

# Integrando Avaliações Textuais de Usuários em Recomendação baseada em Aprendizado por Reforço

Naan Vasconcelos  
naan.vasconcelos@aluno.ufsj.edu.br  
UFSJ  
Minas Gerais, Brazil

Nícollas Silva  
ncsilvaa@dcc.ufmg.br  
UFMG  
Minas Gerais, Brazil

Davi Reis  
davireisjesus@aluno.ufsj.edu.br  
UFSJ  
Minas Gerais, Brazil

Washington Cunha  
washingtoncunha@dcc.ufmg.br  
UFMG  
Minas Gerais, Brazil

Thiago Silva  
thiagosilva@ufsj.edu.br  
UFSJ  
Minas Gerais, Brazil

Elisa Tuler  
etuler@ufsj.edu.br  
UFSJ  
Minas Gerais, Brazil

Leonardo Rocha  
lrocha@ufsj.edu.br  
UFSJ  
Minas Gerais, Brazil

## ABSTRACT

In Multi-Armed-Bandit (MAB) approaches for Recommendation Systems, items are represented as arms to be recommended and the goal is to maximize the expected user's satisfaction (i.e., reward). Despite the reward often being the ratings explicitly assigned by the user, in other scenarios, implicit ratings extracted from user comments by review-aware recommendation systems (RARs) may efficiently elucidate the user's preferences. In this paper, we provide a preliminary study of the impact of using these implicit ratings instead of explicit ones in MAB approaches. Our results point out that implicit ratings decrease the entropy of the datasets, negatively impacting the performance of MAB.

## KEYWORDS

Sistema de recomendação consciente de comentário textual, MAB

## 1 INTRODUÇÃO

Sistemas de Recomendação (SsR) deixaram de ser apenas algoritmos *offline*, que fazem previsões em *batch*, para se tornarem interativos e responsáveis por orientar toda a jornada de interações de usuários com aplicações [33]. Um SsR deve escolher continuamente um item (ou um conjunto de itens), receber o *feedback* do usuário (avaliação do item) e atualizar-se a cada iteração. Por essa razão, trabalhos recentes abordam a tarefa de recomendação com um modelo de Aprendizado por Reforço representada por *Multi-Armed Bandits* (MAB) [3]. Nesse caso, um conjunto fixo e limitado de itens (*arms*) deve ser selecionado para maximizar o ganho esperado (*reward*). Em cada tentativa, o sistema deve decidir quais *arms* selecionar, qual ordem selecionar e se deve continuar ou tentar uma opção diferente. Por esta razão, o MAB normalmente enfrenta o dilema

de escolher entre o *arm* potencialmente mais relevante até o momento (e.g. item popular) para satisfazer ao usuário imediatamente (*exploitation*) ou optar por um *arm* distinto (i.e. item com diferentes avaliações - alta entropia) capaz de agregar mais conhecimento ao sistema (*exploration*) [25].

As abordagens MAB atuais são dependentes de *feedback* (*ratings*) definidos de forma explícita. Todavia, existe uma outra área de pesquisa em recomendação com atenção para as avaliações textuais (*reviews*) fornecidas pelos próprios usuários no intuito de obter *feedback* de forma implícita. Tais modelos são conhecidos como *Review Aware Recommender Systems* - (RARs) [1] e são projetados para melhorar a precisão das recomendações personalizadas ao utilizar as avaliações e opiniões dos usuários como fonte de informação. A principal hipótese desses modelos é que as opiniões dos usuários expressas por meio de texto tendem a ser mais fidedignas à sua real percepção e à relevância de um item [12, 17]. RARs vêm sendo impulsionadas pelos avanços em arquiteturas de redes neurais para tarefas de aprendizado profundo aplicadas a linguagem natural (PLN) [8, 9, 30]. Apesar dos avanços em termos de efetividade e diversidade por meio do uso de *feedbacks* implícitos oriundos de *reviews*, até onde sabemos, nenhuma proposta de MAB os tem utilizado.

