até 0,0009), sinalizando maior repetição e menor personalização quando se privilegia apenas o histórico. Esse comportamento evidencia que o mecanismo de curiosidade efetivamente amplia a descoberta e a singularidade das listas.

Conforme a Tabela 20, a estratégia por Frequência superou a Curiosidade em MAP em todas as execuções e cortes de ranking (12/12 comparações), com ganhos que variam de 0,90 a 2,13 e tendência de aumento do ganho à medida que *k* cresce (por exemplo, Run 1: 0,90  $\rightarrow$  1,02; Run 4: 2,03  $\rightarrow$  2,13 do Top-10 ao Top-50). Em contrapartida, a Tabela 21 mostra que a Curiosidade superou a Frequência em IUS também em 12/12 comparações, com ganhos de 0,20 a 3,55, e efeito mais forte para listas maiores (Top-50 apresenta os maiores ganhos em todas as runs). Em síntese, os resultados das duas tabelas confirmam o trade-off precisão  $\times$  diversidade: frequência maximiza precisão (MAP), enquanto curiosidade maximiza diversidade/personalização (IUS) (ZHOU et al., 2020), com vantagens que se intensificam conforme o tamanho da lista recomendada aumenta.

Os achados confirmam as hipóteses do trabalho: *há um trade-off entre precisão e descoberta*. A curiosidade social amplia diversidade e personalização (IUS menor), ainda

Tabela 20 – Ganho da Frequência em relação à Curiosidade na métrica MAP.

Run	Top-10	Top-20	Top-50
Run 1	0,90	0,94	1,02
Run 2	1,64	1,67	1,78
Run 3	0,91	0,99	1,09
Run 4	2,03	2,08	2,13

Tabela 21 – Ganho da Curiosidade em relação à Frequência na métrica IUS.

Run	Top-10	Top-20	Top-50
Run 1	0,20	2,08	3,55
Run 2	0,74	1,73	3,23
Run 3	0,89	1,85	2,89
Run 4	0,93	2,14	3,34

que com redução de MAP frente à frequência—resultado alinhado à literatura sobre o equilíbrio entre relevância e exploração (XU et al., 2021; ZHAO; LEE, 2016). Desse modo, o objetivo central de investigar e demonstrar a utilidade da curiosidade social para diversificar recomendações musicais foi alcançado.

## 5.1 Limitações do Trabalho

Apesar dos resultados alcançados, este trabalho apresenta algumas limitações que precisam ser consideradas.

A primeira limitação é de natureza computacional. O volume de dados utilizado, aliado ao número elevado de execuções (*runs*) e de conjuntos de recomendações possíveis, exigiria um poder de processamento superior ao disponível. Por essa razão, foi necessário reduzir tanto a quantidade de *runs* quanto a diversidade de experimentos realizados. Essa restrição pode impactar a precisão estatística das análises e limitar a profundidade da comparação entre os diferentes métodos de recomendação e valores de  $k$ .

Outra limitação importante está relacionada à própria base de dados utilizada. O *dataset* do Last.fm, embora bastante abrangente e amplamente empregado na literatura, possui dados antigos (coletados entre 2013 e 2014). Isso significa que ele pode não representar fielmente os padrões atuais de consumo de música, marcados pela ascensão de novas plataformas de *streaming* e mudanças significativas no comportamento do usuário. Portanto, ao extrapolar os resultados para cenários modernos, é preciso cautela, visto que as dinâmicas de engajamento e recomendação podem ter se transformado ao longo dos anos.

Além disso, o modelo adotado assume que há influência direta dos amigos sobre o

comportamento de escuta dos usuários. Essa premissa levou à necessidade de realizar o *merge* entre os arquivos de histórico de acessos do usuário-alvo e os arquivos de histórico de acessos dos seus respectivos amigos. Embora tal abordagem seja adequada para explorar a dimensão social das recomendações, ela também implica uma simplificação: assume-se que todos os amigos têm impacto uniforme sobre o usuário, o que na prática pode variar de acordo com a intensidade e qualidade de cada relação social.

Em resumo, as limitações principais deste estudo podem ser sintetizadas em três pontos: (i) restrições computacionais que limitaram o número de experimentos; (ii) uso de um *dataset* antigo, que pode não refletir os padrões atuais de consumo musical; e (iii) a simplificação do modelo de influência social, que considera de forma homogênea as conexões de amizade ao realizar o *merge* dos históricos de escuta.

