

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI

Rodrigo de Carvalho Santos

**Estratégias para Aprimorar a Diversidade
Categorica e Geográfica de Sistemas de
Recomendação de POIs**

São João del-Rei

2019

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI

Rodrigo de Carvalho Santos

Estratégias para Aprimorar a Diversidade Categórica e Geográfica de Sistemas de Recomendação de POIs

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação do Departamento de Ciência da Computação da Universidade Federal de São João del-Rei como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

Orientador: Leonardo Chaves Dutra da Rocha

Universidade Federal de São João del - Rei – UFSJ

Departamento de Ciência da Computação

Programa de Pós-Graduação em Ciência da Computação

São João del-Rei

2019

Ficha catalográfica elaborada pela Divisão de Biblioteca (DIBIB)
e Núcleo de Tecnologia da Informação (NTINF) da UFSJ,
com os dados fornecidos pelo(a) autor(a)

S237e Santos, Rodrigo de Carvalho.
Estratégias para Aprimorar a Diversidade
Categorica e Geografica de Sistemas de Recomendacao de
POIs / Rodrigo de Carvalho Santos ; orientador
Leonardo Chaves Dutra da Rocha. -- São João del-Rei,
2019.
91 p.

Dissertação (Mestrado - Ciência da Computação) --
Universidade Federal de São João del-Rei, 2019.

1. Mineração de Dados. 2. Sistemas de Recomendação.
3. Recomendação de Pontos de Interesse. I. Rocha,
Leonardo Chaves Dutra da, orient. II. Título.

Rodrigo de Carvalho Santos

Estratégias para Aprimorar a Diversidade Categórica e Geográfica de Sistemas de Recomendação de POIs

Dissertação apresentada como requisito para obtenção do título de mestre em Ciências no Curso de Mestrado do Programa de Pós Graduação em Ciência da Computação da UFSJ.

Trabalho aprovado. São João del-Rei, 12 de agosto de 2019:



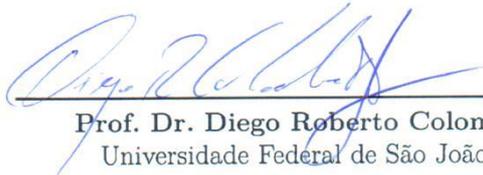
Prof. Dr. Leonardo Chaves Dutra da Rocha
Orientador



Prof. Dr. Adriano César Machado Pereira
Universidade Federal de Minas Gerais



Dr. Fernando Henrique de Jesus Mourão
Seek IA Labs



Prof. Dr. Diego Roberto Colombo Dias
Universidade Federal de São João del-rei

São João del-Rei
12 de agosto de 2019

Este trabalho é dedicado aos meus pais, irmãos, meu sobrinho Guilherme, minha esposa Rosany, meu filho João Francisco e à minha segunda mãe Nilza que não mediram esforços para que eu vencesse esta etapa da minha vida.

Agradecimentos

Primeiramente, agradeço a Deus, sempre, por tudo que acontece em minha vida. Agradeço também à minha esposa Rosany, que nestes dois anos não mediu esforços para me apoiar e incentivar. Sem dúvida, todas as suas demonstrações de amor e paciência tornaram-se força para que pudéssemos concluir esta etapa. De maneira especial, agradeço a meu pai pelo amor incondicional e por sempre ter acreditado em mim. Regracio aos meus irmãos, Rejane e Renan, pelos momentos de conversa e apoio. Além deles, agradeço à minha mãe e minha sogra pelo apoio de sempre. Também de maneira muito especial, agradeço ao meu sobrinho Guilherme pelo afeto e companheirismo e a meu filho, recém nascido, João Francisco.

Agradeço imensamente aos meus professores do Departamento de Ciência da Computação da UFSJ, com os quais tive a honra de conviver por dois anos. Em especial, ao meu orientador e amigo professor Leonardo Rocha, pelo companheirismo, pelos conselhos, pelas correções, pelo comprometimento com este trabalho, pela paciência, pelo imenso conhecimento compartilhado e pelas diversas demonstrações de amizade. Sem dúvida, seus ensinamentos e, principalmente essa amizade cultivada, pretendo levar por toda a vida.

Não menos importantes, agradeço aos colegas que conheci durante o curso e no LabPi: Diogo, Millas, Jardel, Thiago, Alan, Washington, Leandro, João, Christian, Antônio, Massilon, Sávyo e Vinícius. Agradeço também ao Luiz Felipe e Diego Carvalho pelos trabalhos desenvolvidos em conjunto. Em especial, agradeço ao meu colega Nícollas que, conheci há tão pouco tempo, mas que tornou-se fundamental para o desenvolvimento deste trabalho. Todas as horas que você dedicou a me ensinar o que sabia, todos os conselhos, as inúmeras leituras e releituras que fez sobre o que eu escrevia, a paciência em corrigir e ensinar; tudo isso jamais será esquecido e espero que a amizade criada se fortaleça ainda mais ao passar dos anos.

Por fim, agradeço aos colegas do trabalho pela convivência diária e pelo período que pude contar com a compreensão de todos enquanto dividia minha jornada entre trabalho e mestrado. Agradeço ao grande amigo de sempre, Davi pelas longas conversas, conselhos e orientações que ajudaram muito para que este projeto fosse concluído. A todas as outras pessoas que de alguma forma fazem parte da minha vida, e que porventura não as tenha citado, meu muito obrigado!

“Há três caminhos para o fracasso: não ensinar o que se sabe; não praticar o que se ensina; não perguntar o que se ignora.”

Venerável Beda

Resumo

Atualmente, os chamados Sistemas de Recomendação (SsR) têm sido usados para auxiliar usuários a descobrir Pontos de Interesse (POIs) relevantes em redes sociais baseadas em localização (*Location Based Social Network* - LBSN), tais como o Yelp e o FourSquare. Diante dos principais desafios de esparsidade de dados e diferentes fontes de informação, tais como a influência geográfica, nesse cenário, a maioria dos trabalhos sobre recomendações de POIs tem se concentrado apenas na melhoria da efetividade (i.e., acurácia) dos sistemas. No entanto, existe um crescente consenso de que apenas a efetividade não é suficiente para avaliar a utilidade prática desses sistemas. Em cenários reais, as diversidades categóricas e geográficas foram identificadas como as principais dimensões complementares para avaliar a satisfação dos usuários e a utilidade das recomendações. Os poucos trabalhos existentes na literatura estão concentrados em apenas um desses conceitos, isoladamente. Neste trabalho, propomos uma nova estratégia de pós-processamento, que combina esses conceitos a fim de melhorar a utilidade dos POIs recomendados para os usuários. Nesse contexto, o pós-processamento consiste em reordenar uma lista de POIs oriunda de um Recomendador Base (RB) a fim de maximizar as diversidades categórica e geográfica entre os primeiros itens da lista. Nossos resultados experimentais nos conjuntos de dados do *Yelp* mostram que nossa estratégia pode melhorar a utilidade da lista de POIs recomendados aos usuários, considerando diferentes SsR e múltiplas métricas de diversificação. Nosso método é capaz de melhorar a diversidade em até 120% sem grandes perdas de efetividade.

Palavras-chaves: Sistemas de recomendação. Pontos-de-interesse. Diversificação Categórica. Diversificação Geográfica.

Abstract

Nowadays, Recommender Systems (RSs) have been used to help users to discover relevant Points Of Interest (POI) in Location Based Social Network (LBSN), such as Yelp and FourSquare. In face to the main data sparsity challenges and different sources of information, such as geographical influence, in this scenario, most of works about POI recommendations has only focused on improving the system's accuracy. However, there is a consensus that just it is not enough to assess the practical effectiveness. In real scenarios, categorical and geographic diversities have been identified as key dimensions of user satisfaction and recommendation utility. The few existing works are concentrated on just one of these concepts, singly. In this work, we propose a novel postprocessing strategy to combine these concepts in order to improve the usefulness of recommended POIs. In this context, postprocessing involves reordering a POI list from a Base Recommender (RB) to maximize categorical and geographic diversities among the first items on the list. Our experimental results in Yelp datasets show that our strategy can improve the usefulness of the list of recommended user POIs, considering different RSs and multiple diversification metrics. Our method is able to improve the diversity up to 120% without major accuracy losses.

Keywords: Recommender Systems. Points of Interest. Categorical Diversity. Geographic Diversity.

Lista de ilustrações

| | |
|---|----|
| Figura 1 – Exemplo de modelagem da relação entre usuários e itens para ser utilizada por Sistemas de Recomendações. | 28 |
| Figura 2 – Visão geral da hierarquia de classes de técnicas de Sistemas de Recomendação | 29 |
| Figura 3 – Visão Geral do Funcionamento do Geo-Cat. | 47 |
| Figura 4 – Exemplo hipotético de uma distribuição de POIs sobre a divisão da cidade de Las Vegas em sub-regiões quadradas | 49 |
| Figura 5 – Gráfico com quantidades de Usuários, POIs e <i>check-ins</i> , após pré-processamento, das cidades de <i>Charlotte</i> , <i>Las Vegas</i> e <i>Phoenix</i> | 54 |
| Figura 6 – Gráfico que apresenta os valores médios de POIs por Usuários, <i>check-ins</i> por usuários e <i>check-ins</i> por POIs | 55 |
| Figura 7 – Gráfico de distribuição da quantidade de POIs por sub-regiões de cada cidade. | 56 |
| Figura 8 – Exemplo de uma árvore de categorias que ilustra como as categorias de POIs são organizadas na base do <i>Yelp</i> | 56 |
| Figura 9 – Gráfico que apresenta curvas de resultados do Geo-Cat sob diferentes combinações das possíveis configurações para os parâmetros λ e β | 61 |
| Figura 10 – Gráfico de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage da cidade de Charlotte. | 68 |
| Figura 11 – Gráfico de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage da cidade de Las Vegas. | 69 |
| Figura 12 – Gráfico de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage da cidade de Phoenix | 70 |
| Figura 13 – Gráficos de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage | 73 |

Lista de tabelas

| | | |
|----------|---|----|
| Tabela 1 | – Pós-Processamento do Geo-Cat sobre as listas geradas pelos recomendadores base para a cidade de Charlotte. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatisticamente significativos e ▼ representa perdas estatisticamente significativas. Geo-Cat consegue melhorar todos recomendadores base em termos de diversidade, sem que ocorra perda da efetividade. | 63 |
| Tabela 2 | – Pós-Processamento do Geo-Cat sobre as listas geradas pelos recomendadores base para a cidade de Las Vegas. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatisticamente significativos e ▼ representa perdas estatisticamente significativas. Geo-Cat consegue melhorar todos recomendadores base em termos de diversidade, sem que ocorra perda da efetividade. | 64 |
| Tabela 3 | – Pós-Processamento do Geo-Cat sobre as listas geradas pelos recomendadores base para a cidade de Phoenix. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatisticamente significativos e ▼ representa perdas estatisticamente significativas. Geo-Cat consegue melhorar todos recomendadores base em termos de diversidade, sem que ocorra perda da efetividade. | 65 |
| Tabela 4 | – Aplicação da MAUT para selecionar o melhor SR base entre os SsR USG, GeoMF e GeoSoCa. Conforme podemos observar, o USG é aquele que apresenta o melhor resultado sob as várias dimensões de qualidade. | 66 |
| Tabela 5 | – Comparativo entre as <i>Baselines</i> e o Geo-Cat para a cidade de Charlotte. A avaliação estatística foi feita utilizando-se o teste de <i>Wilcoxon</i> com 95% de intervalo de confiança. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatísticos e ▼ representa perdas estatísticas. | 68 |
| Tabela 6 | – Comparativo entre as <i>Base Lines</i> e o Geo-Cat para a cidade de Las Vegas. A avaliação estatística foi feita utilizando-se o teste de <i>Wilcoxon</i> com 95% de intervalo de confiança. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatísticos e ▼ representa perdas estatísticas. | 69 |
| Tabela 7 | – Comparativo entre as <i>Base Lines</i> e o Geo-Cat a para cidade de Phoenix. A avaliação estatística foi feita utilizando-se o teste de <i>Wilcoxon</i> com 95% de intervalo de confiança. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatísticos e ▼ representa perdas estatísticas. | 70 |

| | |
|---|----|
| Tabela 8 – Aplicação da MAUT como ferramenta para selecionar o melhor algoritmo de diversificação para a cidade de Charlotte. | 71 |
| Tabela 9 – Aplicação da MAUT como ferramenta para selecionar o melhor algoritmo de diversificação para a cidade de Las Vegas. | 72 |
| Tabela 10 – Aplicação da MAUT como ferramenta para selecionar o melhor algoritmo de diversificação para a cidade de Phoenix. | 72 |

Lista de abreviaturas e siglas

| | |
|-------|--|
| POI | Point of Interest |
| SsR | Sistemas de Recomendação |
| LBSN | Location-based Social Networks |
| GPS | Global Positioning System |
| MMR | Maximal Marginal Relevance |
| MAUT | Multi-Attribute Utility Theory |
| CB | Content-based |
| CF | Collaborative filtering |
| MAP | Mean Average Precision |
| MAE | Mean Absolute Error |
| RMSE | Root of Mean Square Error |
| EPC | Expected Popularity Complement |
| ILD | Intra-list Distance |
| GC | Genre Coverage |
| TWCF | Time Weight Collaborative Filtering |
| LRT | Location Recommendation Framework |
| LFBCA | Location-friendship BookmarkColoring Algorithm |
| FCF | Friend-based Collaborative Filtering |
| USG | User, social and Geographical influence |
| RB | Recomendador Base |
| KNN | k-Nearest Neighbors |

Sumário

| | | |
|----------|--|-----------|
| 1 | Introdução | 21 |
| 1.1 | Contextualização & Motivação | 21 |
| 1.2 | Objetivo | 22 |
| 1.2.1 | Justificativa | 23 |
| 1.2.2 | Hipóteses e Objetivos Específicos | 23 |
| 1.3 | Principais Contribuições | 24 |
| 1.4 | Organização do Documento | 25 |
| 2 | Referencial Teórico | 27 |
| 2.1 | Conceitos Básicos sobre Recomendação | 27 |
| 2.2 | Classes de Recomendadores | 28 |
| 2.2.1 | Filtragem Baseada em Conteúdo | 29 |
| 2.2.2 | Filtragem Colaborativa | 30 |
| 2.2.3 | Métodos híbridos | 31 |
| 2.3 | Recomendação de POIs | 32 |
| 2.4 | Principais Métricas de Qualidade e Avaliação | 33 |
| 3 | Trabalhos Relacionados | 38 |
| 3.1 | Efetividade em Recomendação de POIs | 38 |
| 3.2 | Diversidade em Sistemas de Recomendação | 41 |
| 3.3 | Diversificação Geográfica em Recomendação de POIs | 44 |
| 4 | Geo-Cat: Diversificação Geográfica e Categórica | 46 |
| 4.1 | Visão Geral | 46 |
| 4.2 | Diversificação Categórica | 47 |
| 4.3 | Diversificação Geográfica | 48 |
| 4.4 | Diversificação Geográfica-Categórica | 51 |
| 5 | Avaliação Experimental | 53 |
| 5.1 | Configuração Experimental | 53 |
| 5.1.1 | Base de Dados | 53 |
| 5.1.2 | Recomendadores Selecionados | 57 |
| 5.1.3 | Métricas de Avaliação | 58 |
| 5.2 | Teoria da Utilidade Multiatributo (MAUT) | 59 |
| 5.3 | Resultados & Discussões | 60 |

| | | |
|----------|---|-----------|
| 5.3.1 | Configurações de Parâmetros dos Métodos de Diversificação | 61 |
| 5.3.2 | Avaliação do Geo-Cat nos SsR Tradicionais | 62 |
| 5.3.3 | Avaliação do Geo-Cat com os Baselines | 66 |
| 6 | Conclusão & Trabalhos Futuros | 74 |
| | Referências | 79 |

1 Introdução

Neste capítulo, primeiramente apresentamos uma contextualização acerca do problema de recomendação de itens – mais especificamente sobre a recomendação de pontos de interesse (POIs). Em seguida, são descritas as principais características que diferenciam a recomendação de POIs das recomendações feitas nos cenários tradicionais. Também são expostos o objetivo principal deste trabalho e um detalhamento de como esse objetivo é abordado, dos objetivos específicos e das hipóteses relacionadas. Por fim, são apresentadas as principais contribuições deste trabalho.

1.1 Contextualização & Motivação

Nos últimos anos, vivenciamos uma grande expansão da *Internet* e, conseqüentemente, dos serviços que ela oferece (1). De maneira geral, o desenvolvimento dessa rede fez com que a *Web* deixasse de ser apenas uma plataforma para simples troca de documentos, para se tornar um ambiente completo com ampla variedade de aplicações. Por essa razão, é notório o aumento da quantidade de usuários que utilizam os serviços oferecidos na *Web*. Com isso, todos os dias milhões de pessoas a utilizam, produzindo, armazenando e consumindo uma enorme quantidade e variedade de dados. Entretanto, esse grande volume de dados e a vasta quantidade de informações disponíveis na *Web* geram um cenário desafiador: há mais dados do que efetivamente as pessoas conseguem analisar (2).

Um exemplo desse cenário pode ser observado em empresas de *Internet* que são especialmente afetadas pelo grande volume de informações que precisam processar e armazenar. A Google, por exemplo, processa mais de 24 petabyte de dados por dia. Além disso, cerca de 800 milhões de usuários mensais do YouTube enviam mais de uma hora de vídeo por segundo (2).