Neste trabalho, investigamos preliminarmente o impacto do uso de *ratings* extraídos de comentários de usuários em SsR interativos baseados em MAB. Avaliamos experimentalmente o desempenho de três estratégias clássicas de MAB ( $\epsilon$ -Greedy, UCB e Thompson Sampling) [21] usando duas coleções de dados populares, comparando as versões originais com versões modificadas, onde os *ratings* foram calculados a partir de comentários textuais por quatro estratégias diferentes (MAN, CARM, HRDR e CARP) [4]. Comparando os resultados sob duas perspectivas, efetividade (*hits*) e diversidade (*user coverage*), os melhores foram obtidos quando as estratégias de MAB utilizavam as bases originais. Por se tratar de um resultado que contrasta com o que é reportado na literatura [10, 17], realizamos uma caracterização comparativa das coleções de dados. Enquanto nas coleções originais identificamos que os *ratings* assinalados são mais dispersos entre as possíveis valores (1-5), nas bases modificadas os *ratings* estavam concentrados entre 4 e 5. Apesar dos novos valores

estarem mais de acordo com as preferências dos usuários [11, 17], provocaram um desequilíbrio no dilema de *exploration/exploitation*: uma redução significativa na entropia de *ratings* impacta no *exploration*. Esses resultados apontam que para que MAB possam utilizar RARS com sucesso, suas estratégias de *exploration* precisam buscar alternativas à entropia.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 Multi-Armed Bandits

*Multi-Armed Bandits* (MAB) é a principal abordagem de Aprendizado por Reforço para SsR interativa [15]. Trata-se de um modelo de decisão sequencial que, continuamente, escolhe uma ação  $a$  entre um conjunto de ações  $\mathcal{A}$ , denominadas *arms*. A cada seleção de uma ação em um ponto  $t$  no tempo resulta em uma recompensa  $R(a_t) \in \mathbb{R}$ . Esses modelos precisam decidir entre: 1) *exploitation*, que visa selecionar os *arms* com as maiores recompensas do passado; e 2) *exploration*, que seleciona diferentes *arms* para obter mais informações sobre o domínio e tomar decisões futuras melhores. A escolha entre essas duas opções caracterizam o dilema de *exploitation-exploration* (i.e. *exp-exp*) e exige que o modelo seja capaz de explorar o máximo conhecimento disponível enquanto também explora o espaço de solução para adquirir ainda mais conhecimento sobre o domínio [32]. Em RS, os itens a serem recomendados são modelados como os *arms* e a recompensa é o *feedback* do usuário para essa recomendação (e.g., *rating*) [19]. Essa resposta dos usuários é, em geral, coletada e salva continuamente em um conjunto  $\mathcal{H}(t)$ . Um item é recomendado de acordo com uma regra de predição  $\pi$ , definida como uma função que explora as informações atuais sobre o usuário:  $i(t) \equiv \pi(\mathcal{H}(t))$ . Por isso, a estratégia ideal deve maximizar as recompensas nas  $T$  interações:

$$i^*(\cdot) = \arg \max_{i(\cdot)} \sum_{t=1}^T \mathbb{E} [r_{u,i(t)} | t] \quad (1)$$

Para definir a utilidade dos itens para cada usuário, as principais abordagens utilizam uma formulação matricial probabilística via PMF (*Probabilistic Matrix Factorization*), modelando a distribuição de *rewards* pelos fatores latentes de usuários e itens, semelhantes aos métodos *model-based* [27]. A recompensa esperada é quase sempre modelada como o produto dos fatores latentes do usuário  $\mathbf{p}_u$  com os fatores do item  $\mathbf{q}_i$ . A função objetivo e dada da seguinte forma:

$$i^*(\cdot) = \arg \max_{i(\cdot)} \sum_{t=1}^T \mathbb{E} [r_{u,i(t)} | t] = \arg \max_{i(\cdot)} \sum_{t=1}^T \mathbb{E} [\mathbf{p}_u^\top \mathbf{q}_{i(t)} | t] \quad (2)$$

Os esforços atuais estão concentrados em como otimizar essa função objetivo, equilibrando o dilema de *exp-exp*. As abordagens tradicionais do MAB, como  $\epsilon$ -*Greedy*, *UCB* e *Thompson Sampling*, utilizadas no presente estudo, consideram essa função objetivo como regra de predição [26]. A diferença entre esses algoritmos está na maneira como eles controlam o dilema de *exp-exp* [2, 20–22]. Enquanto o  $\epsilon$ -*Greedy* explora a regra de predição com probabilidade  $(1 - \epsilon)$ , os algoritmos *UCB* e *Thompson Sampling* primeiro medem uma incerteza  $\Sigma$  em torno das informações disponíveis sobre o usuário e os itens. *Thompson Sampling* é um algoritmo

probabilístico que extrai os vetores de usuários e itens de uma distribuição normal definida pelas informações atuais disponíveis. No entanto, mesmo essa abordagem mede a recomendação com base na combinação de  $p_u$  e  $q_i$ .