## 5.2 Principais Contribuições

Com base nessa fundamentação teórica, foi implementado e avaliado um modelo que incorpora métricas de curiosidade ao processo de recomendação, com o objetivo de equilibrar precisão, novidade e diversidade. Para isso, o sistema foi avaliado a partir de duas métricas complementares, já mencionadas anteriormente: *Mean Average Precision* (MAP) e *Inter-User Similarity* (IUS).

Os resultados mostraram que em termos de MAP a abordagem por frequência obteve valores mais elevados, confirmando sua capacidade de reforçar padrões já consolidados no histórico dos usuários. Entretanto, a métrica IUS evidenciou o diferencial da curiosidade: listas mais personalizadas e menos repetitivas entre usuários, além de recomendações mais diversificadas dentro do perfil individual de cada um. Esses resultados reforçam que, embora a curiosidade não aumente a precisão imediata, ela contribui de maneira significativa para resolver problemas de estagnação e ampliar o potencial de descoberta no consumo de músicas.

Destarte, as principais contribuições do trabalho se referem a Integração da curiosidade social em Sistemas de Recomendação, ou seja, operacionalização de uma *discovery utility* inspirada em variáveis colativas e na Curva de Wundt para música online, conectando evidências de curiosidade em consumo digital; ao ambiente experimental replicável, separação do dataset em treino e teste, com múltiplas *runs* e sumarização com IC95%, permitindo comparar acurácia (MAP) e diversidade entre usuários (IUS) de forma transparente; e bem como a evidência do *trade-off* em precisão e descoberta, que demonstra quantitativamente que indução da curiosidade reduz a sobreposição entre listas (IUS) de forma consistente, ainda que com perda de MAP, resultado coerente com abordagens que personalizam o nível de novidade por perfil.

Em síntese, nossos achados mostram que a incorporação explícita de métricas

de curiosidade em sistemas de recomendação de música é uma estratégia viável para romper a estagnação sem sacrificar por completo a utilidade imediata: ainda que a precisão (MAP) caia frente a abordagens por frequência, o ganho consistente em diversidade e personalização interusuários (IUS) amplia o espaço de descoberta de forma mensurável.

### 5.3 Trabalhos Futuros

Para trabalhos futuros, é proposto avançar da etapa de modelagem para a implementação de um sistema de recomendação híbrido orientado à descoberta, baseado em filtro colaborativo e Deep Learning, que incorpora um modelo de curiosidade social para aprimorar a diversidade das recomendações de músicas na plataforma alvo (Last.fm), alinhando-se às formulações originalmente previstas para o algoritmo completo. Além disto, a avaliação das recomendações deste sistema no conjunto de teste, reportando métricas de relevância e de descoberta, de acordo com as práticas de avaliação em sistemas de recomendação.

## Referências

- ABBAS, F.; NIU, X. One size does not fit all: Modeling users' personal curiosity in recommender systems. *arXiv preprint*, arXiv, arXiv:1907.00119, 2019. Citado 2 vezes nas páginas 18 e 26.
- BARTO, A. G.; SINGH, S.; CHENTANEZ, N. Intrinsically motivated learning of hierarchical collections of skills. In: *Proceedings of the 3rd International Conference on Development and Learning (ICDL)*. San Diego, CA, USA: IEEE, 2004. Citado na página 21.
- BERLYNE, D. E. *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill, 1960. Citado 13 vezes nas páginas 10, 16, 17, 18, 19, 20, 21, 24, 27, 36, 39, 42 e 44.
- BOBADILLA, J. et al. Recommender systems survey. *Knowledge-Based Systems*, Elsevier, v. 46, p. 109–132, 2013. Citado na página 23.
- BRODERS, S. *The Age of Curiosity: The Neural Network of an Idea in Eighteenth-Century English Literature*. [S.l.]: Walter de Gruyter GmbH & Co KG, 2021. v. 72. Citado na página 16.
- BUYALSKAYA, A.; CAMERER, C. F. The neuroeconomics of epistemic curiosity. *Current Opinion in Behavioral Sciences*, v. 35, p. 141–149, 2020. Citado 3 vezes nas páginas 10, 15 e 18.
- COVER, T. M.; THOMAS, J. A. *Elements of Information Theory*. 2nd. ed. [S.l.]: Wiley-Interscience, 2006. Citado na página 39.
- DASTON, L. Wonders and the order of nature 1150–1750. 1998. Citado na página 16.
- DONG, F. et al. A personalized hybrid recommendation system oriented to e-commerce mass data in the cloud. In: IEEE. *2013 IEEE International Conference on Systems, Man, and Cybernetics*. [S.l.], 2013. p. 1020–1025. Citado na página 10.
- FU, Z.; NIU, X. Modeling users' curiosity in recommender systems. *ACM Transactions on Knowledge Discovery from Data*, ACM New York, NY, v. 18, n. 1, p. 1–23, 2023. Citado 2 vezes nas páginas 24 e 26.
- GRIDACH, M. Hybrid deep neural networks for recommender systems. *Neurocomputing*, v. 413, p. 23–30, 2020. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231220309966>>. Citado na página 9.
- ISINKAYE, F.; FOLAJIMI, Y.; OJOKOH, B. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, Elsevier, v. 16, n. 3, p. 261–273, 2015. Citado na página 23.
- JANNACH, D. et al. *Recommender systems: an introduction*. [S.l.]: Cambridge University Press, 2010. Citado 8 vezes nas páginas 9, 11, 22, 23, 24, 46, 47 e 54.