Ainda nesse contexto, aplicações de redes sociais também são bons exemplos de cenários que lidam com um enorme volume de dados constantemente. Por exemplo, o Facebook recebe mais de 10 milhões de fotos por hora e seus usuários clicam no botão “curtir” ou deixam um comentário aproximadamente três bilhões de vezes por dia. Igualmente alta, a quantidade de mensagens no Twitter excede 400 milhões de *tweets* por dia.

De forma análoga, empresas de comércio eletrônico, tais como Amazon e Netflix, também armazenam centenas de milhares de produtos distintos para serem oferecidos aos seus usuários (3, 4). Assim, se por um lado constatamos que há muita informação disponível aos usuários desses sistemas, por outro, essa grande variedade torna a simples seleção de um produto uma tarefa complicada.

Logo, aplicações que filtram as informações disponíveis, mostrando apenas o que é de interesse do usuário, têm se tornado cada vez mais importantes (5). Essas aplicações são denominadas Sistemas de Recomendação (SsR), sendo que o principal objetivo desses sistemas é facilitar a associação entre usuários e itens potencialmente relevantes ou de interesse (4). Em outras palavras, os SsR apoiam-se em informações conhecidas dos usuários ou dos itens para tentar prever a ordenação de um conjunto restrito de itens ou prever o valor de relevância de cada item, considerando os itens, um único usuário ou um grupo de usuários. No cenário da *Web*, encontramos bons exemplos de SsR em aplicações de comércio eletrônico como a Amazon – cuja tarefa é encontrar produtos que sejam mais relevantes aos usuários –, em provedores de conteúdo multimídia como Netflix (filmes) e Spotify (músicas) – visando encontrar conteúdos de maior interesse para os usuários –, e mesmo as redes sociais, como Facebook, em que originalmente são recomendadas notícias e novas amizades (3, 4, 6).

Recentemente, observamos a popularização das redes sociais baseadas em localização LBSN (7), tais como Yelp, Foursquare, Gowalla e o próprio Facebook. Essas aplicações possuem funcionalidades que permitem que usuários compartilhem sua localização geográfica com seus amigos (utilizando os dispositivos de *global positioning system*–GPS–integrados a celulares, *tablets*, etc.), além de possibilitar que realizem *check-ins* em determinados lugares e/ou utilizem de sua localização para divulgar produtos no *marketplace* (8). O registro de todas essas atividades representa um conjunto de *feedbacks* implícitos sobre as experiências dos usuários, gerando a oportunidade para criação de SsR que realizem recomendações de pontos de interesse (*point-of-interest* - POI) (3).

De maneira geral, a qualidade dos SsR é avaliada sobre dois aspectos: efetividade e diversidade. A efetividade, em suma, é mensurada por métricas que avaliam se o usuário consumiu ou não os itens recomendados, tais como *Precision* e *Recall*. Por outro lado, as métricas de diversidade calculam quanto os itens das listas de recomendações são diversos entre si, considerando para isso atributos inerentes dos itens (i.e., gêneros musicais, sinopses de filmes, etc). Mais especificamente, em recomendação de POIs, além da diversificação categórica, a diversidade geográfica dos POIs é tão importante quanto a categórica. Nesse sentido, a diversidade da distribuição geográfica dos POIs recomendados também é considerada importante para aumentar a utilidade dos POIs recomendados.

1.2 Objetivo

O objetivo principal desse trabalho é propor e avaliar novas estratégias capazes de prover diversidade nas listas de recomendações de POIs. Mais especificamente, nossa meta é que essas estratégias aprimorem tanto a diversificação geográfica, quanto a diversidade categórica dos itens recomendados, sem que a efetividade dos Sistemas de Recomendação seja comprometida.

1.2.1 Justificativa

Grande parte dos SsR em POIs tem se concentrado apenas na efetividade da tarefa de correspondência entre POIs e usuários (7, 9, 10, 11), ou seja, tais SsR buscam apenas recomendar POIs que têm maior probabilidade de serem frequentados pelos usuários, baseados no histórico de *check-ins* dos usuários. Entretanto, maximizar apenas a efetividade desses SsR leva a modelos superespecializados nas preferências dos usuários, recomendando apenas POIs semelhantes aos anteriormente visitados (12). Sobretudo, no cenário de POI, essa superespecialização pode ser ainda pior devido à questão geográfica, isto é, SsR tradicionais podem apresentar aos usuários POIs restritos a sub-regiões relacionadas a rotina do usuário (e.g., trabalho, casa, etc.) (13). Dessa forma, existe um crescente consenso de que apenas a efetividade não é suficiente para avaliar a utilidade prática e o valor das recomendações (12, 14, 15) e que dimensões complementares de qualidade precisam ser consideradas. Particularmente, a diversidade tem sido identificada como uma dimensão chave para medir a satisfação dos usuários e a utilidade de uma recomendação, tais como diversidades categórica e geográfica.

Encontramos na literatura alguns trabalhos focados na diversificação categórica de SsR em geral (16, 17, 18), no entanto, não são capazes de prover uma diversificação geográfica. Isso resulta na recomendação de POIs categoricamente “diversos” concentrados em uma pequena porção das sub-regiões em que o usuário alvo é mais ativo. No que se refere à diversificação geográfica, encontramos apenas um trabalho (9), mas também não foi capaz de prover uma diversificação categórica. Em suma, é importante que os SsR de POIs sejam assertivos, recomendando simultaneamente POIs efetivamente relevantes para os usuários e também geográfica e categoricamente diversificados, o que, potencialmente, pode impulsionar as vendas (15) e aumentar a utilidade dos POIs recomendados para o usuário.

1.2.2 Hipóteses e Objetivos Específicos

Elaborar estratégias que combinem a efetividade e diversidade é atualmente um desafio para os SsR, uma vez que elas representam conceitos desvinculados (19) mas, inversamente correlacionados. Se, por um lado, POIs que tendem a aumentar a acurácia dos SsR são aqueles mais populares entre os usuários, por outro, itens que podem prover uma maior diversidade são aqueles menos convencionais. A partir disso, definimos nossa primeira hipótese:

- **(H1): SsR de POIs devem explorar as diversidades categórica e geográfica, simultaneamente, sem perdas relevantes de efetividade.**

Para validar a hipótese **H1**, adaptamos o conceito de Relevância Máxima Marginal (*Maximal Marginal Relevance* - MMR), originalmente considerado na área de Recuperação

de Informação (20, 21, 22). Por meio do MMR, dada uma determinada consulta, é possível recuperar um conjunto de documentos (em nosso caso, POIs) mais diferentes entre si, mas ainda sim relevantes. Assim, para alcançar o objetivo principal, derivamos os seguintes objetivos específicos:

- **Objetivo específico 1:** Definir estratégias que sejam capazes de combinar diversidade categórica com diversidade geográfica;
- **Objetivo específico 2:** Modelar o *trade-off* entre efetividade e diversidade (categórica e geográfica) por meio do MMR, a fim de maximizar a diversidade e manter a relevância dos POIs;
- **Objetivo específico 3:** Estudar e propor estratégias que utilizem as relevâncias estimadas dos POIs recomendados por um SR base em conjunto com estratégias que combinem diversidade categórica e geográfica, por meio do modelo MMR.

O processo de avaliação de SsR constitui-se como um dos grandes desafios ainda em aberto. Atualmente, encontramos na literatura um vasto número de propostas de métricas para avaliação de SsR (23, 24, 25, 26, 16, 27, 28). Ainda assim, não existe um consenso sobre quais métricas de avaliação são necessárias para consolidar um novo SsR. Frequentemente, os pesquisadores se limitam a avaliar métricas inerentes às técnicas implementadas, não validando simultaneamente requisitos de qualidade e satisfação dos usuários (24). A partir disso, definimos nossa segunda hipótese:

1. **(H2): A utilidade dos SsR de POIs deve ser avaliada considerando múltiplas perspectivas de qualidade, tais como acurácia, diversidade categórica e diversidade geográfica.**

Para validar nossa hipótese **H2**, adaptamos a Teoria da Utilidade Multiatributo (*Multi-Attribute Utility Theory - MAUT*), utilizada no domínio de Teoria dos Jogos (29) para manipular problemas em que os resultados são dependentes de dois ou mais atributos (30). Em nosso trabalho, materializamos a MAUT como uma métrica que visa representar a eficácia de um método em função de diferentes dimensões de qualidade (31). Assim, derivamos mais um objetivo específico:

- **Objetivo específico 4:** Avaliar as estratégias propostas por meio da MAUT.

1.3 Principais Contribuições

Assim, destacamos como principais contribuições deste trabalho:

1. O *Geo-Cat*, um novo método de diversificação *geográfica e categórica*, que visa utilizar ambos os conceitos na recomendação de POIs. Trata-se de um método de pós-processamento a ser aplicado sobre uma lista de um recomendador base, no intuito de maximizar a diversidade e manter a relevância dos POIs (i.e., acurácia). Esse *trade-off* é modelado de forma semelhante à Relevância Máxima Marginal (MMR). Enquanto a relevância é medida com base no *score* atribuído pelo recomendador base, a diversidade é uma combinação linear de dois fatores: a diversidade geográfica (9) e a diversidade categórica (16).
2. A adaptação da Teoria da Utilidade Multiatributo (MAUT), originalmente proposta para a área de Teoria dos Jogos (29), para o contexto de recomendação em que SsR possam ser avaliados sob diferentes perspectivas em conjunto. Basicamente, consolidamos a MAUT como uma métrica em que cada dimensão de qualidade é considerada como um atributo. Para cada atributo, realiza-se uma normalização dentre as pontuações alcançadas por todas as estratégias avaliadas. Em seguida, somam-se essas pontuações normalizadas alcançadas por cada estratégia, sendo que uma ponderação pode ser aplicada de acordo com a importância da métrica considerada (31).
3. Uma vasta avaliação experimental dos principais trabalhos de diversificação de SsR de POIs existentes na literatura. Em nossos resultados, demonstramos que o *Geo-Cat* é o método que melhor equilibra todos os fatores qualidade avaliados, destacando-se frente aos principais métodos existentes na literatura.

1.4 Organização do Documento

O restante deste trabalho está organizado da seguinte forma. O Capítulo 2 apresenta um referencial teórico da tarefa de recomendação, apresentando alguns conceitos básicos e uma breve descrição dos cenários onde esses sistemas são utilizados. Além disso, definimos formalmente a tarefa de recomendação de POIs, destacando suas peculiaridades frente aos SsR tradicionais, e finalizamos apresentando as principais métricas de avaliação de SsR. No Capítulo 3, apresentamos os principais e mais referenciados SsR de POIs da literatura, cujo objetivo principal é sempre relacionado à efetividade das propostas. Apresentamos os principais trabalhos que visam abordar a satisfação dos usuários por meio de dimensões complementares de qualidade no contexto de SsR mais gerais, tais como diversidade e novidade. Finalizamos apresentando os trabalhos específicos de diversidade de recomendações em cenários de POIs. No Capítulo 4, primeiramente abordamos parcialmente nossa primeira hipótese de trabalho, apresentando nossa estratégia para se obter SsR mais diversos geográfica e categoricamente. Depois, detalhamos a estratégia de diversidade categórica utilizada, bem como a estratégia de diversidade geográfica e como combinamos esses dois conceitos. No Capítulo 5, abordamos a segunda hipótese do trabalho, apresentando nossa proposta

de adaptação da MAUT para o cenários de recomendação para que múltiplas perspectivas possam ser avaliadas simultaneamente. Além disso, consolidamos ambas as hipóteses de trabalho, apresentando uma vasta avaliação experimental comparando nossa estratégias com os principais trabalhos de diversificação de SsR existentes na literatura. Nessa avaliação, demonstramos o quão nossa estratégia é melhor frente aos demais trabalhos da literatura e como a MAUT pode ser importante em avaliações futuras de SsR. Por fim, o Capítulo 6 apresenta as conclusões deste trabalho, resumizando os resultados obtidos em cada capítulo. Nesse também, são apresentados alguns possíveis trabalhos futuros a serem realizados.

2 Referencial Teórico

Neste capítulo, apresentamos os conceitos básicos que tangem os SsR e mostramos um breve panorama dos cenários em que esses sistemas são utilizados. Primeiramente, definimos formalmente a tarefa de recomendação. Em seguida, é feita a especificação dos conjuntos de métodos clássicos de recomendação utilizados pela literatura: Filtragem Colaborativa (CF), Filtragem Baseada em Conteúdo (CB) e os Métodos Híbridos. A seguir, com base na definição apresentada anteriormente, definimos a tarefa de recomendação de POIs e suas peculiaridades. Por fim, são destacadas as principais métricas utilizadas para avaliar a qualidade de um SR.

2.1 Conceitos Básicos sobre Recomendação

Atualmente, os SsR têm se tornado uma ferramenta fundamental nos mais variados serviços para assistir usuários a “escolher bem” quais itens consumir (4). Em suma, provedores de serviço utilizam SsR para: (1) maximizar as vendas de itens, (2) vender itens mais diversificados, (3) aumentar a satisfação dos usuários, (4) fidelizar usuários, e (5) entender melhor o que os usuários desejam (4). Portanto, observamos que ferramentas capazes de filtrar as informações disponíveis, mostrando apenas itens que são de interesse do usuário, se tornam cada vez mais importantes (5). A ideia principal dos SsR é utilizar as várias fontes de dados disponíveis para tentar inferir quais são as melhores indicações para os clientes. Em geral, no cenário de recomendações, os clientes são chamados de “usuários”, enquanto os produtos a serem recomendados são chamados de “itens” (32, 4).

Formalmente, o problema de recomendação de itens pode ser formulado como se segue: Seja U o conjunto de todos os usuários e seja I o conjunto de todos os itens possíveis de serem recomendados. A tarefa de recomendação consiste em encontrar para cada usuário um subconjunto personalizado de itens $R_u \subset I$, com k itens, que maximiza a função $f(u, i)$ para cada $u \in U$ e cada $i \in R_u$; em que $f(u, i)$ é a função que mede a utilidade do item i para o usuário u . Em outras palavras, a recomendação também pode ser definida como o problema de estimar *ratings* (avaliação) para itens que ainda não foram consumidos pelo usuário (33).

De uma forma bem resumida, as avaliações feitas por um usuário alvo sobre itens consumidos no passado são utilizadas como informação base para estimar as avaliações para outros itens. Cada usuário u do conjunto U é interpretado como um perfil que inclui várias características do próprio usuário. Da mesma forma, cada elemento i do conjunto I é entendido como um conjunto de características de cada item (34). Além disso, a informação de quais usuários consumiram quais itens no passado também é considerada para cada usuário e item.

| | I ₁ | I ₂ | I ₃ | I ₄ | I ₅ | I ₆ | I ₇ | I ₈ |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| U ₁ | 3 | | 2 | | 3 | | 2 | |
| U ₂ | 1 | 2 | | 3 | | | 1 | 2 |
| U ₃ | ? | 3 | 1 | ? | 2 | ? | 3 | ? |
| U ₄ | | | 1 | 1 | | 3 | | 2 |
| U ₅ | 2 | | 2 | | 1 | 1 | 3 | |
| U ₆ | | 1 | | 2 | | 1 | 2 | |

Figura 1 – Exemplo de modelagem da relação entre usuários e itens para ser utilizada por Sistemas de Recomendações.

O princípio básico para se fazer recomendações é baseado nas relações existentes entre usuários e itens. A Figura 1 ilustra um exemplo de cenário em que novas recomendações precisam ser geradas para um usuário alvo. As linhas representam os usuários e as colunas, os itens existentes na base. Cada posição nessa tabela representa a relação entre um usuário u e um item i , sendo que essa relação pode ser expressada de forma implícita: (1) binária se um usuário consumiu ou não um determinado item (0 ou 1); (2) pela quantidade de vezes que um usuário consumiu cada um dos itens (adotado em nosso exemplo), ou ainda de forma explícita pela sua avaliação acerca de cada item. Nosso usuário alvo, destacado em vermelho, consumiu no passado os itens 2, 3, 5 e 7. Além disso, temos que o item 1 foi consumido pelos usuários 1, 2 e 5, o item 4 pelos usuários 2, 4 e 6, o item 6 pelos usuários 4, 5 e 6 e, por fim, o item 8 pelos usuários 2 e 4. Essas informações são “compreendidas” por um SR, capacitando-o para fazer predições sobre quais itens, ainda não consumidos, são mais relevantes para um usuário (32). Nas próximas seções, apresentamos as diferentes abordagens para explorar essas informações.

2.2 Classes de Recomendadores

Os SsR diferem-se entre si quanto à abordagem adotada para prever um conjunto de n itens para um usuário (25, 35, 4, 24, 32). De uma forma mais geral, os SsR podem ser classificados de duas maneiras: não-personalizados ou personalizados. Os sistemas não-personalizados não consideram as informações (consumo) existentes sobre o usuário alvo, ou mesmo essas informações podem não estar disponíveis. As recomendações são baseadas apenas nas preferências dos usuários em geral (36, 37). Um exemplo desses SsR são aqueles que assumem que os itens mais consumidos por todos os usuários representam as melhores recomendações.

Por outro lado, as recomendações personalizadas baseiam-se em informações conhecidas de um usuário para modelar um perfil para este e, a partir desse perfil, visa apresentar ao usuário os itens que são potencialmente mais relevantes. As recomendações personalizadas são comumente subdivididas em três subclasses de recomendações: (1) baseadas em conteúdo (*content-based* - *CB*); (2) filtragem colaborativa (*collaborative filtering* - *CF*); e (3) métodos híbridos (*hybrid approaches*) (37, 33, 4). De forma resumida, os algoritmos da classe CB recomendam itens com base nas características dos itens consumidos. De maneira diferente, os algoritmos da classe CF recomendam itens com base nos relacionamentos entre usuários. Esses relacionamentos podem ser estabelecidos por meio da similaridade de consumo entre usuários (5). Os métodos híbridos tentam melhorar a acurácia das recomendações, combinando técnicas de CB e CF (38). Na Figura 2 ilustramos a divisão de classes de recomendação existentes. Nas próximas seções, apresentamos um detalhamento sobre cada uma dessas classes.