### 2.2 Recomendações Review-Aware

Sistemas de Recomendação *Review-Aware* [23] partem da premissa que as preferências dos usuários sobre itens consumidos podem ser melhor capturadas por meio das avaliações textuais (*reviews*) fornecidas pelos usuários ao invés dos *ratings* diretamente assinalados [12, 17]. Essa área vem sendo impulsionada pelos expressivos avanços em propostas de arquiteturas de redes neurais aplicadas a PLN [18]. Há um número crescente de SsR que vêm adaptando essas propostas de PLN, considerando diferentes abordagens de arquitetura, tais como redes *Attention-Based* [14], redes convolucionais temporais (TCN) [5], grafos [31], extração de aspectos [6] e tópicos [24], análise de sentimento [7] e, mais recentemente, aprendizagem contrastiva [29]. Recentemente, em [4] realizou-se uma revisão sistemática da literatura dessas abordagens, comparando-as experimentalmente.

Um algoritmo de destaque nessa avaliação é o **HRDR** [16], que transforma os comentários em representações vetoriais por meio de *embeddings*. Em seguida, aplica-se uma camada de convolução 2D para capturar padrões locais e, em seguida, um mecanismo de atenção para calcular pesos de atenção, focando em aspectos relevantes das representações. As representações ajustadas pelo mecanismo de atenção são combinadas, resultando em representações finais dos usuários e itens. De forma semelhante, o **CARP** [13] também vetoriza os comentários de usuários e itens que são processados por camadas CNN para capturar o contexto e extrair características importantes. O modelo calcula pesos de atenção para diferentes partes dos comentários usando mecanismos de autoatenção. Essas representações de usuários e itens capturam as variações e aspectos relevantes dos comentários. Outro algoritmo de destaque é o **CARM** [14] que também utiliza uma camada CNN sobre representações vetoriais para aprender características importantes dos comentários. As características convolucionais extraídas são combinadas com os fatores latentes dos usuários e itens, obtidos através da PMF. Essas características combinadas são então passadas por uma camada linear (*fully connected*) e uma função de ativação Tanh, transformando as características aprendidas em uma representação mais compacta. Por fim, o **MAN** [28] utiliza redes convolucionais temporais sobre as representações vetoriais para capturar padrões locais nos comentários. Uma função de atenção calcula uma matriz de atenção entre as representações dos usuários e itens, as quais são novamente processadas por uma camada convolucional temporal para ajustar os pesos da atenção. Este processo resulta em características de interação entre usuários e itens.

Uma característica comum às abordagens de MAB é que todas consideram os *rewards* de forma explícita, ou seja, por meio das *ratings* atribuídas pelos usuários aos itens recomendados. Nesse trabalho, nossa meta é alterar esses *rewards* explícitos por valores implícitos, extraídos dos *reviews* dos usuários por meio de abordagens acima descritas.

Dataset	MusicalInstruments			MusicalInstrumentsMAN			MusicalInstrumentsCARM			MusicalInstrumentsHRDR			MusicalInstrumentsCARP		
Measure	Hits			Hits			Hits			Hits			Hits		
T	5	10	20	5	10	20	5	10	20	5	10	20	5	10	20
Linear UCB	<b>0.267</b>	<b>0.408</b>	<b>0.584</b>	0.188	0.321	0.484	0.191	0.314	0.463	0.185	0.272	0.446	0.192	0.272	0.461
e-Greedy	<b>0.033</b>	<b>0.062</b>	<b>0.125</b>	0.026	0.053	0.106	0.025	0.051	0.102	0.025	0.051	0.103	0.025	0.047	0.105
TS	<b>0.122</b>	<b>0.187</b>	0.287	0.102	0.163	0.270	0.116	<b>0.187</b>	<b>0.301</b>	0.114	0.181	0.279	0.004	0.015	0.048
Measure	UsersCoverage			UsersCoverage			UsersCoverage			UsersCoverage			UsersCoverage		
T	5	10	20	5	10	20	5	10	20	5	10	20	5	10	20
Linear UCB	<b>0.068</b>	<b>0.145</b>	<b>0.228</b>	0.052	0.131	0.197	0.056	0.121	0.186	0.051	0.105	0.177	0.057	0.128	0.191
e-Greedy	<b>0.031</b>	<b>0.058</b>	<b>0.106</b>	0.025	0.050	0.092	0.025	0.049	0.092	0.024	0.047	0.089	0.023	0.044	0.090
TS	<b>0.101</b>	<b>0.140</b>	<b>0.191</b>	0.084	0.121	0.174	0.089	0.128	0.188	0.092	0.131	0.179	0.004	0.015	0.044