- LOEWENSTEIN, G. The psychology of curiosity: A review and reinterpretation. *Psychological bulletin*, American Psychological Association, v. 116, n. 1, p. 75, 1994. Citado 3 vezes nas páginas 10, 16 e 19.
- MACEDO, L.; CARDOSO, A. Modeling forms of surprise in learning. In: *Proceedings of the International Conference on Cognitive Modeling (ICCM)*. Bethesda, MD, USA: [s.n.], 2001. Citado na página 21.
- MACKAY, D. J. *Information Theory, Inference and Learning Algorithms*. [S.l.]: Cambridge University Press, 2005. Citado na página 39.
- MAHER, M. L.; MERRICK, K.; SAUNDERS, R. Achieving creative behaviour using curious learning agents. In: *Papers from the 2008 AAI Spring Symposium on Creative Intelligent Systems*. Stanford, CA, USA: AAI Press, 2008. Citado na página 21.
- MERRICK, K. E.; HUNTINGTON, E. Attention focus in curious, reconfigurable robots. In: *Proceedings of the Australasian Conference on Robotics and Automation (ACRA)*. Canberra, Australia: [s.n.], 2008. Citado na página 21.
- MERRICK, K. E.; MAHER, M. L. Motivated reinforcement learning for non-player characters in persistent computer game worlds. In: *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE)*. Hollywood, CA, USA: [s.n.], 2006. Citado na página 21.
- MOREIRA, G. de S. P.; FERREIRA, F.; CUNHA, A. M. da. News session-based recommendations using deep neural networks. In: *Proceedings of the 3rd Workshop on Deep Learning for Recommender Systems*. New York, NY, USA: Association for Computing Machinery, 2018. (DLRS 2018), p. 15–23. ISBN 9781450366175. Disponível em: <<https://doi.org/10.1145/3270323.3270328>>. Citado na página 9.
- RENNER, B. Measuring social curiosity. *Journal of Individual Differences*, Hogrefe & Huber Publishers, v. 27, n. 4, p. 229–241, 2006. Citado na página 19.
- RENNER, B.; SCHÜTZ, A. The multifaceted nature of social curiosity: Evidence for a four-factor model. *Journal of Individual Differences*, Hogrefe Publishing, v. 33, n. 1, p. 29–38, 2012. Citado na página 19.
- RICCI, F.; ROKACH, L.; SHAPIRA, B. Introduction to recommender systems. In: *Recommender Systems Handbook*. Boston, MA: Springer US, 2011. p. 1–35. Citado na página 23.
- SCHEDL, M. The lfm-1b dataset for music retrieval and recommendation. In: *Proceedings of the ACM ICMR*. New York, NY, USA: Association for Computing Machinery, 2016. (ICMR'16). Citado 4 vezes nas páginas 28, 29, 30 e 36.
- SCHMIDHUBER, J. A possibility for implementing curiosity and boredom in model-building neural controllers. In: *Proceedings of the International Conference on Simulation of Adaptive Behavior (SAB)*. Paris, France: MIT Press, 1991. Citado na página 21.
- SCHMIDHUBER, J. Artificial curiosity based on the minimum description length principle. In: *Proceedings of the International Conference on Artificial Neural Networks (ICANN)*. Edinburgh, UK: IEE/IEEE, 1999. Citado na página 21.

SHANKAR, A. Capitalism, curiosity, and specialization. *The Newsletter of the Technical Committee on Cognitive and Development Systems*, IEEE Computational Intelligence Society, v. 15, n. 1, p. 12, 2018. Citado na página 10.