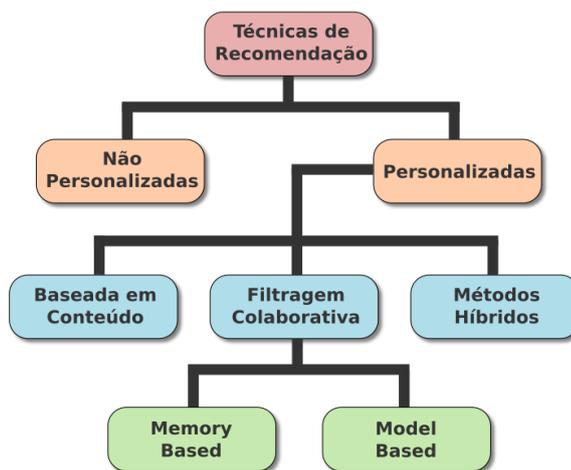


Figura 2 – Visão geral da hierarquia de classes de técnicas de Sistemas de Recomendação

2.2.1 Filtragem Baseada em Conteúdo

Técnicas de recomendações baseadas em CB são abordagens que visam tanto encontrar itens relevantes quanto estimar o *rating* que o usuário u atribuiria para o item i , baseado nas características descritivas de i (39). Por meio dessas características, encontram-se na base de dados os itens similares a i já consumidos pelo usuário u no passado. A partir dos *ratings* atribuídos por u a esses itens similares, estima-se o *rating* do item i (32, 36). De forma prática, considere, por exemplo, um SR baseado em CB para filmes. Se um usuário avaliou bem um filme de ficção, então o sistema tende a recomendar outros filmes do mesmo gênero que tenham conteúdo similares.

Do ponto de vista dos SsR baseados em CB, um item com n atributos é compreendido como um vetor $X_i = (x_1, x_2, \dots, x_n)$. Esses atributos podem ser binários, nominais ou numéricos. Por exemplo, em um SR para recomendação de POIs, os atributos po-

dem ser as categorias às quais o POI pertence, tais como restaurantes, bares, salões e outros (39). Da mesma forma, um usuário pode ser descrito formalmente como um vetor $Y_u = (y_1, y_2, \dots, y_n)$, em que cada elemento é uma combinação de atributos dos itens consumidos pelo usuário no passado. Logo, o objetivo dos SsR baseados em CB é tentar encontrar itens cujos atributos assemelhem com o perfil do usuário (24).

Uma das principais vantagens das técnicas baseadas em CB é sua habilidade em realizar recomendações de novos itens inseridos na base de dados, uma vez que o sistema estima o *rating* desses itens baseado em outros itens similares previamente avaliados pelo usuário (32, 33). No mesmo exemplo para cenário de filme, se um filme lançamento é inserido na aplicação e, mesmo que nenhum outro usuário tenha assistido esse filme, se as características dele forem similares a outros filmes assistidos por um usuário u , ele será recomendado. Todavia, técnicas de recomendação baseadas em CB também apresentam algumas desvantagens: (1) recomendar apenas itens muito parecidos com os que já foram consumidos pelo usuário. Por exemplo, se um usuário nunca assistiu um filme de terror, esse tipo de filme raramente terá chance de ser recomendado para o usuário; e (2) não faz boas recomendações para novos usuários, uma vez que os modelos de treinamento usam o histórico de avaliações do usuário para estimar os *ratings* (32) de itens. Logo, sem histórico não há como o modelo fazer previsões.

2.2.2 Filtragem Colaborativa

Em recomendações baseadas em CF, o sistema recomenda itens com base em medidas de similaridade, tais como correlação de *Pearson* ou similaridade de cossenos, entre os usuários ou itens (37). Dessa maneira, os itens recomendados para um usuário u são itens que foram bem avaliados por outros usuários que têm perfis semelhantes ao de u . Em síntese, essas técnicas de recomendação assumem que usuários com históricos de consumos parecidos têm interesses afins (5). Mais formalmente, define-se que a função de utilidade $f(u, p)$ do item p para o usuário u é estimada com base na função de utilidade $f(s, p)$ associada ao item p para o usuário $s \in S$, em que S representa o conjunto de usuários similares a u (33). Por exemplo, em um cenário de recomendação de POIs, um SR baseado em CF primeiramente encontra os usuários mais similares ao usuário u , ou seja, aqueles que frequentaram os mesmos POIs ou os mais similares aos frequentados pelo usuário u . Em seguida, apenas os POIs que foram melhor avaliados pelos usuários mais similares e que ainda não foram visitados por u são recomendados.

Dentro da subclasse de CF, podemos ainda dividir os algoritmos em outras duas subclasses: *memory-based* e *model-based* (40). Basicamente, os SsR *memory-based* utilizam métricas de similaridade para calcular a diferença entre usuários ou entre itens, com base nos itens previamente consumidos e avaliados pelos usuários (33), a fim de fazer a previsão dos *ratings*. O modelo de previsão pode ser feito focado nos usuários (*user-based*)

ou nos itens (*item-based*). Nos modelos focados no usuário, os *ratings* atribuídos pelos usuários mais similares ao usuário alvo u são agregados. Os modelos focados em item visam agregar os *ratings* recebidos pelos itens mais similares a um item que está sendo avaliado para recomendação ou não. Os SsR *model-based*, em sua maioria baseados em algoritmos de aprendizado de máquina, utilizam todos os *ratings* assinalados pelos usuários para realizar as previsões de recomendação. Dentre os algoritmos de aprendizado de máquina utilizados, aqueles baseados em decomposição de fatores latentes (5) são os mais utilizados na literatura. Assim como os *memory-based*, os modelos de previsão podem ser feitos focados nos usuários ou nos itens (34, 5).

Uma das maiores vantagens dos métodos CF é o potencial para descobrirem padrões implícitos que seriam dificilmente descobertos usando métodos de CB (41). Além disso, segundo (5), métodos CF são também capazes de alcançar itens novos e inesperados por meio de usuários vizinhos, melhorando a novidade e a diversidade nos SsR. Se por um lado esses métodos utilizam o “conhecimento” prévio para tentar melhorar a qualidade das recomendações, por outro, estes não são eficientes quando usados para fazer recomendações a novos usuários ou para recomendar novos itens que ainda não foram avaliados previamente (33). Nesses casos, não há como ter um “aprendizado” pela falta de informações prévias.

2.2.3 Métodos híbridos

Conforme supracitado nas seções 2.2.1 e 2.2.2, tanto as recomendações CB quanto as CF possuem pontos positivos e pontos negativos. Consequentemente, para mitigar os problemas causados pelos pontos negativos de ambas são desenvolvidos métodos híbridos de recomendação. Em outras palavras, métodos híbridos são aqueles que combinam técnicas de recomendação CB e CF para suavizar os pontos fracos de cada uma. Se por um lado, CF é mais indicada para recomendar itens já avaliados por outros usuários ou para recomendar itens para usuários que já têm um histórico de avaliações de outros itens, por outro lado, CB é a melhor estratégia para recomendar itens novos, que ainda não foram avaliados, ou para recomendar itens a novos usuários (32, 4, 37).

Por tudo isso, há na literatura quatro maneiras distintas de implementar CB e CF para criar métodos híbridos de recomendação:

1. Implementar CB e CF separados: primeiramente, métodos CF e CB são implementadas separadamente e em seguida os resultados das previsões são associados por meio de uma combinação linear dos *ratings*, ou ainda por uma função de qualidade que determina qual é a melhor recomendação a se utilizar (42, 43);
2. Incorporar características de CB em alguma técnica de CF: algumas características das técnicas de CB são adicionadas aos métodos de CF (i.e. integrar atributos de

conteúdo no perfil do usuário a fim de calcular a função de similaridade entre usuários ou itens) (33, 44);

3. Incorporar características de CF às técnicas de CB: algumas características das técnicas de CF são adicionadas aos métodos de CB (i.e. estratégias de redução de dimensionalidade para grandes coleções de conteúdos) (34, 45);
4. Construir um modelo de recomendação que integre técnicas de CB e CF: Incorporar em um método de recomendação tanto técnicas de CF, quanto as de CB (i.e. modelos capazes de usar informações de conteúdo dos itens e também analisar as preferências do usuário, baseando-se nos hábitos de consumo deste) (46, 34, 47, 33).

2.3 Recomendação de POIs

Formalmente, o problema de recomendação de POIs pode ser definido como um conjunto de n usuários $U = (u_1, u_2, \dots, u_n)$ e um conjunto de m POIs $P = (p_1, p_2, \dots, p_m)$. O conjunto de POIs visitados pelo usuário u é denotado por P_u . Além disso, é importante destacar que cada POI possui uma localização geográfica representada por latitude e longitude. As informações sobre as frequências de *check-ins* dos usuários em POIs são convertidas para uma matriz de frequências C . Cada entrada C_{up} em C representa a relação entre um usuário u e um POI p (48). Conseqüentemente, a frequência de *check-ins* representa as preferências dos usuários em vários POIs. Além disso, cada usuário pode ter uma lista de amigos, e dessa maneira as relações sociais são representadas por uma matriz S . Nesse caso, cada entrada S_{uv} expressa a relação social entre o usuário u e v . Na maioria dos casos, $S_{uv} = 1$ significa que existe relação social entre os usuários u e v ; 0 significa que não existe tal relação (49). Portanto, o principal objetivo de um SR de POIs é aprender sobre as preferências dos usuários de acordo com seus históricos de *check-in* e recomendar POIs interessantes para o usuário (49).

Dentro do domínio de recomendação, o cenário de POIs se difere dos demais devido às suas características peculiares. Em primeiro lugar, o problema de esparsidade (50) é ainda maior no domínio de POIs do que em outros cenários de recomendação, uma vez que o usuário frequenta apenas uma fração muito pequena de todos os POIs localizados em uma região (51, 3). Por exemplo, a esparsidade da base de dados da Netflix é cerca de 99%, enquanto que, a esparsidade da base de dados do Gowalla é em torno de $1 - (2.08 \times 10^{-4})$ (52). É importante destacar também que a fonte de informação sobre os usuários difere dos demais SR. Nesse cenário, as principais fontes de informação para medir o interesse do usuário são:

1. Influência Social: Assume que as opiniões dos amigos presentes na rede social do usuário são mais relevantes que a dos não-amigos (53, 54).

2. Influência temporal: Refere-se ao fato de que os usuários tendem a visitar lugares, por exemplo restaurantes, em um período esporádico de tempo (55, 56).
3. Influência Geográfica: É a principal característica a ser considerada, pois ela é intrínseca desse cenário de recomendação, uma vez que os usuários e POIs estão situados em um espaço físico (49).

Sobretudo, a influência geográfica é o fator mais importante que distingue a recomendação de POIs das recomendações tradicionais. Além disso, estudos apontam que usuários normalmente frequentam áreas restritas e preferem visitar POIs localizados próximos a outros POIs que eles já visitaram (48, 49). De modo semelhante, a influência temporal também é considerada importante para a recomendação de POIs porque as restrições físicas e de horários resultam em padrões de comportamento específicos dos usuários (48).

2.4 Principais Métricas de Qualidade e Avaliação

Konstan e Riedl (1999) sugerem que as abordagens existentes para avaliar SsR podem ser divididas em duas categorias: (1) *avaliação online*, em que o desempenho é avaliado nos usuários de um sistema de recomendação em execução; e (2) *avaliação offline*, em que o desempenho de um mecanismo de recomendação é avaliado em base de dados existentes. Em geral, nas avaliações *online* as recomendações são apresentadas para usuários em tempo real. Nesse caso, os usuários não precisam avaliar o item recomendado. Em vez disso, o SR observa com que frequência o usuário aceita uma recomendação. A principal vantagem da avaliação *online* é que ela mensura implicitamente a satisfação dos usuários. Porém, a avaliação online é difícil de ser realizada devido à necessidade de um sistema totalmente funcional e uma comunidade de usuários. Além disso, um teste online deve estar preocupado com a interface da aplicação, pois influencia a satisfação dos usuários e a intenção de fornecer *feedback*. Consequentemente, a maioria dos pesquisadores prefere a avaliação offline, devido a simplicidade e eficiência em executá-la. Nesse caso, usa-se base de dados existentes coletadas durante a execução de um sistema real (24). Contudo, avaliações *offline* não consideram fatores humanos durante a avaliação e consequentemente podem não ser recomendadas para alguns cenários.

Na avaliação offline, a recomendação pode ser vista como recuperação de informações, i.e. a seleção do subconjunto de ativos que são relevantes para o usuário. Nessa perspectiva, as métricas para avaliação são bem conhecidas. Em geral, os principais SsR existentes se preocupam em melhorar a acurácia sob a premissa de apresentar itens potencialmente úteis aos usuários (24). Para estes, as principais métricas de avaliação estão relacionadas aos conceitos de acurácia, precisão e revocação (23). Precisão e revocação visam quantificar informações sobre listas de recomendações geradas (33). Por sua vez, as métricas relacionadas a acurácia consistem em avaliar a predição realizada pelos SsR,

como no caso das métricas de MAE e RMSE (33). Todas essas métricas relacionadas à efetividade estão devidamente detalhadas como se segue:

- **Precision:** representa a probabilidade que um item recomendado seja relevante, sendo definida como a razão dos itens relevantes recomendados (N_{rs}) pelo número de itens da lista de recomendação (N) (23). Um item recomendado é considerado relevante quando pertence ao conjunto dos itens de testes do usuário.

$$Precision = \frac{N_{rs}}{N} \quad (2.1)$$

- **Recall:** representa a probabilidade de um item relevante ser recomendado, sendo definida como a razão dos itens relevantes recomendados (N_{rs}) pelo número de itens relevantes existentes (N_r) (23).

$$Recall = \frac{N_{rs}}{N_r} \quad (2.2)$$

- **F-measure (F1-Score):** representa a combinação desejada de *precision* e *recall*, por meio da média harmônica penalizada (24).

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.3)$$

- **Distribuição de Probabilidade de Precision e Recall:** medida estatística básica que nos permite verificar quantas vezes cada valor X de *precision* e *recall* ocorre durante o processo de avaliação. Basicamente, para cada valor existente, contabilizamos o número de vezes em que ele aparece e analisamos essa distribuição de probabilidade. A função $f(u)$ recebe 1, caso o valor de *precision* e/ou *recall* seja igual a X , e 0, caso contrário.

$$P(X) = \sum_{u \in U} f(u) \quad (2.4)$$

- **Função de Distribuição Acumulada de Precision e Recall:** medida estatística básica conhecida como CDF (*Cumulative Distribution Function*), que possibilita uma visão geral das distribuições. Basicamente, avalia-se o número de usuários para os quais são atribuídos valores de *precision* e *recall* maiores ou iguais a um limiar X que varia de 0 (menor valor) a 1 (maior valor).

$$F(x) = |Precision(u) \leq x| \quad F(x) = |Recall(u) \leq x| \quad (2.5)$$

- **Average Precision:** representa a precisão média obtida por um conjunto de k itens recomendados, considerando a relevância dos itens, para um usuário u . Basicamente,

somam-se todos os valores de *precision* obtidos para os k itens relevantes dentre todos os n itens avaliados, normalizando pelo número total de itens relevantes no domínio (I^+). (23).

$$AP(u) = \sum_{i=1}^n \frac{Precision@k \times rel(i)}{|I^+|} \quad (2.6)$$

- **Mean Average Precision (MAP):** consiste em calcular a média dos valores de *Average Precision*, considerando todas as listas de recomendações R geradas para os m usuários do domínio. (23).

$$MAP = \frac{1}{m} \sum_{u=1}^m AP(R_u) \quad (2.7)$$

- **Hit Rate (Acurácia):** representa a quantidade de itens recomendados corretamente dentre o conjunto de itens recomendados R . Essa métrica é avaliada com base no conjunto de itens consumidos (I_{teste}) pelo usuário alvo que estão presentes no conjunto teste. (25).

$$Hits(u) = |R_u \cap I_{teste}| \quad (2.8)$$

- **Mean Absolute Error (MAE):** representa a distância entre o *rating* real (r), atribuído pelo usuário u ao item i , para o *rating* previsto pelo recomendador (\hat{r}) (26). Para esta métrica leva-se em conta todos os n itens recomendados para o usuário.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{r}_i - r_i| \quad (2.9)$$

- **Root Mean Squared Error (RMSE):** representa a distância quadrática entre o *rating* real (r) para o *rating* previsto pelo recomendador (\hat{r}). A diferença dos valores quadrados é utilizada a fim de penalizar erros maiores (26).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{r}_i - r_i)^2} \quad (2.10)$$

Recentemente, observamos uma crescente preocupação, por parte dos pesquisadores, de que apenas essas métricas de acurácia não são suficientes para avaliar a eficácia prática das recomendações (12, 57). Na maioria dos casos, o objetivo de recomendação é inerentemente ligado a uma noção de descoberta, uma vez que a recomendação traz maiores benefícios quando esta expõe o usuário a uma experiência relevante, que ele não teria encontrado sozinho (33). Nesse contexto, métricas que remetem surpresa são características altamente desejáveis, pois refletem a satisfação dos usuários. Esses conceitos de surpresa estão relacionados às métricas de novidade, diversidade e serendipidade (58). Apesar de existirem diversas implementações para as métricas de novidade e diversidade,

grande destaque pode ser dado ao *framework* proposto em Vargas e Castells (2011). Por outro lado, o conceito de serendipidade relacionado a recomendação está bem consolidado em Zhang et al. (2012). Todas essas métricas estão relacionadas à satisfação dos usuários e tratam-se de dimensões de qualidade complementares. São elas:

- **Novidade:** se refere a quão diferente um item recomendado é com relação a todos os outros previamente consumidos pelo usuário alvo. É mensurado por meio da distância entre os itens recomendados e os itens do perfil de cada usuário, como proposto no *framework* de Vargas e Castells (2011). Nessa métrica, utiliza-se a similaridade de *Pearson*, medindo a relação entre a covariância e variância de dois itens.

$$nov(R|u) = EPC = C \sum_{i_k \in R} disc(k)p(rel|i_k, u)(1 - p(seen|i_k)) \quad (2.11)$$

- **Diversidade:** se aplica a um conjunto de itens, medindo quão diferente os itens são com relação uns aos outros (4). Assim como em (16), diversidade é calculada como a distância média esperada de um item para uma lista de itens (ILD), calculada como o complemento da similaridade dos itens recomendados.

$$ILD = div(R) = \frac{2}{|R|(|R| - 1)} \sum_{i_k \in R, l < k} d(i_k, i_l) \quad (2.12)$$

- **Serendipidade:** é uma forma de mensurar o quão surpreso o usuário ficou com o sucesso das recomendações. Calculamos a serendipidade por meio do complemento da similaridade de cosseno dos itens presentes no histórico de um usuário e as recomendações geradas (27). Os valores mais baixos indicam que as recomendações desviam do comportamento tradicional de um usuário, e, portanto, trazem maior surpresa.

$$ser = 1 - \sum_{u \in S} \frac{1}{|S||H_u|} \sum_{h \in H_u} \sum_{i \in R_{u,20}} \frac{CosSim(i, h)}{20} \quad (2.13)$$

- **Cobertura de Catálogo:** é a fração dos itens relevantes recomendados pelo menos uma vez, levando em conta todos os usuários. Valores mais altos de cobertura do catálogo indicam que o algoritmo contrabalança o viés de popularidade, cobrindo uma grande parte do conjunto geral de itens (28).

$$CC = \frac{|\bigcup_{u \in U} R^+|}{|U|} \quad (2.14)$$

- **Cobertura de Gênero (*Genre Coverage- GC*):** é proporção média de gêneros (i.e., características dos itens) relevantes recomendados ao usuário, a fim de medir os interesses do usuário (28). Essa métrica é definida para o conjunto de usuários U como a média do número de gêneros relevantes para cada usuário, em que: I_u^+ é o conjunto

de itens relevantes avaliados por u ; R_u é o conjunto de itens recomendados para u ; e $genres(i)$ são os gêneros associados ao item i .

$$GC = \frac{1}{|U|} \sum_{u \in U} \frac{|U_{i \in R_u} genres(i) \cap U_{i \in I_u^+} genres(i)|}{|U_{i \in I_u^+} genres(i)|} \quad (2.15)$$

Todas as métricas acima apresentadas são comuns em avaliações de SsR de qualquer contexto. Entretanto, conforme previamente mencionado, no cenário de POIs temos ainda a especificidade de que é importante considerar se as recomendações não estão superespecializadas a sub-regiões específicas, tais como aquelas relacionadas a rotina do usuário (e.g., trabalho, casa, etc.). Dessa forma, recentemente uma nova métrica que avalia a diversidade geográfica das recomendações é proposta:

- **Diversidade Geográfica Proporcional (PR_g):** é a proporcionalidade da distribuição geográfica dos POIs em sub-regiões, a fim de mensurar se os POIs recomendados estão distribuídos proporcionalmente nas subáreas de atividade do usuário alvo (ua) (9). Basicamente, essa métrica é o complemento da desproporcionalidade $DP_g(P)$, proposta por *Dang et al.* (17), normalizada pela proporcionalidade máxima possível ($|T_u| + 1/2$). Dessa forma, valores mais altos de PR_g indicam que os POIs recomendados estão proporcionalmente distribuídos pelas subáreas de atividade de ua ; por outro lado, valores baixos de PR_g podem ser causados por dois motivos: (1) os POIs recomendados estão distribuídos geograficamente em subáreas aleatórias que não têm relação com as subáreas de atividade do ua ; e/ou (2) os POIs recomendados estão localizados em poucas subáreas e não apresentam diversidade geográfica. Conforme veremos no Capítulo 4, essa métrica é a base para o funcionamento da estratégia de diversidade geográfica proposta em Hay e Yamana (2017).

$$PR_g(P) = 1 - DP_g(P)/(|T_u| + 1/2) \quad (2.16)$$

3 Trabalhos Relacionados

Nesse capítulo, inicialmente apresentamos uma revisão bibliográfica dos principais e mais referenciados SsR de POIs da literatura. Em todas as propostas, o foco da avaliação é sempre com relação da efetividade dos algoritmos propostos frente a outros trabalhos. Em seguida, apresentamos as principais estratégias que abordam a satisfação dos usuários por meio de dimensões complementares de qualidade em recomendação, tais como novidade, diversidade e serendipidade, relacionados a SsR em geral. Por fim, finalizamos este capítulo, apresentando trabalhos específicos de diversidade de recomendações em cenários de POIs.

3.1 Efetividade em Recomendação de POIs

Conforme descrito na seção 2.3, pesquisadores diferem os SsR tradicionais dos SsR de POIs por meio de características consideradas peculiares para esse tema (49) (i.e. matrizes demasiadamente esparsas, localização geográfica). Nesse sentido, embora existam diversas abordagens eficazes de recomendação para os cenários tradicionais, elas não são igualmente efetivas para o caso da recomendação de POIs. Por essa razão, encontramos na literatura diversos trabalhos com propostas de SsR específicos para POIs, cujas principais fontes de informações para medir o interesse do usuário sobre os POIs são: (1) influência temporal; (2) influência social; e (3) influência geográfica.

Encontramos na literatura alguns trabalhos que exploram a influência temporal em recomendação de POIs, com destaque para dois deles: Ding e Li (2005) e Gao et al. (2013). O primeiro (59) é um dos pioneiros na exploração da influência temporal em recomendação de POIs. Nele, os autores apresentam o *Time Weight Collaborative Filtering-TWCF*, um método de filtragem colaborativa temporal baseado no usuário, que calcula a similaridade entre os usuários, ponderando os *check-ins* com uma função de decaimento ao longo do tempo. No segundo trabalho (60), os autores assumem a premissa de que usuários com históricos de consumo semelhantes, em um mesmo período de tempo compartilham dos mesmos interesses. Dessa forma, os autores apresentam o *Location Recommendation Framework (LRT)*, um método baseado em fatoração de matrizes, no qual as preferências dos usuários em horários diferentes são modeladas por meio de diferentes matrizes de *check-ins* para cada espaço de tempo t separadamente, em que $t \in \{0, 1, \dots, 23\}$ representa a hora do dia na qual o *check-in* foi realizado. Nos resultados de avaliações experimentais realizadas sobre dados dos usuários, foi possível identificar importantes padrões de comportamento temporais dos usuários e como eles podem melhorar a efetividade das recomendações, quando comparados com estratégias que não utilizam a informação temporal. Conforme veremos mais adiante, existem outros trabalhos na literatura que também exploram a

influência temporal, porém, conjugando com outras informações (i.e. influência geográfica).

No que diz respeito à influência social, existem muitos trabalhos propostos na literatura que merecem destaque (61, 53, 62, 63). Wang, Terrovitis e Mamoulis (2013) apresentam um método de filtragem colaborativa, denominado *Location-friendship Bookmark-Coloring Algorithm - LFBCA*, que visa mitigar o problema da esparsidade em recomendações de POIs. Para isso, o método proposto cria recomendações com base em quatro fatores: (1) comportamento do usuário passado (locais visitados), (2) a localização de cada local, (3) as relações sociais entre os usuários e (4) a similaridade entre os usuários. Por outro lado, Ye, Yin e Lee (2010) assumem que existem fortes relações sociais e geográficas entre os usuários e seus POIs favoritos. Diante de tais considerações, propõem também uma abordagem de filtragem colaborativa chamada *Friend-based Collaborative Filtering (FCF)* para SsR em LBSNs. De forma resumida, o FCF explora as avaliações de POIs que foram visitados por amigos sociais de um usuário alvo. Por fim, os trabalhos de Ying et al. (2014) e Noulas et al. (2012) indicam abordagens de recomendação baseadas em grafos. No primeiro, Ying et al. (2014) propõem um modelo que considera as conexões entre usuários e POIs como arestas do grafo; de forma semelhante, Noulas et al. (2012) propõem um grafo com arestas distintas para representar as relações usuários-usuários e as relações usuários-POIs. Apesar de todos esses trabalhos apresentarem resultados interessantes do ponto de vista da efetividade, as estratégias propostas dependem de que as relações sociais sejam densas e explícitas. Além disso, conforme veremos adiante, outros trabalhos, que utilizam outras informações, têm se mostrado mais efetivos.

Do ponto de vista dos trabalhos que exploram exclusivamente a influência geográfica, destacamos algumas estratégias (64, 65, 66). Liu et al. (2014) apresentam um método de filtragem colaborativa, baseado em matriz de fatoração ponderada denominado *Instance-Region Neighborhood Matrix Factorization - IReNMF*. As intuições por trás do *IReNMF* são: (1) usuários possuem preferências semelhantes nos POIs vizinhos (influência ao nível da localização); e (2) os POIs de uma mesma região geográfica podem compartilhar preferências de usuários (influência ao nível da região). Li et al. (2015) apresentam um modelo baseado em ranking que aprende as preferências dos usuários por POIs e inclui a influência geográfica da vizinhança de POIs. Esse método, denominado *Ranking Geographical Matrix Factorization - RankGeoFM*, baseado em filtragem colaborativa e utiliza uma matriz das preferências geográficas dos usuários em conjunto com a matriz de preferência de POIs tradicional. A estratégia *Geographical Matrix Factorization - GeoMF* é proposta por Lian et al. (2014). O *GeoMF* é uma estratégia híbrida que utiliza de fatoração de matrizes para incluir a influência geográfica e as preferências dos usuários para predizer as listas de POIs mais relevantes. Para isso, os autores assumem que existem áreas de atividade dos usuários e influência dos POIs. Para extrair a influência geográfica o *GeoMF* divide uma determinada macrorregião (i.e. cidade) inteira em sub-regiões de mesmo tamanho e define que cada usuário possui uma área de atividade e cada POI possui uma área de influência sobre as regiões. Por meio de uma

vasta avaliação, os autores demonstram que a efetividade do *GeoMF* é bem superior às propostas que exploram exclusivamente outros tipos de informação (i.e. social e temporal).

Encontramos ainda na literatura um vasto número de propostas que combinam as informações geográficas e temporais (55, 67) e informações geográficas e sociais (68, 51, 69, 70, 71), com destaque, em termos de efetividade, para o último conjunto de trabalhos. Cheng et al. (2012) apresentam um método híbrido denominado *Multicenter Gaussian Model - MGM*. Com base na observação de que os *check-ins* do usuário geralmente são distribuídos em vários centros, como casa e local de trabalho, o *MGM* aprende regiões de atividade para cada usuário usando várias distribuições gaussianas, considerando também nessas distribuições as relações sociais dos usuários.

De maneira semelhante, Ye et al. (2011) apresentam o *User, social and Geographical influence - USG*, um dos trabalhos de recomendação de POIs mais referenciado na literatura e que, conforme apresentado por Liu et al. e Han et al. apresenta recomendações mais qualificadas que outros algoritmos de recomendação (72, 9). Trata-se de uma técnica híbrida que recomenda POIs levando em consideração a influência geográfica, social e as preferências do usuário. Dessa forma, essa estratégia combina esses três fatores em uma única função objetivo. A influência geográfica é extraída a partir de uma função de probabilidade de um determinado usuário visitar um POI candidato, levando em consideração o conjunto de POIs visitado pelo usuário. A influência social é extraída par a par entre os amigos de cada usuário para fazer a recomendação de POIs baseadas em amigos mais similares ao usuário alvo. Mais especificamente, a similaridade entre os amigos do usuário é baseada nos *check-ins* em POIs comuns ao usuário alvo e seus amigos. Uma estratégia muito semelhante ao USG é proposta em Zhang e Chow (2013), o *iGSLR*. Essa estratégia utiliza como informação social a similaridade entre amigos, calculada com base na distância de sua residência. No caso de ausência dessa informação, considera-se como residência a localização do POI mais frequentado pelo usuário. Já para a influência geográfica, o *iGSLR* aprende uma distribuição de distância do histórico de *check-in* usando uma função de estimação de densidade. A probabilidade de um usuário visitar um novo POI específico é assim estimada com base nos valores nas distâncias entre POI candidato e os POIs visitados pelo usuário. Li et al. (2016) apresentam o *Augmented Square error based Matrix Factorization - ASMF*, um algoritmo de recomendação de POIs que considera as informações sociais de três tipos: amigos sociais, amigos baseados em localização (i.e., geograficamente próximos) e informações de vizinhança dos amigos. Todas essas informações são incorporadas à matriz de visitas a POIs dos usuários para que métodos de fatoração possam ser aplicados. O argumento principal dos autores é que adicionar essas informações é relevante para contornar o problema de Cold Start (54, 26).

Finalmente, Zhang e Chow (2015) apresentam o *Geographical, Social and Categorical - GeoSoCa*, que considera aspectos sociais, categóricos e geográficos para prever

a relevância de um POI não visitado para um usuário. Para a influência social, modela-se a distribuição dos *check-ins* dos usuários entre os POIs para estimar a correlação entre os usuários. A partir dessa distribuição, define-se uma métrica de correlação categórica dos POIs, associando a popularidade de um POI e suas categorias para um usuário e, assim, estimar essa correlação. A influência geográfica é semelhante à apresentada por Zhang e Chow (2013), por meio de uma função de estimação de densidade.

Recentemente, Liu et al.(2017) apresentaram uma vasta avaliação experimental dos diversos algoritmos de recomendação de POIs em LBSN apresentados acima. Nessa avaliação, os autores consideraram tanto a eficiência quanto a eficácia dos algoritmos, descrevendo em quais cenários cada um dos algoritmos apresentados é mais adequado. Uma das principais conclusões desse trabalho é que, atualmente, a influência geográfica é aquela que alcança os melhores resultados de efetividade, uma vez que o usuário sempre levará em conta a distância de um local para visitá-lo. Além disso, as técnicas que combinam a influência geográfica com informações sociais também se apresentam muito promissoras. A partir desse trabalho e de toda a revisão bibliográfica apresentada acima, selecionamos para a avaliação experimental de nossa proposta os algoritmos específicos de recomendação de POIs: (i) **GeoMF**, que explora apenas a informação geográfica; (ii) **USG** e **GeoSoCa**, que exploram informações sociais e geográficas. Um fato relevante e comum a esses trabalhos é que todos são focados apenas na efetividade, sem avaliar dimensões complementares de qualidade, tais como diversidade e novidade.

3.2 Diversidade em Sistemas de Recomendação

As abordagens descritas na seção 3.1 utilizam características peculiares dos problemas de recomendação de POIs para tentar aumentar a efetividade (i.e. acurácia) das listas de recomendações. Em outras palavras, tais abordagens têm como principal objetivo aumentar a relevância dos POIs recomendados para o usuário. Entretanto, de acordo com *McNee et al. (2016)* e *Ge et al. (2010)*, elevar apenas a relevância dos itens recomendados não é suficiente para se ter recomendações de boa qualidade. Os autores de ambos os artigos argumentam que a satisfação dos usuários nem sempre está relacionada com altos índices de acurácia das recomendações, e que há outros fatores importantes para os usuários que devem ser considerados, tais como diversidade e novidade. De forma semelhante, *Wu et al. 2012* consideram que apenas acurácia não é suficiente, e que outras dimensões complementares de qualidade das recomendações também devem ser levadas em consideração. Nesse sentido, os usuários ficam mais satisfeitos com itens recomendados que atendem ou superam suas expectativas, do que com recomendações tediosas e sem novidades (57, 14). Dessa maneira, a novidade também é muito desejável na maioria dos cenários de recomendação. De acordo com *Vargas et al. (2014)*, o propósito da recomendação está inerentemente ligado a uma noção de descoberta, já que a recomendação faz mais sentido quando expõe o

usuário a uma experiência relevante que ele não teria encontrado por si mesmo. Portanto, recomendações simplesmente de grande efetividade geralmente são de pouca utilidade (16). Por tudo isso, faz-se necessário que, além de incluir nas listas itens relevantes para o usuário, os SsR incluam também itens que apresentem novidade e diversificação.

Embora relacionados, os conceitos acerca de diversidade e novidade são diferentes. A novidade normalmente refere-se a quão diferente um item é em relação aos itens já vistos anteriormente por um usuário. Por outro lado, a diversidade, foco do presente trabalho, é aplicada a um conjunto de itens e está relacionada à diferença entre os itens do conjunto (16). Por exemplo, em um cenário de recomendação de músicas, se o usuário ouve apenas canções do gênero *pop*, um sistema que recomenda músicas de outros gêneros (i.e. *rock*, *funk*, entre outros) apresenta mais itens novos e, conseqüentemente, diversos para o usuário. Logo, quando um conjunto é diversificado, cada item é novo em relação ao restante do conjunto, ou seja, existe uma relação entre os conceitos de novidade e diversidade (16).