**Table 1: Musical Instruments - Efetividade e diversidade das estratégias MAB aplicadas nas coleções originais e suas versões modificadas. Valores em negrito correspondem aos melhores valores validados estatisticamente por Wilcoxon com p-value = 0.05**

Dataset	Tucson			TucsonMAN			TucsonCARM			TucsonHRDR			TucsonCARP		
Measure	Hits			Hits			Hits			Hits			Hits		
T	5	10	20	5	10	20	5	10	20	5	10	20	5	10	20
Linear UCB	<b>0.133</b>	<b>0.225</b>	<b>0.430</b>	0.062	0.142	0.212	0.080	0.151	0.230	0.082	0.147	0.233	0.072	0.136	0.208
e-Greedy	<b>0.023</b>	<b>0.056</b>	<b>0.118</b>	0.022	0.032	0.060	0.010	0.020	0.049	0.018	0.036	0.072	0.018	0.032	0.064
TS	<b>0.073</b>	<b>0.128</b>	<b>0.223</b>	0.057	0.090	0.143	0.044	0.078	0.128	0.052	0.091	0.147	0.045	0.076	0.133
Measure	UsersCoverage			UsersCoverage			UsersCoverage			UsersCoverage			UsersCoverage		
T	5	10	20	5	10	20	5	10	20	5	10	20	5	10	20
Linear UCB	<b>0.113</b>	<b>0.155</b>	<b>0.265</b>	0.058	0.128	0.170	0.073	0.128	0.167	0.077	0.125	0.171	0.065	0.117	0.156
e-Greedy	<b>0.023</b>	<b>0.054</b>	<b>0.104</b>	0.022	0.032	0.059	0.010	0.020	0.046	0.018	0.034	0.065	0.018	0.031	0.062
TS	<b>0.069</b>	<b>0.113</b>	<b>0.180</b>	0.056	0.087	0.131	0.042	0.071	0.111	0.050	0.084	0.128	0.044	0.072	0.118

**Table 2: Tucson - Efetividade e diversidade das estratégias MAB aplicadas nas coleções originais e suas versões modificadas. Valores em negrito correspondem aos melhores valores validados estatisticamente por Wilcoxon com p-value = 0.05**

### 3 AVALIAÇÃO EXPERIMENTAL

O objetivo dessa seção é responder a pergunta de pesquisa *Qual o impacto do uso de ratings extraídos de comentários de usuários e itens em SsR interativos baseados em MAB?* Para isso, nossos experimentos consistem em comparar o desempenho das estratégias MAB utilizando as bases de dados originais que contém *ratings* assinalados de forma explícita, com versões modificadas dessas mesmas coleções, onde os *ratings* são calculados a partir dos comentários textuais. Os modelos de MAB utilizados nos experimentos estão destacados na Seção 2.1: *ε-Greedy*, *UCB* e *Thompson Sampling*. Tratam-se das estratégias que são base de funcionamento da grande maioria das estratégias atuais. Para a geração das bases de dados modificadas, os cálculos dos *ratings* foram extraídos pelas quatro estratégias descritas na Seção 2.2: *MAN*, *CARM*, *HRDR* e *CARP*, escolhidos por apresentarem bons resultados em uma recente e vasta avaliação experimental [4].