SHRESTHA, P. et al. Curiosity-inspired personalized recommendation. In: *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. [S.l.: s.n.], 2020. p. 33–40. Citado 5 vezes nas páginas 10, 16, 18, 26 e 55.

SILVIA, P. J. Interest—the curious emotion. *Current Directions in Psychological Science*, SAGE, v. 17, n. 1, p. 57–60, 2008. Citado 3 vezes nas páginas 17, 18 e 19.

SOUSA, A. M.; ALMEIDA, J. M.; FIGUEIREDO, F. Analyzing and modeling user curiosity in online content consumption: A lastfm case study. In: *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. [S.l.: s.n.], 2019. p. 426–431. Citado 6 vezes nas páginas 9, 20, 24, 25, 26 e 39.

SOUSA, A. M.; ALMEIDA, J. M.; FIGUEIREDO, F. Metrics of social curiosity: The whatsapp case. *Online Social Networks and Media*, v. 29, p. 100200, 2022. ISSN 2468-6964. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2468696422000040>>. Citado 14 vezes nas páginas 9, 12, 20, 24, 26, 37, 38, 39, 40, 41, 42, 43, 44 e 59.

STORCK, J.; HOCHREITER, S.; SCHMIDHUBER, J. Reinforcement-driven information acquisition in non-deterministic environments. In: *Proceedings of the International Conference on Artificial Neural Networks (ICANN)*. Paris, France: Springer, 1995. Citado na página 21.

UĞUR, E.; ŞAHİN, E.; ÇAKMAK, M. Curiosity-driven learning of affordances in robots. In: *Proceedings of the IEEE International Conference on Development and Learning (ICDL)*. Monterey, CA, USA: IEEE, 2007. Citado na página 21.

WOJTOWICZ, Z.; LOEWENSTEIN, G. Curiosity and the economics of attention. *Current Opinion in Behavioral Sciences*, Elsevier, v. 35, p. 135–140, 2020. Citado na página 15.

WU, Q.; MIAO, C. Curiosity: From psychology to computation. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 46, n. 2, p. 1–26, 2013. Citado 8 vezes nas páginas 9, 11, 16, 17, 20, 21, 22 e 36.

WUNDT, W. M. *Grundzüge der physiologischen Psychologie*. [S.l.]: W. Engelman, 1874. v. 1. Citado na página 9.

XU, J. et al. Modeling curiosity in social media: A case study on twitter. In: *Proceedings of the International AAAI Conference on Web and Social Media*. [S.l.: s.n.], 2018. v. 12, n. 1, p. 603–606. Citado 2 vezes nas páginas 20 e 27.

XU, K. et al. Enhancing recommender systems with a stimulus-evoked curiosity mechanism. *IEEE Transactions on Knowledge and Data Engineering*, v. 33, n. 6, p. 2437–2451, 2021. Citado 8 vezes nas páginas 9, 12, 18, 25, 26, 36, 59 e 61.

ZHANG, F.-g.; SHENG-HUA, X. Analysis of trust-based e-commerce recommender systems under recommendation attacks. In: IEEE. *The First International Symposium on Data, Privacy, and E-Commerce (ISDPE 2007)*. [S.l.], 2007. p. 385–390. Citado na página 10.

ZHAO, P.; LEE, D. L. How much novelty is relevant? it depends on your curiosity. In: *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York, NY, USA: Association for Computing Machinery, 2016. (SIGIR '16), p. 315–324. ISBN 9781450340694. Disponível em: <<https://doi.org/10.1145/2911451.2911488>>. Citado 8 vezes nas páginas 9, 16, 17, 18, 24, 26, 36 e 61.

ZHOU, T. et al. *How Much Novelty is Needed? Curiosity-Driven Recommender Systems*. New York, NY, USA, 2020. 388–392 p. Citado 9 vezes nas páginas 9, 12, 25, 36, 49, 50, 56, 59 e 60.