Sobretudo, o conceito de diversificação foi introduzido em SsR a partir de trabalhos relacionados na área de recuperação de informação (RI). Uma frequente abordagem utilizada é a MMR, proposta por Carbonel e Goldstein (1998), que consiste em uma combinação linear de relevância e diversidade - Equação 3.1. O MMR é utilizado para encontrar a funcionalidade f de um documento d para uma *query* q , baseado nos D documentos disponíveis no sistema. Definindo o $\lambda = 0$, o MMR se comporta como um modelo de predição tradicional, não alterando os documentos recuperados. Por outro lado, aplicando o $\lambda = 1$, faz com que o MMR foque no processo de diversificação e altere a ordem dos documentos recuperados. Nesse caso, o maior desafio consiste em definir os valores de λ e as melhores métricas para medir a diversidade dos documentos.

$$f(q, d, D) = (1 - \lambda) \cdot rel(q, d) - \lambda \cdot \max_{d_j \in D} sim(d, d_j) \quad (3.1)$$

A maioria das abordagens minimizam a redundância dos documentos aplicando o método MMR como uma etapa de pós-processamento das ferramentas de busca (20, 21, 22). A adaptação dessa proposta para SsR consiste em reordenar uma lista de recomendação com n itens, cuja ordem original é definida por um recomendador base (RB), com base no objetivo de diversificação e os top- k itens reordenados, para $k \leq n$, são apresentados para o usuário. Os métodos de reordenação aplicados no cenário de recomendação são, principalmente, destinados a maximizar: (1) a diversificação categórica do item; (2) a distância entre os itens da lista (*intra-list distance*); ou (3) a proporcionalidade das categorias dos itens.

No primeiro aspecto, Vargas *et al.* (2014) propõem o *framework* **Binomial**, que minimiza a redundância dos tópicos retornados para uma consulta. Em sua essência, tal abordagem significa considerar a recomendação de itens aleatórios como a melhor abordagem em termos de diversidade de gêneros e usar uma distribuição binomial como

modelo para a distribuição de gêneros resultantes da amostragem aleatória de itens (18). Basicamente, o Binomial é baseado em uma métrica de diversidade definida como o produto de dois componentes: cobertura de categorias e a não redundância. A cobertura é definida pelo produto dos gêneros não cobertos e suas probabilidades. Por sua vez, o conceito de não redundância é definido pelo nível de *tolerância* restante do usuário para cada gênero já coberto. A função apresentada pela equação 3.2 combina linearmente a relevância e a diversidade binomial. Observa-se que esta função é similar a MMR, sendo que a diversidade é representada pelo quanto o item i pode aumentar a diversificação da lista R existente.

$$div(i; R) = BinomDiv(R \cup \{i\}) - BinomDiv(R) \quad (3.2)$$

De acordo com o segundo aspecto, destacamos o *framework* de **Intra-List Distance - LD**, proposto por *Vargas et al. (2014)*. Nesse, a diversidade é baseada na distância entre os itens da lista de recomendações, medida pela correlação de *Pearson*. A equação 3.3 representa o cálculo do ILD. Especialmente, em alguns casos, considerando os gêneros aos quais o item pertence como atributos do item, é possível medir a distância baseada nos gêneros dos itens. Por exemplo, em um cenário de recomendação de filmes, um filme que pertence aos gêneros comédia e romance é considerado mais distante de outro filme que pertence aos gêneros ação e policial, e mais próximo de um filme que pertence aos gêneros drama e romance. Em outras palavras, quanto mais próximo um item é considerado de outro, mais parecidos eles são entre si, ou seja, quanto menor o valor de ILD entre dois itens, mais parecidos eles são.

$$ILD = div(R) = \frac{2}{|R|(|R| - 1)} \sum_{i_k \in R, l < k} d(i_k, i_l) \quad (3.3)$$

Por sua vez, *Dang et al. (2012)* propõem um novo **Proportional Framework - PM2** que explora a diversidade baseada na proporcionalidade da relevância de cada documento d por cada tópico t . Basicamente, a fim de reduzir a redundância dos documentos, o PM2 também considera a proporcionalidade de um tópico t a outro não relacionado a d . A Equação 3.4 representa essa proporção, semelhante ao método MMR. $P(d|t)$ representa a probabilidade da relevância de um documento d para um tópico t , o termo $qt[i]$ representa o número de interações de um tópico t em uma *query* q . No domínio de recomendação, esses tópicos podem ser representados como categorias ou gêneros dos itens disponíveis.

$$f(q, d) = \lambda \cdot qt[i^*] \cdot P(d|t_{i^*}) + (1 - \lambda) \cdot \sum_{i \neq i^*} qt[i] \cdot P(d|t_i) \quad (3.4)$$

Todas essas estratégias podem ser adaptadas para o problema da diversidade em recomendação de POIs, considerando como categorias: museus, restaurantes, teatro, igrejas etc. Contudo, embora alcancem resultados satisfatórios em cenários tradicionais, esses

métodos não provêm, necessariamente, uma diversificação geográfica, uma dimensão de qualidade complementar fundamental no cenário de POIs. Alguns trabalhos até tentaram adaptar um grupo específico de métodos de reordenação que se baseiam na comparação par a par da distância do item (20, 73, 74). Basicamente, essas propostas trocam o conceito de “distância” para a distância geográfica entre dois POIs. Entretanto, esses métodos simplesmente maximizam a distância entre POIs, recomendando POIs localizados longe uns dos outros. Tal abordagem não possui aplicabilidade prática, uma vez que os usuários preferem POIs próximos à sua localização atual (51).

3.3 Diversificação Geográfica em Recomendação de POIs

Sob a perspectiva de recomendação de POIs, a diversificação das recomendações é tão importante quanto a relevância dos itens recomendados (9), pois, recomendar sempre os mesmos locais em uma região geográfica muito restrita, tende a perder o interesse do usuário. Ainda assim, conforme observamos em seções anteriores, as características geográficas são na maioria das vezes utilizadas para melhorar apenas a relevância dos itens recomendados, e não a diversificação geográfica deles (75, 56, 51, 3). Além disso, outros SsR têm o propósito de diversificar os resultados pelas categorias dos POIs, e não pela localização destes (71, 53). Tendo em vista esses aspectos, é evidente que o problema da diversificação geográfica ainda é pouco estudado e permanece como um desafio no cenário de recomendação de POIs (9, 49).

Pelo melhor de nosso conhecimento, apenas o trabalho apresentado por *Han J. and Yamana H.* (9) visa melhorar a diversificação geográfica, sendo portanto o trabalho mais relevante no contexto desta dissertação de mestrado. Os autores assumem que a rotina de visita dos usuários em seus distritos de maior atividade (e.g., casa, trabalho etc) faz com que os principais métodos recomendem apenas POIs localizados nessa pequena porção de regiões. Em particular, os autores argumentam que recomendar POIs próximos à localização atual de um usuário - uma forma popular de aplicação para dispositivos móveis - é menos útil. Ao invés disso, defendem que recomendar POIs, assumindo que os interesses dos usuários são distribuídos em várias sub-regiões de um distrito, alcançam melhores resultados. A estratégia, denominada **Geo-Div**, apresentada divide a área onde os POIs estão localizados em regiões menores de tamanho fixo chamadas subáreas, onde cada uma mede $0,5 \times 0,5 \text{ km}^2$. Além disso, é usado o conceito de densidade de atividade do usuário por subárea proposto por Zhang e Chow (2015). Em outras palavras, de acordo com a frequência de *check-ins* do usuário em uma subárea é calculado um valor, entre 0 e 1, que representa o quanto um usuário é ativo em uma subárea. Em seguida, são selecionadas apenas as subáreas onde há grande concentração de atividade do usuário. Em seguida, um recomendador base seleciona os 80 POIs mais relevantes localizados geograficamente no distrito ativo do usuário. A partir dessas 80 POIs, os autores propõem uma reordenação dessa lista, considerando a diversificação geográfica proporcional dos

POIs recomendados. Esse método é baseado no MMR e define a diversidade do item de acordo com a proporcionalidade da recomendação, demonstrado nas Equações 3.5 e 3.6.

$$f(i, R) = rel(i)^{1-\lambda} + div_g(i, R)^\lambda \quad (3.5)$$

$$div_g(i, R) = max(0, PR_g(\{i\} \cup R) - PR_g(R)) \quad (3.6)$$

De forma resumida, cada POI relevante p_i , entre os 80 candidatos selecionados pelo recomendador base, recebe um valor de proporcionalidade $cin[0, 1]$. Esse valor é atribuído com base na quantidade de POIs que o usuário visitou, localizados geograficamente em até 500 metros de distância de p_i . Consequentemente, quanto mais POIs foram visitados nas proximidades de p_i , mais o valor de c se aproxima de 1; de forma contrária, quanto menos POIs visitados, mais o valor c se aproxima de 0. Em seguida, com base no valor de relevância calculado pelo recomendador base e no valor de c de cada POI a proporcionalidade da diversidade é calculada. Sendo assim, os POIs que apresentam maiores valores de proporcionalidade são colocados no topo da lista final de recomendações. Esse método melhora a distribuição proporcional dos primeiros itens recomendados, porém, é notório que nesse a diversificação geográfica ainda é limitada. De acordo com a proposta apresentada, ao selecionar apenas as subáreas com altos valores de atividade do usuário e recomendar POIs que pertencem apenas a estas subáreas o método proposto limita a diversificação geográfica. Apesar disso, esse trabalho ainda é a principal fonte de inspiração para esta dissertação.

Conforme veremos no capítulo seguinte, neste trabalho propomos o *Geo-Cat*, um método de pós-processamento a ser aplicado sobre uma lista de um recomendador base, no intuito de maximizar a diversidade categórica e geográfica, mantendo a relevância dos POIs. Modelamos esse problema de forma semelhante à MMR. Enquanto a relevância é medida com base no *score* atribuído por um recomendador base, a diversidade é uma combinação linear de dois fatores: a diversidade geográfica (9) e a diversidade categórica (16).

4 Geo-Cat: Diversificação Geográfica e Categórica

Neste capítulo, apresentamos uma visão geral de nossa estratégia para se obter SsR de POIs mais diversos geográfica e categoricamente. Apresentamos em detalhes a estratégia de diversidade categórica adaptada, bem como a estratégia de diversidade geográfica e como esses dois conceitos são combinados formando, assim, nossa proposta: o *Geo-Cat*. Basicamente, este capítulo trata da primeira contribuição deste trabalho, quando apresentamos como cada um dos três primeiros objetivos específicos traçados na seção 1.2.2 foram alcançados e, conseqüentemente, abordando parcialmente a primeira hipótese desta pesquisa: (H1): *É possível explorar a diversidade categórica e geográfica, simultaneamente, sem perdas relevantes de efetividade.*

4.1 Visão Geral

Conforme temos discutido e, principalmente, destacado no capítulo anterior, os principais SsR de POIs focam na construção de modelos que privilegiam métricas de efetividade, tais como *Precision* e *Recall* (71, 51, 24). Por outro lado, as estratégias de diversidade visam adicionar outras dimensões de qualidade ao processo de recomendação, agregando assim mais satisfação para os usuários quanto aos itens recomendados (15). Contudo, apesar de estratégias como *Binom*, *ILD* e *PM2* serem muito eficientes nos cenários clássicos de recomendação (i.e. filmes, músicas, etc), elas não foram projetadas para o domínio de recomendação de POIs. De fato, tais estratégias não consideram o impacto do fator geográfico que define os pontos de interesse e, portanto, não são tão adequadas para POIs. Por outro lado, a estratégia *Geo-Div*, estado da arte em diversificação geográfica, não considera os aspectos intrínsecos dos POIs, como suas categorias (i.e. bares, museus, etc) ou mesmo perfil dos usuários que os visitam. Por essa razão, os resultados apresentados por *Han e Yamana. (2017)* (9) indicam que os POIs recomendados são melhor distribuídos geograficamente, porém, não alcançam bons valores de diversidade categórica. Dessa forma, a estratégia de recomendação apresentada neste trabalho, a qual denominamos de **Geo-Cat**, é projetada para considerar ambos os conceitos de diversificação: categórica e geográfica.

Assim como os principais métodos voltados para diversificação em SsR, o **Geo-Cat** também consiste de uma estratégia de pós-processamento que amplia os objetivos iniciais de efetividade de algum RB. Conforme ilustrado na Figura 3, o processo de diversificação proposto primeiramente utiliza um RB, tradicional ou específico, para compor uma lista temporária com n POIs. Nos RsB tradicionais, as matrizes de *check-ins* do usuário, POIs

e/ou itens, são usadas como as únicas fontes de informação para gerar as recomendações. Por outro lado, SsR específicos para POIs podem usar outras fontes de influência, tais como geográficas, sociais ou temporárias, em conjunto com essas matrizes para tentar obter uma maior efetividade nos POIs recomendados. Em seguida, a lista temporária é processada pelo **Geo-Cat** que, em suma, faz a reordenação dos itens da lista temporária baseado no resultado da função objetivo que considera tanto a diversificação categórica (16) quanto a geográfica (9). Nesse sentido, o Geo-Cat utiliza a nota de relevância (rel) que representa o *rating* atribuído pelo RB para cada POI em conjunto com os valores que cada POI agrega às diversidades categórica (cat) e geográfica (geo). Dessa maneira, o Geo-Cat reordena a lista final de POIs em ordem decrescente pelo valor obtido na função objetivo do Geo-Cat. Por fim, são recomendados ao usuário os top- k primeiros POIs da lista final, onde $k \leq n$. Nas seções seguintes descrevemos detalhadamente as estratégias de diversificação categórica e geográfica, isoladamente e, posteriormente, a abordagem utilizada para combinar ambos os conceitos.

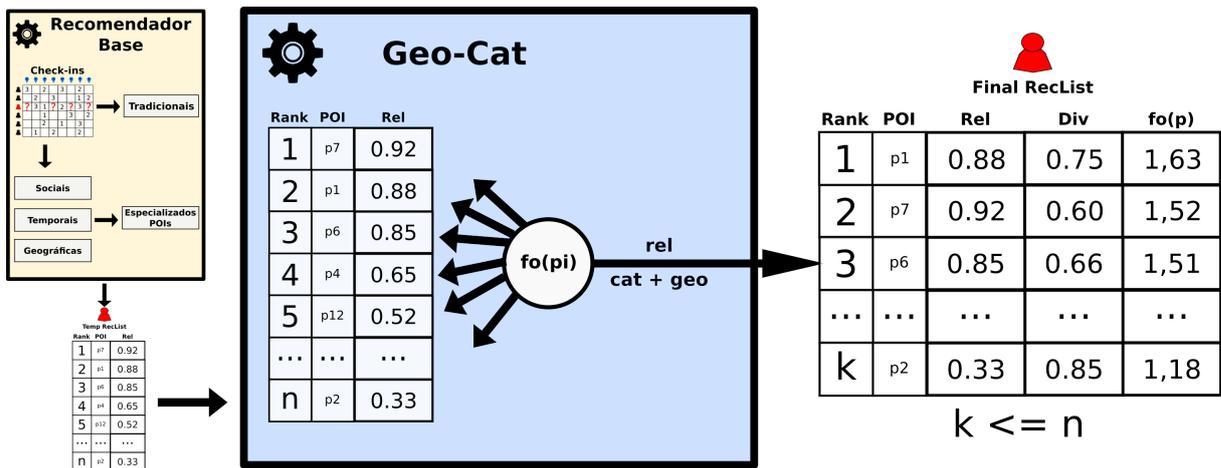


Figura 3 – Visão Geral do Funcionamento do Geo-Cat.

4.2 Diversificação Categórica

Esse modelo visa alcançar a diversidade categórica por meio dos fatores que estão intrinsecamente relacionados aos POIs. Especificamente, estamos nos referindo às suas categorias (e.g., restaurante, cinema, museu, entre outros) e aos conteúdos dos POIs (i.e. descrições textuais, atributos etc). Nossa proposta, neste caso, é combinar esses dois fatores em apenas uma métrica que visa otimizar a diversidade categórica dos recomendadores. Para o primeiro fator, adaptamos a métrica de *Genre Coverage*, proposta por *Puthiya et al.* (28), enquanto para o segundo fator adaptamos a métrica de *Intra-List Distance* (ILD), proposta por *Vargas et al.* (2014). Nossa proposta consiste em agregar essas duas características por meio de uma combinação linear, dando pesos iguais a ambas, como mostra a Equação 4.1. Dessa forma, representa-se quanto um item i influencia a lista de

recomendação R sob os pontos de vista do *Genre Coverage* e do *ILD*.

$$div_{cat}(i, R) = GenreCov(\{i\} \cup R) + \frac{ILD(\{i\} \cup R)}{|R|} \quad (4.1)$$

A métrica de *Genre Coverage* (28) visa mensurar a diversidade das categorias de POIs recomendados. Basicamente, essa métrica representa a média de categorias relevantes presentes na lista de recomendação para o usuário. Valores altos de *Genre Coverage* indicam que os POIs recomendados cobrem um número maior de categorias relevantes. Neste trabalho, são consideradas como relevantes aquelas categorias às quais pertencem os POIs visitados um número maior de vezes que a média de visitas do usuário alvo. Por exemplo, se os dados do usuário mostram que ele visita em média cinco vezes cada POI, então são consideradas relevantes as categorias dos POIs que foram visitados cinco ou mais vezes por este usuário. Dessa maneira, para cada novo POI inserido na lista, o valor total de *Genre Coverage* é acrescido do valor calculado para o novo item incluído na lista de recomendação. Em outras palavras, o valor de *Genre Coverage* aumenta à medida que novos POIs, de categorias diferentes das quais os POIs que já foram inseridos na lista pertencem são incluídos na lista de recomendações.