Coleção	# Usuários	# Itens	Esparsidade
Amazon - Musical Instruments	27.530	10.620	99,92%
Yelp 2021 - Tucson	8.540	8.867	99,99%

**Table 3: Visão geral das coleções utilizadas nas avaliações.**

Avaliamos a combinação das estratégias de MAB e *Review-Aware* em duas coleções de cenários distintos, uma de comércio eletrônico e outra de pontos de interesse, conforme descritos na Tabela 3. Em ambas as coleções, todos os itens com *rating* também possuem um comentário textual correspondente. Para cada uma delas, foram geradas outras quatro versões modificadas pelas quatro estratégias de *Review-Aware*.

As três estratégias de MAB foram aplicadas sobre as cinco versões da base (uma original e quatro modificadas). Consideramos duas métricas distintas de avaliação, uma de precisão e outra de diversidade: 1) *Hits* métrica de precisão que corresponde ao número de recomendações que pertencem ao histórico de cada usuário; 2) *Users Coverage*, métrica de diversidade que corresponde à porcentagem de usuários distintos que estão interessados nos itens recomendados. Para comparar os resultados alcançados com validação estatística, adotamos o teste de *Wilcoxon*, visto a distribuição não normal dos conjuntos de dados de recomendação [4].

#### 3.1 Resultados Experimentais

As Tabelas 1 e 2 apresentam os resultados de nossa avaliação para as coleções *Musical Instruments* e *Tucson*, respectivamente. Os resultados remetem a média de *Hits* e *UserCoverage* acumulados após *T* interações dos usuários com o sistema. Tanto em efetividade quanto em diversidade, os melhores resultados foram alcançados pelas estratégias MAB nas coleções originais, nas quais os *ratings* foram assinalados de forma explícita pelos usuários. As diferenças entre os resultados alcançados nas bases originais e modificadas são ainda maiores à medida que ocorrem mais iterações. A principal razão é que as estratégias MAB aprendem menos sobre os usuários quando consideram bases com *ratings* calculados por estratégias RARs. Tratam-se de resultados que contrastam com aqueles reportados na literatura para outros cenários, por exemplo recomendação *offline*, que apontam *ratings* extraídos a partir dos *reviews* dos usuários são capazes de elucidar melhor as preferências dos usuários [12, 17, 28].

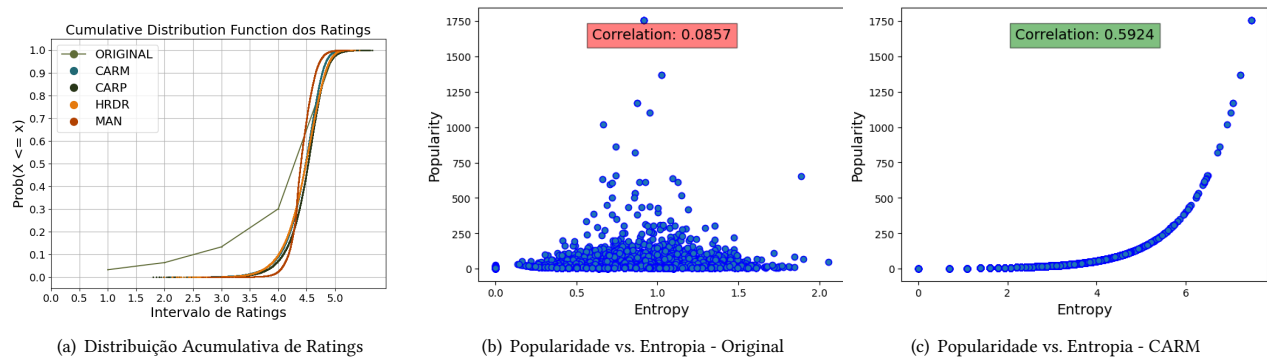


Figure 1: Caracterização Comparativa entre a coleção original Musical Instruments e suas versões modificadas por RARs.

Para compreender o comportamento das estratégias MAB nas coleções modificadas, realizamos uma caracterização comparativa delas com suas versões originais. Por restrição de espaço, os resultados serão reportados apenas para coleção *Musical Instruments*, mas as conclusões são similares na *Tucson*. Primeiramente, comparamos a distribuição de *ratings*, apresentadas na Figura 1(a). Observamos que enquanto na base original essa distribuição é bem dispersa entre os valores de 1 a 5, a distribuição está concentrada entre os valores de 3,5 a 5 nas bases modificadas. Os valores calculados estão coerentes com a literatura [12, 17], que aponta que os mesmos refletem de forma mais fidedigna as preferências dos usuários. Todavia, essa alteração causou um desequilíbrio em como as estratégias MAB lidam com o dilema *exp-exp*.