# Apêndices

# APÊNDICE A – Resultados das métricas

## A.1 Métrica MAP por Run

### Curiosidade

Tabela 22 – Resultados completos da métrica MAP para o método Curiosidade (por Run e recomendação)

<b>Run / Recomendação</b>	<b>Top-10</b>	<b>Top-20</b>	<b>Top-50</b>
Run 1 - One	0.1489	0.1489	0.1411
Run 1 - Two	0.1384	0.1389	0.1311
Run 1 - Three	0.0048	0.0000	0.0000
Run 1 - Four	0.4025	0.3964	0.3851
Run 2 - One	0.1374	0.1389	0.1365
Run 2 - Two	0.1473	0.1459	0.1384
Run 2 - Three	0.0000	0.0000	0.0000
Run 2 - Four	0.3945	0.3894	0.3765
Run 3 - One	0.0000	0.0000	0.0000
Run 3 - Two	0.2033	0.2059	0.2041
Run 3 - Three	0.0066	0.0067	0.0068
Run 3 - Four	0.5110	0.5008	0.4841
Run 4 - One	0.1417	0.1446	0.1420
Run 4 - Two	0.1534	0.1507	0.1471
Run 4 - Three	0.0000	0.0000	0.0000
Run 4 - Four	0.3673	0.3592	0.3460

## Frequência

Tabela 23 – Resultados completos da métrica MAP para o método Frequência (por Run e recomendação)

<b>Run / Recomendação</b>	<b>Top-10</b>	<b>Top-20</b>	<b>Top-50</b>
Run 1 - One	0.6920	0.6961	0.6929
Run 1 - Two	0.6264	0.6308	0.6380
Run 1 - Three	0.0000	0.0000	0.0000
Run 1 - Four	0.0000	0.0000	0.0000
Run 2 - One	0.6337	0.6386	0.6442
Run 2 - Two	0.6581	0.6620	0.6668
Run 2 - Three	0.0000	0.0000	0.0000
Run 2 - Four	0.5000	0.5000	0.5000
Run 3 - One	0.0000	0.0000	0.0000
Run 3 - Two	0.6193	0.6303	0.6443
Run 3 - Three	0.0065	0.0068	0.0073
Run 3 - Four	0.7500	0.7836	0.8072
Run 4 - One	0.6607	0.6672	0.6682
Run 4 - Two	0.6456	0.6509	0.6553
Run 4 - Three	0.0000	0.0000	0.0000
Run 4 - Four	0.7000	0.7000	0.6649

## A.2 Métrica IUS por Run

### Curiosidade

Tabela 24 – Resultados completos da métrica IUS para o método Curiosidade (por Run e recomendação)

<b>Run / Recomendação</b>	<b>Top-10</b>	<b>Top-20</b>	<b>Top-50</b>
Run 1 - One	0.000329	0.000249	0.000215
Run 1 - Two	0.000183	0.000133	0.000156
Run 1 - Three	0.000499	0.000204	0.000176
Run 1 - Four	0.000193	0.000160	0.000157
Run 2 - One	0.000246	0.000209	0.000200
Run 2 - Two	0.000236	0.000201	0.000200
Run 2 - Three	0.000250	0.000208	0.000205
Run 2 - Four	0.000220	0.000220	0.000208
Run 3 - One	0.000207	0.000208	0.000205
Run 3 - Two	0.000244	0.000240	0.000237
Run 3 - Three	0.000250	0.000233	0.000229
Run 3 - Four	0.000248	0.000227	0.000232
Run 4 - One	0.000270	0.000208	0.000205
Run 4 - Two	0.000254	0.000238	0.000236
Run 4 - Three	0.000247	0.000232	0.000229
Run 4 - Four	0.000260	0.000241	0.000235

## Frequência

Tabela 25 – Resultados completos da métrica IUS para o método Frequência (por Run e recomendação)

<b>Run / Recomendação</b>	<b>Top-10</b>	<b>Top-20</b>	<b>Top-50</b>
Run 1 - One	0.000488	0.000639	0.000909
Run 1 - Two	0.000509	0.000656	0.000880
Run 1 - Three	0.000242	0.000598	0.000825
Run 1 - Four	0.000203	0.000397	0.000587
Run 2 - One	0.000419	0.000586	0.000884
Run 2 - Two	0.000412	0.000573	0.000857
Run 2 - Three	0.000423	0.000558	0.000850
Run 2 - Four	0.000398	0.000567	0.000844
Run 3 - One	0.000467	0.000656	0.000907
Run 3 - Two	0.000426	0.000620	0.000913
Run 3 - Three	0.000445	0.000649	0.000981
Run 3 - Four	0.000451	0.000663	0.000714
Run 4 - One	0.000496	0.000719	0.000977
Run 4 - Two	0.000500	0.000725	0.000985
Run 4 - Three	0.000497	0.000721	0.000979
Run 4 - Four	0.000498	0.000722	0.000981