A métrica de *Intra-List Distance* (ILD) (16) visa identificar um conjunto de POIs capaz de cobrir o máximo possível de perfis de POIs. As categorias dos POIs são consideradas como atributos e os POIs são modelados como vetores desses atributos. A *ILD* entre os POIs é calculada com base na dissimilaridade de categorias par a par dos POIs e, consequentemente, dada pela quantidade de categorias diferentes às quais os POIs pertencem. Portanto, são considerados menos similares POIs que pertencem a um número menor de categorias em comum. Por essa razão, aumentar os valores de *ILD* indica que os POIs recomendados são mais distintos entre si. O *ILD* é calculado sobre toda a lista de recomendação, e não individualmente para cada POI. Por essa razão, precisamos dividi-lo pelo tamanho da lista para utilizar o valor médio de cada item presente, conforme apresentado na equação 4.1.

Na prática, ambas as métricas são essenciais, pois é possível que os POIs recomendados pertençam a várias categorias relevantes, tendo altos valores de *Genre Coverage*, mas pertençam a categorias similares. Contudo, com a utilização de *ILD*, o modelo proposto consegue capturar essas diferenças entre as categorias e gerar listas de recomendação com mais diversidade categorica.

4.3 Diversificação Geográfica

De maneira complementar, nosso modelo visa também maximizar a diversidade geográfica dos POIs recomendados, considerando para isso o nível de atividade dos usuários em determinadas sub-regiões. Especificamente, dividimos uma área (e.g., cidade) em sub-

regiões de $0.5 \times 0.5 \text{ km}^2$ e seguimos a direção do trabalho de *Han J. e Yamana H. (2017)*, que define a métrica de diversificação geográfica baseada na proporcionalidade (PR_g). Na Figura 4 apresentamos um exemplo de como ficaria a cidade de *Las Vegas*, caso ela fosse dividida em sub-regiões quadradas de $0.5 \times 0.5 \text{ km}^2$ (linhas azuis) e como ficaria a distribuição de POIs nessas sub-regiões (pontos vermelhos). Observando a figura, há uma grande discrepância no número de POIs por sub-regiões, ou seja, há tanto algumas sub-regiões que acumulam muitos POIs, quanto outras que não têm nenhum. Além disso, percebemos que a maioria dos POIs estão localizados geograficamente nas sub-regiões mais centrais da Figura 4.

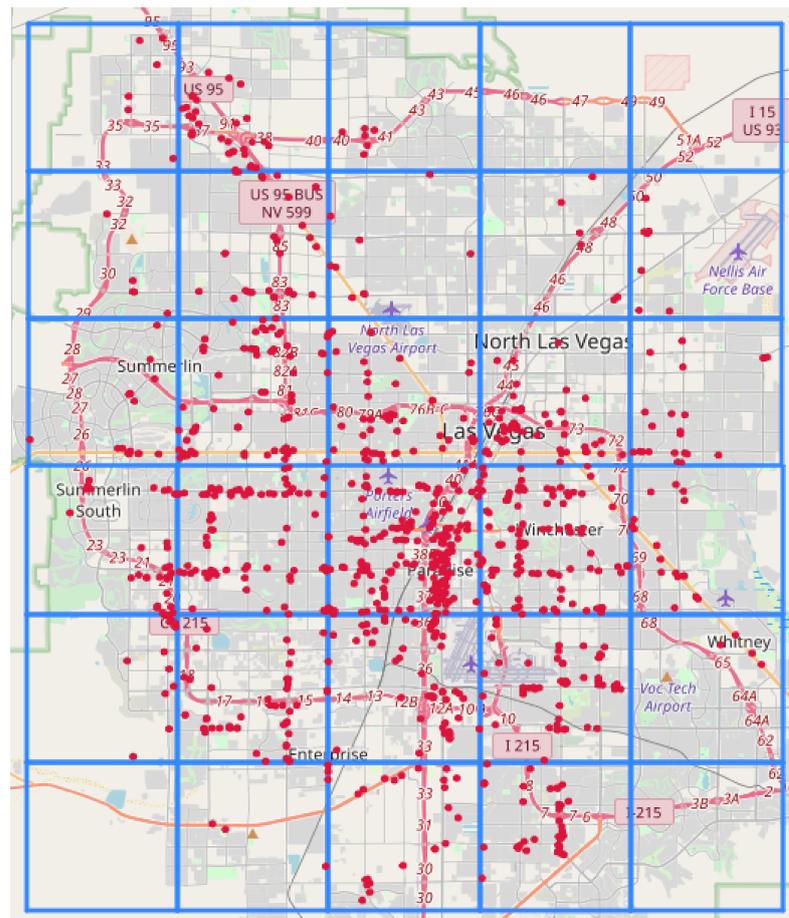


Figura 4 – Exemplo hipotético de uma distribuição de POIs sobre a divisão da cidade de Las Vegas em sub-regiões quadradas

Fundamentalmente, a métrica de PR_g é o complemento da métrica de desproporcionalidade DP_g normalizada, como mostra a Equação 4.2 proposta por *Dang V. e Croft W. (2012)*. Essa desproporcionalidade DP_g mensura se o comportamento do usuário está ou não dividido proporcionalmente nas sub-regiões de uma cidade. Para isso, DP_g leva em conta cada POI l presente no histórico de visitas do usuário e a cobertura desse POI na região, conforme mostra a Equação 4.3. A cobertura c de um POI é calculada por um método iterativo no qual c é inicializado com o resultado da divisão entre o número de POIs visitados pelo usuário e a quantidade de lugares recomendados. Em seguida, esse valor é alterado de acordo com os vizinhos de cada POI: (1) se o POI recomendado possui vizinhos, c é dividido

pela quantidade de vizinhos do POI e distribuído para cada um deles, que armazenam a soma desses valores ao longo das iterações; ou (2) se o POI não possui vizinhos, o valor de c é acumulado (n_{NR}) para ser utilizado no cálculo da desproporcionalidade. Assim como nos trabalhos da literatura (9), definimos como vizinhos de um POI l todos os POIs que foram consumidos pelo mesmo usuário e estão em um raio de 100 metros de distância de l .

$$PR_g(P) = 1 - DP_g(P)/(|T_u| + 1/2) \quad (4.2)$$

$$DP_g(P) = \sum_{l \in T_u} f_l \cdot (v_l - s_l)^2 + n_{NR}^2/2 \quad (4.3)$$

Além disso, esses valores de cobertura são utilizados para definir a variável binária f_l de cada POI. Nesse contexto, f_l é igual a 1 quando o valor de cobertura acumulado pelo POI é menor que 1. Caso contrário, o valor assumido é 0. Dessa forma, o modelo garante que a desproporcionalidade seja calculada apenas para POIs que receberam valores de cobertura bem distribuídos. POIs que receberam um valor de cobertura muito alto significa que estão muito próximos geograficamente dos POIs recomendados e devem ser penalizados com baixa proporcionalidade. Por sua vez, v_l representa o valor esperado de cobertura acumulada e s_l representa os valores de cobertura mensurados. Por essa razão, o valor de v_l observado sempre será 1, pois quando o valor de cobertura c não é distribuído de forma uniforme entre os POIs do conjunto de treinamento, a desproporcionalidade é máxima (9). Todas essas variáveis compõem a Equação 4.3, que mensura a desproporcionalidade de cada POI l do conjunto de treino do usuário.

Dessa forma, o conceito de maximizar a diversidade geográfica significa recomendar de forma proporcional POIs localizados em distintas sub-regiões que já são frequentadas pelo usuário, aumentando assim sua proporcionalidade em todas as regiões. Para exemplificar, consideramos três sub-regiões $\{a, b, c\}$ frequentadas pelo usuário u e uma lista de recomendações com n POIs. Consideramos também que o histórico de *check-ins* do usuário mostra que no passado ele visitou, 10, 5 e 5 POIs nas sub-regiões a , b e c , respectivamente. Portanto, para maximizar a diversidade geográfica, os POIs recomendados devem ser distribuídos de forma proporcional às visitas de u nas três sub-regiões, ou seja, metade dos POIs devem estar localizados na sub-região a , 1/4 na sub-região b e 1/4 na sub-região c . Caso os POIs sejam distribuídos aleatoriamente em várias sub-regiões ou os POIs recomendados estejam muito concentrados em poucas sub-regiões, o valor de PR_g será penalizado, conforme descrito anteriormente nesta seção. Por essas razões, propomos na Equação 4.4 maximizar o valor de PR_g . Para cada POI i a ser testado, calcula-se a diferença entre o novo PR_g , adicionando i , e o PR_g atual. Assim, espera-se recomendar um número de

POIs por sub-região que seja proporcional ao comportamento do usuário no passado.

$$div_{geo}(i, R) = PR(\{i\} \cup R) - PR(R) \quad (4.4)$$

4.4 Diversificação Geográfica-Catagórica

Com as Equações 4.1 e 4.4 definidas, o *Geo-Cat* consiste então em uma combinação linear desses conceitos supracitados, como mostra a Equação 4.5. A primeira parte dessa equação refere-se à diversidade geográfica, mais especificamente à proporcionalidade geográfica de distribuição dos POIs proposta por Han J. e Yamana H. (2017). A segunda parte corresponde à diversidade categórica dos POIs, em que utilizamos as duas principais métricas existentes na literatura: *ILD* e *Genre Coverage*. Para equilibrar os fatores geográficos e categóricos, adicionamos um parâmetro β a ser ajustado com algum valor real entre 0 e 1. Assim, a configuração desse parâmetro permite modificar o comportamento da diversidade do nosso método *Geo-Cat*. Dessa maneira, valores próximos de 0 resultam em recomendações mais diversas categoricamente, enquanto valores próximos de 1 resultam em recomendações com maior diversidade geográfica.

$$div_{geo-cat}(i, R) = \beta \cdot div_{geo}(i, R) + (1 - \beta) \cdot div_{cat}(i, R) \quad (4.5)$$

Visando equilibrar os conceitos de diversidade e efetividade (i.e. acurácia), propomos um método de reordenação para maximizar a diversidade geográfica-catagórica de uma lista de POIs. Nossa função objetivo, apresentada na equação 4.6, corresponde matematicamente ao método de diversificação do *Geo-Cat* acrescido da relevância do POI para o usuário. Essa relevância, representada na variável $rel(i)$, representa o *score* atribuído por qualquer recomendador base ao POI_i . Nosso intuito é manter a qualidade da predição realizada pelo recomendador base, ao mesmo tempo em que maximizamos o fator de diversificação proposto. Esse *trade-off* é equilibrado pelo parâmetro λ , um valor real entre 0 e 1. Quando λ é próximo de 1, são produzidas listas de recomendação que atribuem mais peso para a diversidade. Por outro lado, se λ é próximo de 0, são geradas listas de recomendações cujos pesos são maiores para a nota de relevância do POI, criando recomendações menos diversas e mais parecidas com as que foram geradas originalmente pelo recomendador base. Essa proposta é semelhante a utilizada em Han e Yamana (2017) para melhorar a diversidade geográfica das recomendações de POIs. Contudo, no *Geo-Cat* reconstruímos a segunda parte da equação com o objetivo de considerar simultaneamente a diversidade categórica e geográfica.

$$f_o(i, R) = rel(i)^{1-\lambda} + div_{gc}(i, R)^\lambda \quad (4.6)$$

$$div_{gc}(i, R) = \max(0, div_{geo-cat}(\{i\} \cup R)) \quad (4.7)$$

Para maximizar o fator de diversidade, definimos a Equação 4.7 e propomos uma otimização gulosa para selecionar POIs que potencializam o método *Geo-Cat*. Esse processo é representado no Algoritmo 1. Inicialmente, obtém-se uma lista temporária oriunda de qualquer recomendador base (linha 1). Na primeira iteração, o *Geo-Cat* calcula a função objetivo para cada um dos n POIs da lista temporária para encontrar o POI com o maior valor dentre todos (linha 4). Em seguida, esse POI é incluído na lista de recomendações (como primeiro POI) e removido da lista temporária (linhas 5 e 6). Na segunda iteração, o algoritmo calcula a função objetivo (linha 4), acrescentando cada um dos POIs remanescentes na lista com o primeiro POI. Com isso, o algoritmo detecta qual é o POI que, junto com os que já foram adicionados anteriormente na lista de recomendações, alcança o maior resultado na função objetivo. Esse processo é repetido por k vezes, a fim de montar uma lista com os top- k POIs que maximizam essa função objetivo.

Algorithm 1 GEO-CAT DIVERSIFICATION (u, R', k)

```

1:  $R' \leftarrow getRec(RB)$ 
2:  $R \leftarrow \emptyset$ 
3: for  $i < k$  do
4:    $POI^* \leftarrow \max_{p \in R'} (rel(p)^{1-\lambda} + div_{gc}(p, R)^\lambda)$ 
5:    $R \leftarrow POI^*$ 
6:    $R' \leftarrow R' \setminus POI^*$ 
7: end for
8: return  $R$ 

```

Dessa forma, ao longo deste capítulo, apresentamos como alcançamos os objetivos específicos 1, 2 e 3, e que se constituíram como a primeira contribuição deste trabalho. A partir dessas descrições, abordamos parcialmente nossa primeira hipótese (*H1*): *É possível explorar a diversidade categórica e geográfica, simultaneamente, sem perdas relevantes de efetividade*. Para a consolidação dessa hipótese, resta ainda uma avaliação quanto a assertividade de nossa estratégia, a qual será apresentada no próximo capítulo. Além disso, para que essa avaliação seja realizada adequadamente, precisamos definir um novo processo, uma vez que introduzimos uma nova abordagem que combina conceitos distintos (i.e. efetividade, diversidade categórica e diversidade geográfica) que devem ser considerados em conjunto. Assim, no próximo capítulo também detalhamos nossa proposta de avaliação que consiste, basicamente, na adaptação da MAUT para os cenários de recomendação.

5 Avaliação Experimental

Neste capítulo, primeiramente apresentamos nossa configuração experimental, descrevendo as bases de dados utilizadas, bem como os algoritmos e métricas selecionadas para todo o processo avaliativo. Além disso, realizamos um estudo preliminar para configurar os parâmetros do nosso método de recomendação *Geo-Cat*. Em seguida, avaliamos nossa proposta em duas partes, por meio de: (1) um comparativo entre os resultados alcançados por diferentes recomendadores base e pelo *Geo-Cat* após realizar o pós-processamento; e (2) outro comparativo entre resultados alcançados pelo *Geo-Cat* com diferentes algoritmos de diversificação estado da arte. Por fim, agregamos os resultados por meio da métrica MAUT, no intuito de identificar a estratégia capaz de equilibrar os principais requisitos de qualidade.

5.1 Configuração Experimental

No intuito de avaliar o modelo proposto neste trabalho, definimos um projeto experimental a ser aplicado em dados de cidades reais. Para isso, selecionamos os principais métodos de recomendação da literatura como recomendadores base, bem como as técnicas estado da arte para diversificação para serem utilizadas como linhas de base em comparação com nossa estratégia. Para avaliar o desempenho de nosso método de diversificação, selecionamos diversas métricas da literatura e, além disso, adaptamos a MAUT para avaliar um método de recomendação sobre múltiplas dimensões de qualidade e satisfação simultaneamente, tais como acurácia, novidade e diversidade.

5.1.1 Base de Dados

Nossa base de dados é constituída pelos dados disponíveis na base do *Yelp Challenge*¹]Disponível em <https://www.yelp.com/dataset/challenge> relativos às cidades de *Charlotte*, *Las Vegas* e *Phoenix* nos Estados Unidos. A principal razão pela qual essas cidades foram selecionadas é que elas estão entre as que têm maior número de POIs, usuários e *check-ins* dentro todas cidades disponíveis nessa base de dados. Nessas três cidades, assim como é feito por outros trabalhos da literatura (9, 71, 51), foi realizado o pré-processamento dos dados para selecionar em cada cidade apenas: (1) os POIs que foram visitados por cinco ou mais usuários; e (2) os usuários que realizaram vinte ou mais *check-ins* na cidade alvo. A quantidade de *check-ins*, usuários e POIs remanescentes, após o pré-processamento, são apresentadas no gráfico da Figura 5.

¹ [

Observando o gráfico da Figura 5, temos que as cidades escolhidas para avaliação possuem características distintas, possibilitando avaliar o desempenho do *Geo-Cat* para diferentes proporções entre as quantidades de usuários e POIs. Após o pré-processamento, são selecionados na cidade de *Charlotte* 148.495 *check-ins*, 7.151 POIs e 2.445 usuários. Por sua vez, Las Vegas é a segunda maior cidade em número de POIs e a primeira em quantidade de usuários, comparando com todas as outras cidades da base de dados. Sobretudo, em Las Vegas foram selecionados 856.686 *check-ins* para 15.168 usuários e 11.562 POIs. *Phoenix*, por sua vez, tem a maior quantidade de POIs, mas fica em segundo lugar em número de usuários. Em *Phoenix*, selecionamos 808.441 *check-ins*, 14.388 usuários e 36.827 POIs. Nota-se que, com exceção de Las Vegas, todas as outras cidades selecionadas possuem mais POIs que usuários na base de dados. Cabe ressaltar ainda que neste trabalho, seguindo a ordem cronológica, foram separados 70% dos primeiros *check-ins* de cada usuário para realizar o treinamento dos modelos e os 30% restantes foram utilizados para testes, porque entende-se que na prática devem ser analisados os *check-ins* do passado para tentar prever comportamentos futuros (71).