Assim, em nossa segunda caracterização, comparamos a correlação entre popularidade dos itens, comumente utilizada como uma estratégia de *exploitation*, e a entropia entre os itens, comumente utilizada como uma abordagem de *exploration*, conforme apresentado nas Figuras 1(b) - base original - e 1(c) - base modificada pelo algoritmo CARM (resultados semelhantes obtidos pelos demais algoritmos de RARS). Conforme podemos observar, na base modificada essa correlação é muito alta quando comparada à base original, o que significa que as bases modificadas apresentam uma baixa entropia, impactando diretamente no *Exploration*.

*Exploration* geralmente apresenta uma correlação significativa entre seu comportamento e a entropia dos itens. Em um cenário sem entropia, onde os *rewards* são mais previsíveis e não há incerteza, o comportamento dos algoritmos mudam significativamente. O *e-Greedy*, por exemplo, rapidamente converge para a *exploitation*, com um valor de  $\epsilon$  tendendo a zero. O UCB também reduz o *exploration*, pois os intervalos de confiança são muito estreitos devido à certeza das recompensas. No caso do Thompson Sampling, as distribuições posteriores se tornam muito concentradas em torno do valor verdadeiro da recompensa de cada opção. Isso resultaria em menos amostragem de opções não ótimas, pois a incerteza é praticamente inexistente. Portanto, a alteração da entropia dos itens nas bases modificadas impacta diretamente o equilíbrio entre *exploration* e *exploitation*, sendo crucial para o design de sistemas de recomendação baseados em algoritmos de aprendizado por reforço eficazes. Tais análises respondem a pergunta de pesquisa estudada neste trabalho.

## 4 CONCLUSÕES E TRABALHOS FUTUROS

Em Sistemas de Recomendação (SsR), Multi-Armed-Bandits (MAB) são modelos de decisão sequencial que continuamente escolhe entre um conjunto de itens, aqueles que maximizam o *reward* esperado (e.g., a satisfação do usuário), visando equilibrar a seleção entre itens com as maiores recompensas no passado (*exploitation*) ou itens inexplorados (*exploration*). Esses modelos utilizam como *reward* valores numéricos explicitamente assinalados pelos usuários aos itens. Neste trabalho apresentamos um estudo preliminar do impacto do uso de *ratings* extraídos de comentários textuais por meio de recomendação *Review-Aware* (RARS) em SsR interativos.

Comparamos experimentalmente o desempenho de três estratégias clássicas de MAB utilizando coleções de dados com *ratings* assinalados de forma explícita (originais) com suas versões na qual os *ratings* eram calculados por meio de quatro estratégias de RARS. Os melhores resultados foram obtidos nas bases originais, contrariando com resultados reportados na literatura para outros cenários além de MAB. Nesse sentido, caracterizamos comparativamente as coleções e constatamos uma redução significativa na entropia de *ratings* dos itens nas bases modificadas. Em um cenário de entropia reduzida, a dinâmica entre *exploration* e *exploitation* é drasticamente simplificada, com os algoritmos rapidamente se concentrando em *exploitation* e reduzindo o aprendizado sobre os usuários. Dessa forma, nosso trabalho abre a possibilidade de novas propostas de alterações nas estratégias de *exploration*, que permitam que as abordagens MAB sejam também capazes de usufruir da habilidade de estratégias RARS em capturar melhor as preferências dos usuários. Também pretendemos estudar como combinar as duas os *ratings* explícitos com a informação obtida por meio das RARS.

## AGRADECIMENTOS

Este trabalho foi financiado por CNPq, CAPES, Fapemig, FAPESP, CIIA-Saúde e AWS.

## REFERENCES

- [1] Sumaia Mohammed Al-Ghuribi and Shahrul Azman Mohd Noah. 2019. Multi-criteria review-based recommender system—the state of the art. *IEEE Access* 7 (2019), 169446–169468.
- [2] Yan Andrade, Nicollas Silva, Thiago Silva, Adriano C. M. Pereira, Diego Roberto Colombo Dias, Elisa Tuler de Albergaria, and Leonardo Rocha. 2023. A Complete Framework for Offline and Counterfactual Evaluations of Interactive Recommendation Systems. In *Proceedings of the 29th Brazilian Symposium on Multimedia*. <https://doi.org/10.1145/3617023.3617049>
- [3] Andrea Barraza-Urbina and Dorota Glowacka. 2020. Introduction to Bandits in Recommender Systems. In *Fourteenth ACM RecSys*. 748–750.
- [4] Guilherme Bittencourt, Guilherme Fonseca, Yan Andrade, Nicollas Silva, and Leonardo Rocha. 2023. A Survey on Review - Aware Recommendation Systems. In *Proceedings of the 29th Brazilian Symposium on Multimedia*. <https://doi.org/10.1145/3617023.3617050>
- [5] Yijin Cai, Yilei Wang, Weijin Wang, and Wenting Chen. 2022. RI-GCN: Review-aware Interactive Graph Convolutional Network for Review-based Item Recommendation. In *2022 IEEE International Conference on Big Data (Big Data)*.
- [6] Xu Chen, Zheng Qin, Yongfeng Zhang, and Tao Xu. 2016. Learning to rank features for recommendation over multiple categories. In *SIGIR*.
- [7] Jin Yao Chin, Kaiqi Zhao, Shafiq Joty, and Gao Cong. 2018. ANR: Aspect-based neural recommender. In *CIKM*.
- [8] Washington Cunha, Celso França, Guilherme Fonseca, Leonardo Rocha, and Marcos André Gonçalves. 2023. An effective, efficient, and scalable confidence-based instance selection framework for transformer-based text classification. In *SIGIR*.
- [9] Claudio MV de Andrade, Fabiano M Belém, Washington Cunha, Celso França, Felipe Viegas, Leonardo Rocha, and Marcos André Gonçalves. 2023. On the class separability of contextual embeddings representations—or “The classifier does not matter when the (text) representation is so good!”. *IP&M* (2023).
- [10] Vinicius HS Durelli, Rafael S Durelli, Andre T Endo, Elder Cirilo, Washington Luiz, and Leonardo Rocha. 2018. Please please me: does the presence of test cases influence mobile app users' satisfaction?. In *Proceedings of the XXXII Brazilian Symposium on Software Engineering*.
- [11] Lucas GS Félix, João Victor Silveira, Washington Luiz, Diego Dias, and Leonardo Rocha. 2018. Avaliação Automática de Conteúdo de Aplicações de Reclamação Online. In *Anais do VI Symposium on Knowledge Discovery, Mining and Learning*. SBC, 49–56.
- [12] Balraj Kumar and Neeraj Sharma. 2016. Approaches, issues and challenges in recommender systems: a systematic review. *Indian J. Sci. Technol* 9, 47 (2016), 1–12.
- [13] Chenliang Li, Cong Quan, Li Peng, Yunwei Qi, Yuming Deng, and Libing Wu. 2019. A capsule network for recommendation and explaining what you like and dislike. In *SIGIR*.
- [14] Duantengchuan Li, Hai Liu, Zhaoli Zhang, Ke Lin, Shuai Fang, Zhifei Li, and Neal N Xiong. 2021. CARM: Confidence-aware recommender model via review representation learning and historical rating behavior in the online platforms. *Neurocomputing* 455 (2021), 283–296.
- [15] Yanghao Li, Cuiling Lan, Junliang Xing, Wenjun Zeng, Chunfeng Yuan, and Jiaying Liu. 2016. Online human action detection using joint classification-regression recurrent neural networks. In *European Conference on Computer Vision*. Springer, 203–220.
- [16] Hongtao Liu, Yian Wang, Qiyao Peng, Fangzhao Wu, Lin Gan, Lin Pan, and Pengfei Jiao. 2020. Hybrid neural recommendation with joint deep representation learning of ratings and reviews. *Neurocomputing* 374 (2020), 77–85.