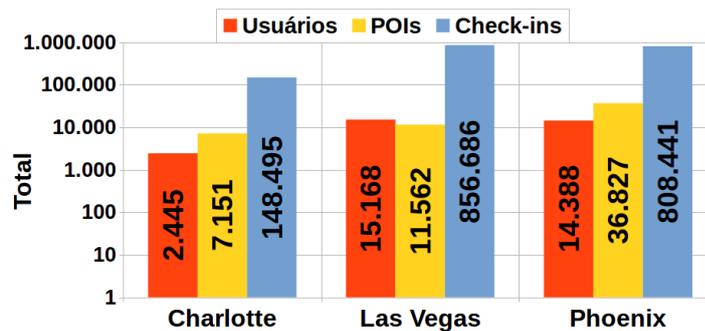


Figura 5 – Gráfico com quantidades de Usuários, POIs e *check-ins*, após pré-processamento, das cidades de *Charlotte*, *Las Vegas* e *Phoenix*

De maneira complementar, analisamos as relações entre as quantidades de POIs, usuários e *check-ins* das três cidades selecionadas e o resultado dessa análise é apresentado no gráfico da Figura 6. Por conseguinte, observa-se neste que há em média 20,77 *check-ins* por POI em *Charlotte* e 21,95 em *Phoenix*, ao passo que em *Las Vegas* a proporção média de *check-ins* por POI é de 74,09. Nota-se também que o valor médio de POIs por usuários em *Charlotte* e *Phoenix* são próximos (2,92 e 2,56 respectivamente). Contudo, Las Vegas destaca-se das outras duas cidades porque possui o valor médio de POIs por usuários mais baixo de todos (apenas 0,76). Por fim, foram analisadas as médias de *check-ins* por POIs e verificou-se que em Las Vegas o valor é de 74,09, enquanto que para as cidades de *Charlotte* e *Phoenix* a média é de 20,77 e 21,95, respectivamente. Portanto, em média os POIs localizados em Las Vegas recebem mais *check-ins* do que os POIs das outras cidades. Todavia, conforme apresentado na seção 5.3 deste trabalho, as diferenças numéricas apuradas no gráfico da Figura 6 não afetam a estabilidade dos resultados alcançados pelo *Geo-Cat* em nenhuma das três cidades selecionadas.

Além disso, para compreender melhor a distribuição de POIs por sub-regiões nas

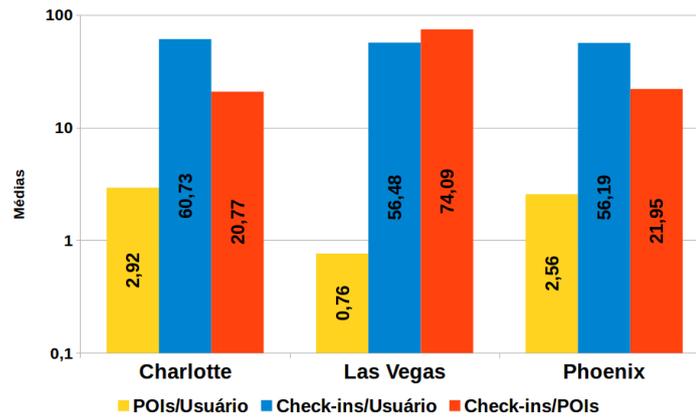


Figura 6 – Gráfico que apresenta os valores médios de POIs por Usuários, *check-ins* por usuários e *check-ins* por POIs

idades analisadas, neste trabalho foram plotados os gráficos apresentados nas Figuras 7a, 7b e 7c para as cidade de *Charlotte*, *Las Vegas* e *Phoenix*, respectivamente. Nesses gráficos são apresentadas as quantidades totais de POIs no eixo Y pelas subáreas de cada cidade no eixo X. Ressaltamos que, as subáreas estão ordenadas das que têm maior quantidade de POIs para as que têm menor quantidade. Por exemplo, a interpretação do gráfico apresentado na Figura 7b mostra o fato que em *Las Vegas*, das 1.100 sub-regiões menos de 20% delas possuem 25 POIs ou mais. Ademais, a grande maioria das sub-regiões de *Las Vegas* possuem poucos POIs. Dessa maneira, observa-se que as curvas de distribuições apresentadas nesses gráficos comprovam que em todas as cidades há poucas sub-regiões que acumulam muitos POIs e que a grande maioria delas possui um número pequeno de POIs. Isso reforça a discussão sobre a métrica de PR_g apresentada na Seção 2.4, de que diversificar a recomendação de POIs simplesmente pela proporção de sua distribuição por sub-regiões distintas aleatoriamente não necessariamente resulta em bons resultados. Conseqüentemente, esses dados também reforçam a necessidade de realizar uma diversificação geográfica bem trabalhada, considerando os níveis de atividade dos usuários em cada sub-região.

Outro aspecto abordado na análise dos dados do *Yelp* é a questão das categorias às quais os POIs pertencem. Por exemplo, um POI pode pertencer às categorias *restaurants*, *food* e *nightlife* de forma concomitante, ou apenas a uma delas. Sobretudo, as categorias são organizadas hierarquicamente em uma árvore multiníveis semelhante à que é apresentada na Figura 8. Como é possível observar na árvore, há uma categoria raiz “mãe” de todas as outras categorias. Abaixo dela há várias categorias como *Artes* e *Comida*, consideradas de nível 1 e, mais abaixo, há outras categorias mais específicas como *Cinema* e *Teatro*, consideradas de nível 2, portanto, “filhas” de alguma categoria do nível 1.

Assim como foi implementado por *Han et al.* (9), em nossos experimentos descartam-se as categorias que estão abaixo do segundo nível da árvore. Essa filtragem é essencial para a execução dos testes porque, como descrito anteriormente, as categorias são utilizadas

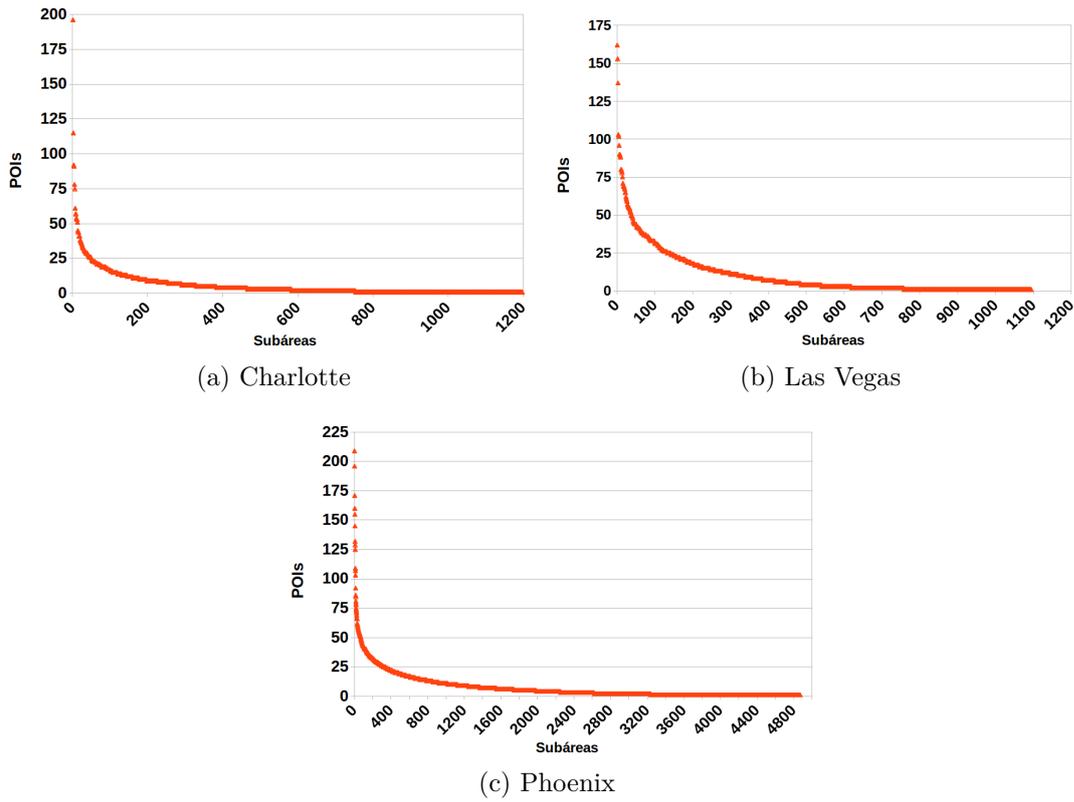


Figura 7 – Gráfico de distribuição da quantidade de POIs por sub-regiões de cada cidade.

para avaliar a similaridade entre POIs. Dessa forma, é possível medir a distância entre as categorias de um POI i em relação às categorias de outro POI j usando a Equação 5.1 proposta por *Castillo et al. (2008)*, em que cat_i representa a categoria do POI i , cat_j indica a categoria do POI j e $sp(cat_i, cat_j)$ é a distância do caminho mais curto entre as categorias i e j na árvore. Ressalta-se ainda que, quando o i ou o j têm mais de uma categoria, é considerado o menor valor da distância entre eles.

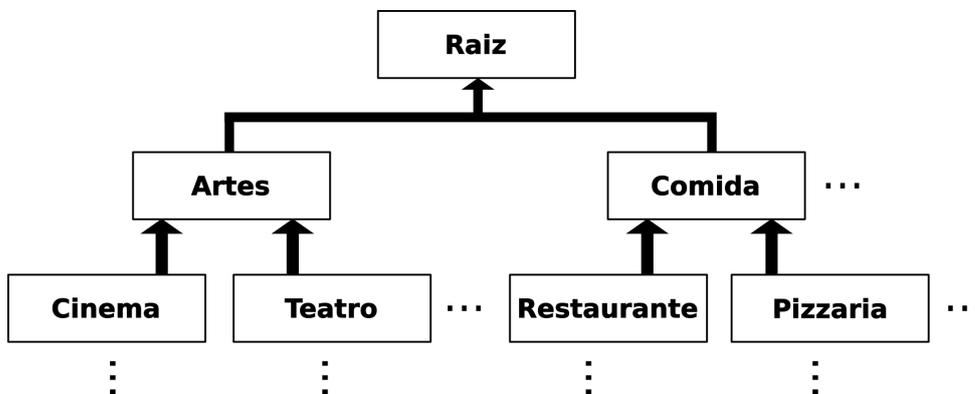


Figura 8 – Exemplo de uma árvore de categorias que ilustra como as categorias de POIs são organizadas na base do *Yelp*

$$dissim_{cat}(i, j) = 1 - \frac{1}{(1 + sp(cat_i, cat_j))} \quad (5.1)$$

5.1.2 Recomendadores Seleccionados

O impacto da aplicação do *Geo-Cat* é avaliado sob duas perspectivas. Em primeiro lugar, avaliamos a aplicação do método *Geo-Cat* sobre sete algoritmos de recomendação, dos quais quatro são métodos de recomendações clássicos (*most popular*, *user-knn*, *item attribute-knn* e *WRMF*) e três são métodos de recomendação específicos para o cenário de POIs (*USG*, *GeoSoCa* e *GeoMF*). Os métodos de recomendação tradicionais foram executados por meio de implementações disponíveis na biblioteca *MyMediaLite* (76). Por sua vez, os métodos específicos para POIs foram executados com base no código do USG disponibilizado por *Han et al.* (9) e nos códigos de (*GeoSoCa* e *GeoMF*) disponibilizados por *Liu et al.* (72). A seguir, apresentamos uma breve descrição sobre cada um dos métodos seleccionados, bem como as configurações de parâmetros utilizadas em cada um deles quando necessárias.

- **Most Popular** (76): recomenda os top- k POIs mais populares entre aqueles que ainda não foram visitados pelo usuário alvo. Em suma, esse é um método de recomendação não personalizado. Por essa razão, quando o método é aplicado sobre os POIs de uma cidade, desde que os dados de *check-in* dos usuários não sofram alterações, os mesmos POIs são recomendados para todos os usuários.
- **User-kNN** (32): recomenda os POIs mais relevantes para usuários mais similares ao usuário alvo, sob um raio de vizinhança N . Nesse caso, configuramos o valor do parâmetro $N = 80$; essa configuração especifica que são consideradas as preferências dos 80 usuários mais similares ao usuário alvo.
- **WRMF** (76): recomenda os POIs mais bem avaliados com base nos fatores latentes dos usuários e dos POIs, resultantes da aplicação de um método de fatoração de matrizes.
- **Item Attribute kNN** (76): recomenda os top- k POIs mais similares aos top- N POIs que foram melhor avaliados pelo usuário alvo. Especificamente, esse método considera os *atributos* (e.g., categorias) dos POIs como base para calcular a similaridade entre os POIs. Nesse caso, configuramos o valor do parâmetro $N = 100$. Essa configuração especifica que são considerados os 100 POIs melhores avaliados pelo usuário alvo.
- **USG** (51): recomenda os POIs tal qual o *User-kNN*; porém, calcula a similaridade dos usuários considerando, além das preferências dos top- N usuários mais similares, informações geográficas dos POIs e os relacionamentos sociais do usuário alvo. Nesse caso, configuramos o valor do parâmetro $N = 80$; essa configuração funciona exatamente como explicado anteriormente para o algoritmo **User-KNN**.
- **GeoMF** (66): recomenda os POIs tal qual o *WRMF*; entretanto, esse método é considerado uma estratégia híbrida que por meio da fatoração de matrizes considera, além das preferências do usuário, a influência geográfica dos POIs, modelando para isso as regiões de atividade do usuário alvo (72).

- **GeoSoCa** (71): recomenda os POIs de forma similar ao *User-KNN*; entretanto, esse método modela três tipos de informações: geográfica, social e categórica. Basicamente, esse método cria um modelo personalizado para representar o comportamento geográfico do usuário alvo. Além disso, a influência social é representada por uma matriz de *check-ins* dos amigos do usuário e a influência categórica é usada por meio da criação de uma distribuição acumulativa dos *check-ins* do usuário alvo em cada categoria (72).

Em um segundo momento, comparamos os resultados alcançados pelo *Geo-Cat* com as técnicas de pós-processamento que são estado da arte em diversificação. Selecionamos os algoritmos *PM2* (17), *LD* (18), *Binom* (18) e *Geo-Div* (9), detalhados na Seção 2. Nesse caso, apesar de todos os métodos de recomendação específicos para POIs apresentarem resultados semelhantes, consideramos o *USG* como recomendador base para todos algoritmos de diversificação por três razões principais: (1) trata-se de um algoritmo específico para o cenário de POIs; (2) simples de ser instanciado; e (3) apresenta os melhores resultados na primeira etapa de avaliação, conforme veremos adiante.

5.1.3 Métricas de Avaliação

A fim de contrastar os resultados do **Geo-Cat** com as estratégias de diversificação existentes na literatura, selecionamos duas métricas relacionadas à efetividade (*Precision* e *Recall*) e quatro métricas relacionadas à satisfação dos usuários, sendo três métricas relacionadas à diversidade (*ILD*, *Genre Coverage* e PR_g), além de uma métrica relacionada à novidade (*EPC*) das listas de recomendações. Para todas essas métricas, avaliamos as listas de recomendações geradas por todos os métodos, considerando os tamanhos 5, 10 e 20. Além disso, essas mesmas métricas são utilizadas para avaliar o comportamento do **Geo-Cat** sobre todos os recomendadores base supracitados na seção 5.1.2.

- **Precision (Prec)**: representa a probabilidade de um POI recomendado ser relevante para o usuário. A relevância do POI é medida verificando se o POI está ou não no conjunto de teste do usuário (58).
- **Recall (Rec)**: representa a probabilidade que um POI relevante tem de ser recomendado para o usuário(58). Utilizamos a mesma noção de relevância do *precision*, mencionada acima.
- **EPC**: avalia a novidade da lista de recomendação de acordo com a popularidade dos POIs. Quanto mais alto o valor de EPC, menos populares são os POIs presentes na lista de recomendações (16).
- **ILD**: avalia a dissimilaridade entre as categorias dos POIs da lista de recomendação, tal qual mencionado anteriormente (18).

- **Genre Coverage (GC)**: representa a média de categorias relevantes recomendadas para o usuário (28). Novamente, a relevância, como explicada anteriormente, está relacionada às categorias dos POIs presentes no teste.
- **PRg**: avalia a distribuição dos POIs recomendados nas subáreas. Valores altos indicam que os POIs estão geograficamente distribuídos nas subáreas onde o usuário é mais ativo. Valores baixos indicam que os POIs recomendados estão concentrados em poucas subáreas ou espalhados aleatoriamente (17, 9).

5.2 Teoria da Utilidade Multiatributo (MAUT)

Conforme temos discutido ao longo deste trabalho, o processo de avaliação de SsR ainda pode ser considerado com um grande desafio em aberto. Diversas são as propostas de métricas de avaliação encontradas na literatura; entretanto, ainda não se tem um consenso sobre quais delas são necessárias para consolidar um novo SsR. É comum encontrar trabalhos que se limitam a avaliar métricas inerentes às técnicas implementadas, não validando simultaneamente requisitos de qualidade e satisfação dos usuários (24).

Diferente da literatura, este trabalho propõe avaliar todas as métricas de qualidade em conjunto. O objetivo é analisar o quão capazes são os SsR em equilibrar diversos critérios qualidade, tais como efetividade das recomendações e satisfação dos usuários. Com esse objetivo, propomos a adaptação da MAUT, originalmente proposta e avaliada no domínio de Teoria dos Jogos (29) para manipular problemas em que os resultados são dependentes de dois ou mais atributos (30). Mais especificamente, nossa proposta consiste em materializar essa teoria por meio de uma métrica que visa representar a eficácia de um método em função de diferentes dimensões de qualidade (31). Em nosso caso, o objetivo é utilizar essa métrica para comparar o *Geo-Cat* com outras estratégias de pós-processamento que também visam prover a diversificação das recomendações.