- [17] Washington Luiz, Felipe Viegas, Rafael Odon de Alencar, Fernando Mourão, Thiago Salles, Dárlinton B. F. Carvalho, Marcos André Gonçalves, and Leonardo Rocha. 2018. A Feature-Oriented Sentiment Rating for Mobile App Reviews. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web, WWW 2018*. ACM.
- [18] Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 1532–1543.
- [19] Javier Sanz-Cruzado, Pablo Castells, and Esther López. 2019. A simple multi-armed nearest-neighbor bandit for interactive recommendation. In *Proceedings of the 13th ACM Conference on Recommender Systems*. 358–362.
- [20] Nicollas Silva, Heitor Werneck, Thiago Silva, Adriano C. M. Pereira, and Leonardo Rocha. 2021. A contextual approach to improve the user's experience in interactive recommendation systems. In *WebMedia '21: Brazilian Symposium on Multimedia and the Web, Belo Horizonte, Minas Gerais, Brazil, November 5-12, 2021*, Adriano César Machado Pereira and Leonardo Chaves Dutra da Rocha (Eds.). ACM, 89–96. <https://doi.org/10.1145/3470482.3479621>
- [21] Thiago Silva, Nicollas Silva, Carlos Mito, Adriano C. M. Pereira, and Leonardo Rocha. 2022. Interactive POI Recommendation: applying a Multi-Armed Bandit framework to characterise and create new models for this scenario. In *WebMedia '22*. <https://doi.org/10.1145/3539637.3557060>
- [22] Thiago Silva, Nicollas Silva, Heitor Werneck, Adriano CM Pereira, and Leonardo Rocha. 2020. The impact of first recommendations based on exploration or exploitation approaches in recommender systems' learning. In *Proceedings of the Brazilian Symposium on Multimedia and the Web*. 173–180.
- [23] Mehdi Srifi, Ahmed Oussous, Ayoub Ait Lahcen, and Salma Mouline. 2020. Recommender systems based on collaborative filtering using review texts—a survey. *Information* 11, 6 (2020), 317.
- [24] Yunzhi Tan, Min Zhang, Yiqun Liu, and Shaoping Ma. 2016. Rating-boosted latent topics: Understanding users and items with ratings and reviews.. In *IJCAI*, Vol. 16. 2640–2646.
- [25] Huazheng Wang, Qingyun Wu, and Hongning Wang. 2016. Learning hidden features for contextual bandits. In *CIKM*. 1633–1642.
- [26] Huazheng Wang, Qingyun Wu, and Hongning Wang. 2017. Factorization bandits for interactive recommendation. In *Thirty-First AAAI Conference on Artificial Intelligence*.
- [27] Qing Wang, Chunqiu Zeng, Wubai Zhou, Tao Li, S Sitharama Iyengar, Larisa Shwartz, and Genady Ya Grabarnik. 2018. Online interactive collaborative filtering using multi-armed bandit with dependent arms. *TKDE* (2018).
- [28] Peilin Yang, Yingyuan Xiao, Wenguang Zheng, Xu Jiao, Ke Zhu, Chenchen Sun, and Li Liu. 2023. MAN: Main-auxiliary network with attentive interactions for review-based recommendation. *Applied Intelligence* 53, 10 (2023), 12955–12970.
- [29] Junliang Yu, Hongzhi Yin, Min Gao, Xin Xia, Xiangliang Zhang, and Nguyen Quoc Viet Hung. 2021. Socially-aware self-supervised tri-training for recommendation. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2084–2092.
- [30] Bruna Stella Zanotto, Ana Paula Beck da Silva Etges, Avner Dal Bosco, Eduardo Gabriel Cortes, Renata Ruschel, Ana Claudia De Souza, Claudio MV Andrade, Felipe Viegas, Sergio Canuto, Washington Luiz, et al. 2021. Stroke outcome measurements from electronic medical records: cross-sectional study on the effectiveness of neural and nonneural classifiers. *JMIR Medical Informatics* (2021).
- [31] Chenyan Zhang, Shan Xue, Jing Li, Jia Wu, Bo Du, Donghua Liu, and Jun Chang. 2023. Multi-Aspect enhanced Graph Neural Networks for recommendation. *Neural Networks* 157 (2023), 90–102.
- [32] Xiaoxue Zhao, Weinan Zhang, and Jun Wang. 2013. Interactive collaborative filtering. In *CIKM*.
- [33] Sijin Zhou, Xinyi Dai, Haokun Chen, Weinan Zhang, Kan Ren, Ruiming Tang, Xiuqiang He, and Yong Yu. 2020. Interactive recommender system via knowledge graph-enhanced reinforcement learning. In *SIGIR*. 179–188.