Formalmente, vamos considerar M como o conjunto de métricas que representa diferentes dimensões de qualidade (i.e., Prec, Rec, EPC, ILD, GC e PRg). Tomamos $m_{i,j}$ como a utilidade da métrica i para o método j (i.e., PM2, LD, Binom, Geo-Div e Geo-Cat). Conforme detalhado no Capítulo 2 e posteriormente sumarizado na Seção 5.1.3, normalmente, quanto maior o valor alcançado por um método para uma dada métrica, melhor é o método. Dessa forma, o valor de $m_{i,j}$ pode ser mapeado diretamente pelo valor obtido pelo método j na métrica i (4). Considerando o desempenho de todos os métodos para cada uma das métricas, nossa proposta consiste em calcular uma pontuação de utilidade (*utility score* - $U_{i,j}$) de um método específico em face dos demais métodos. Para isso, utilizamos o cálculo da *utility score* de um método por meio da normalização min-max dada ela Equação 5.2, em que $\min(m_i)$ e $\max(m_i)$ correspondem ao menor e maior valor obtido por um método

para uma métrica i , respectivamente. Essa normalização possibilita ranquear ou classificar os métodos de uma determinada métrica do melhor ($U_{i,j} = 1$) para o pior ($U_{i,j} = 0$).

$$U_{i,j} = (m_{i,j} - \min(m_i)) / (\max(m_i) - \min(m_i)) \quad (5.2)$$

A partir disso, a função de utilidade de um método U_j pode ser definida na forma de um modelo aditivo, como é apresentado na Equação 5.3, na qual w_i representa o peso da métrica i e normalmente satisfaz a condição $\sum_{i=1}^{|M|} w_i = 1$.

$$U_j = \sum_{i=1}^{|M|} U_{i,j} \cdot w_i \quad (5.3)$$

O peso padrão de cada atributo é $1/|M|$, e pode ter variações se, porventura, algumas das métricas forem consideradas mais importantes que as demais. Em nosso caso, assumimos que todas as métricas são independentes e igualmente importantes no processo de decisão (30). Assim, após obter os resultados do modelo aditivo da Equação 5.3 para cada método de pós-processamento, consideramos como melhor aquele que obtém maior resultado.

Dessa forma, nesta seção abordamos parcialmente a segunda hipótese deste trabalho: (H2): *É possível avaliar SsR considerando múltiplas perspectivas de qualidade, tais como acurácia, diversidade categórica e diversidade geográfica.* Para consolidar nossa segunda hipótese, resta avaliar a aplicação prática da MAUT em cenários de recomendação. Nas seções seguintes apresentamos os resultados das avaliações realizadas considerando a MAUT.

5.3 Resultados & Discussões

Primeiramente, nesta seção apresentamos os resultados de um estudo preliminar realizado para determinar a melhor configuração de parâmetros para o *Geo-Cat*, bem como para todos os outros algoritmos de diversificação. Em seguida, avaliamos nossa proposta em duas partes, por meio de: (1) um comparativo entre os resultados alcançados pelos recomendadores base e pelo *Geo-Cat* após realizar o pós-processamento das listas de recomendação. A partir dessa etapa, utilizando a MAUT, selecionamos o recomendador base que obteve os melhores resultados; e (2) outro comparativo entre resultados alcançados pelo *Geo-Cat* com os dos algoritmos de diversificação estado da arte, utilizando como recomendador base aquele que obteve os melhores resultados na etapa anterior. Por fim, agregamos os resultados por meio da métrica MAUT no intuito de identificar a melhor estratégia de recomendação capaz de equilibrar os principais requisitos de qualidade avaliados.

5.3.1 Configurações de Parâmetros dos Métodos de Diversificação

O comportamento do método de diversificação *Geo-Cat* pode ser modificado pela configuração dos parâmetros λ e β . Nesse sentido, λ é usado para configurar de forma geral o peso do fator diversidade, e β , para guiar o comportamento da diversidade para o fator geográfico e/ou categórico. Por exemplo, quanto mais alto é o valor de λ , mais diversificadas são as recomendações. Ao mesmo tempo, quanto mais alto é o valor de β , mais diversificadas geograficamente são as recomendações. Dessa maneira, para validar os parâmetros ideais para nossa abordagem, variamos os valores de λ e β entre 0 e 1 com intervalos de 0,25 a cada teste realizado sobre um conjunto de validação, resultando em 21 combinações possíveis (quando $\lambda = 0$ o valor de β não é considerado).

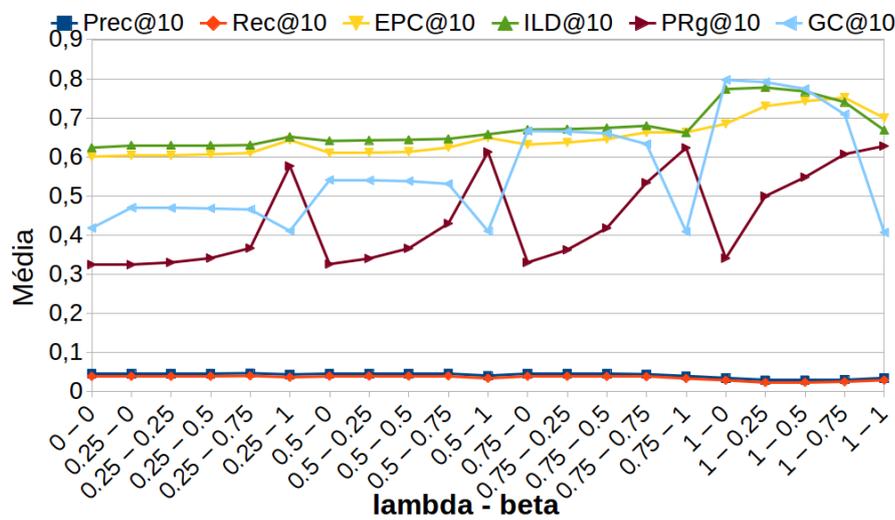


Figura 9 – Gráfico que apresenta curvas de resultados do *Geo-Cat* sob diferentes combinações das possíveis configurações para os parâmetros λ e β .

Dessa forma, é apresentado no gráfico da Figura 9 o comportamento do *Geo-Cat* sob diferentes configurações de parâmetros, considerando as recomendações de 10 POIs, sob todas as métricas de avaliação selecionadas (observamos comportamento muito semelhante para a recomendação de 5 e 20 POIs). Primeiramente, observa-se pouca variação em *precision* e *recall*. Os valores dessas métricas tendem a cair quando $\lambda = 1$ uma vez que o *Geo-Cat*, nesse caso, considera que apenas a diversidade é importante. Observa-se também que aumentar o valor do parâmetro λ (primeiro valor apresentado no eixo X) faz com que as métricas de diversidade alcancem melhores resultados. Por fim, percebe-se que o parâmetro β (segundo valor apresentado no eixo X) funciona perfeitamente como um *threshold* entre as diversidades geográfica e categórica. Logo, conforme o valor de β aumenta, observa-se que o valor de PR_g também aumenta, pois estamos atribuindo mais peso para a diversidade geográfica.

Para definir os melhores parâmetros para nossa estratégia, utilizamos a métrica MAUT considerando todas as métricas de acurácia, novidade e diversidade ao mesmo tempo. Também consideramos que todas as métricas de avaliação são igualmente importantes e possuem pesos iguais. De acordo com os resultados obtidos, a combinação

$\lambda = 0,75$ e $\beta = 0,75$ apresenta os melhores resultados em média para todas as cidades, considerando recomendações de 5, 10 e 20 POIs. Da mesma forma, executamos os mesmos testes de parâmetros para os algoritmos *baselines* de diversificação e pelas mesmas razões supracitadas selecionamos a melhor configuração para cada um: **Binom** ($\lambda = 0,75$); **Geo-Div** ($\lambda = 0,5$); **LD** ($\lambda = 0,25$); **PM2** ($\lambda = 1,0$).

5.3.2 Avaliação do Geo-Cat nos SsR Tradicionais

Nesta primeira análise, nosso objetivo foi avaliar se a utilização de nossa estratégia de pós-processamento *Geo-Cat* foi capaz de melhorar a qualidade das recomendações geradas pelos SsR base, sejam eles tradicionais adaptados para o problema de recomendação de POIs, sejam aqueles específicos para POIs. Os resultados dos testes realizados para as cidades de *Charlotte*, *Las Vegas* e *Phoenix* são apresentados nas Tabelas 1, 2 e 3, respectivamente. Em todas as cidades, para cada métrica e tamanho de lista de recomendação, comparamos os resultados do SR base com e sem o *Geo-Cat* utilizando o teste estatístico de *Wilcoxon* com 95% de confiança. De maneira geral, podemos observar que o *Geo-Cat* conseguiu melhorar os índices de diversidade de grande parte das métricas avaliadas para os SsR base utilizados. Além disso, são apresentados ganhos nessas métricas considerando todos os tamanhos N da lista de recomendação.

Primeiramente, considerando as métrica de efetividade (*Precision* e *Recall*), observamos que a utilização do *Geo-Cat* foi capaz de prover resultados estatisticamente melhores ou equivalentes àqueles obtidos pelos SsR base na maioria dos casos. Pequenas perdas de efetividade, estatisticamente relevantes, são observadas em poucos casos nas três cidades; entretanto, observa-se que em *Las Vegas* e *Phoenix* essas perdas, muito pequenas, ocorrem apenas nos SR base específicos para POIs (i.e. *USG*, *GeoSoCa* e *GeoMF*). Além disso, destacamos os resultados obtidos pelos métodos *Most Popular* e *Item Attribute-KNN*. Tradicionalmente, o primeiro apresenta apenas os POIs mais populares (i.e., mais frequentados); e o segundo é um SR baseado no conteúdo descritivo dos POIs (i.e. categorias) para fazer a recomendação. Em ambos os casos, conseguimos com nosso método apresentar POIs distintos dos anteriores, aumentando a diversidade e a novidade das recomendações, e ainda assim, com um alto índice de efetividade, aumentando as métricas de *Precision* e *Recall*.

Considerando as métricas de diversidade e de novidade, observamos novamente que a utilização do *Geo-Cat* foi capaz de prover resultados muito superiores aos obtidos pelos outros SsR. Em especial, nos SR tradicionais mais focados em efetividade, como os recomendadores *WRMF* e *USER-KNN* e nos SR específicos para POIs, como *USG*, *GeoSoCa* e *GeoMF*, igualmente focados em efetividade, conseguimos melhorar a diversidade e novidade dos itens sem apresentar grandes perdas de efetividade. Por exemplo, no caso da cidade de Phoenix, considerando a lista de 10 recomendações do *USG*, obtivemos uma melhora de 120% para a métrica de PR_g . Algumas exceções são observadas, por exemplo

| | Charlotte | | | | | |
|----------------------------|-----------|----------|----------|----------|----------|----------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 |
| ItemAttributeKNN | 0,0149 | 0,0052 | 0,9097 | 0,1805 | 0,0824 | 0,2232 |
| Geo-cat (ItemAttributeKNN) | 0,0244 ▲ | 0,0085 ▲ | 0,8863 ▼ | 0,4817 ▲ | 0,3629 ▲ | 0,4393 ▲ |
| MostPopular | 0,0455 | 0,0151 | 0,2506 | 0,6702 | 0,0842 | 0,3045 |
| Geo-cat (MostPopular) | 0,0437 ● | 0,0148 ● | 0,4851 ▲ | 0,6162 ▼ | 0,314 ▲ | 0,4619 ▲ |
| UserKNN | 0,0623 | 0,0223 | 0,456 | 0,6205 | 0,1682 | 0,2956 |
| Geo-cat (UserKNN) | 0,0544 ▼ | 0,0202 ● | 0,6564 ▲ | 0,6417 ▲ | 0,3967 ▲ | 0,4602 ▲ |
| WRMF | 0,064 | 0,0233 | 0,534 | 0,5996 | 0,1797 | 0,2819 |
| Geo-cat (WRMF) | 0,0566 ▼ | 0,0203 ● | 0,6769 ▲ | 0,6473 ▲ | 0,4008 ▲ | 0,4584 ▲ |
| USG | 0,0584 | 0,0218 | 0,436 | 0,6266 | 0,216 | 0,2988 |
| Geo-cat (USG) | 0,0551 ▼ | 0,0218 ● | 0,6635 ▲ | 0,6509 ▲ | 0,4189 ▲ | 0,4791 ▲ |
| GeoSoCa | 0,0255 | 0,0095 | 0,802 | 0,7704 | 0,0574 | 0,0958 |
| Geo-cat (GeoSoCa) | 0,0324 ▲ | 0,0090 ● | 0,8684 ▲ | 0,7709 ● | 0,2429 ▲ | 0,2311 ▲ |
| GeoMF | 0,039 | 0,0159 | 0,9375 | 0,7646 | 0,0503 | 0,1063 |
| Geo-cat (GeoMF) | 0,0375 ▼ | 0,0173 ▲ | 0,8608 ▼ | 0,767 ● | 0,3479 ▲ | 0,318 ▲ |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 |
| ItemAttributeKNN | 0,0141 | 0,0092 | 0,9073 | 0,2106 | 0,1303 | 0,2919 |
| Geo-cat (ItemAttributeKNN) | 0,0218 ▲ | 0,0146 ▲ | 0,8855 ▼ | 0,4377 ▲ | 0,4422 ▲ | 0,5473 ▲ |
| MostPopular | 0,0412 | 0,0272 | 0,3378 | 0,6185 | 0,1656 | 0,4503 |
| Geo-cat (MostPopular) | 0,0408 ● | 0,0274 ● | 0,4828 ▲ | 0,6041 ▼ | 0,3895 ▲ | 0,5736 ▲ |
| UserKNN | 0,0538 | 0,0389 | 0,5057 | 0,596 | 0,2666 | 0,4475 |
| Geo-cat (UserKNN) | 0,0524 ● | 0,0384 ● | 0,6409 ▲ | 0,6256 ▲ | 0,501 ▲ | 0,5966 ▲ |
| WRMF | 0,0562 | 0,0401 | 0,5766 | 0,5916 | 0,2708 | 0,4207 |
| Geo-cat (WRMF) | 0,0519 ● | 0,0373 ● | 0,673 ▲ | 0,6367 ▲ | 0,5113 ▲ | 0,5965 ▲ |
| USG | 0,051 | 0,0376 | 0,4959 | 0,5889 | 0,3173 | 0,4489 |
| Geo-cat (USG) | 0,0529 ● | 0,0405 ● | 0,6553 ▲ | 0,6331 ▲ | 0,5405 ▲ | 0,6109 ▲ |
| GeoSoCa | 0,0230 | 0,0163 | 0,8561 | 0,7705 | 0,0986 | 0,1827 |
| Geo-cat (GeoSoCa) | 0,0219 ● | 0,0184 ▲ | 0,8698 ▲ | 0,7765 ▲ | 0,2882 ▲ | 0,3499 ▲ |
| GeoMF | 0,0349 | 0,0282 | 0,9172 | 0,766 | 0,087 | 0,1928 |
| Geo-cat (GeoMF) | 0,036 ▲ | 0,0295 ● | 0,8696 ▲ | 0,7716 ▲ | 0,4012 ▲ | 0,4583 ▲ |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 |
| ItemAttributeKNN | 0,0123 | 0,0163 | 0,9092 | 0,2705 | 0,185 | 0,3915 |
| Geo-cat (ItemAttributeKNN) | 0,0173 ▲ | 0,0225 ▲ | 0,8918 ▼ | 0,387 ▲ | 0,4346 ▲ | 0,6199 ▲ |
| MostPopular | 0,0351 | 0,045 | 0,4179 | 0,5245 | 0,2698 | 0,559 |
| Geo-cat (MostPopular) | 0,0358 ● | 0,0467 ● | 0,5119 ▲ | 0,564 ▼ | 0,4155 ▲ | 0,7038 ▲ |
| UserKNN | 0,0461 | 0,0651 | 0,5591 | 0,5832 | 0,36 | 0,5855 |
| Geo-cat (UserKNN) | 0,0472 ▲ | 0,0679 ▲ | 0,643 ▲ | 0,6133 ▲ | 0,5418 ▲ | 0,7103 ▲ |
| WRMF | 0,0469 | 0,0669 | 0,6159 | 0,592 | 0,3498 | 0,5655 |
| Geo-cat (WRMF) | 0,0482 ▲ | 0,0677 ● | 0,6792 ▲ | 0,6212 ▲ | 0,5501 ▲ | 0,7096 ▲ |
| USG | 0,0442 | 0,0644 | 0,5537 | 0,5763 | 0,4155 | 0,5834 |
| Geo-cat (USG) | 0,0464 ● | 0,0681 ● | 0,6534 ▲ | 0,6156 ▲ | 0,5959 ▲ | 0,7223 ▲ |
| GeoSoCa | 0,0202 | 0,0295 | 0,8646 | 0,7681 | 0,1531 | 0,3138 |
| Geo-cat (GeoSoCa) | 0,0201 ● | 0,0315 ● | 0,8738 ▲ | 0,7767 ▲ | 0,2984 ▲ | 0,4572 ▲ |
| GeoMF | 0,0312 | 0,0495 | 0,9112 | 0,7658 | 0,1396 | 0,3273 |
| Geo-cat (GeoMF) | 0,0337 ● | 0,0499 ● | 0,8768 ▲ | 0,7762 ▲ | 0,3809 ▲ | 0,559 ▲ |

Tabela 1 – Pós-Processamento do Geo-Cat sobre as listas geradas pelos recomendadores base para a cidade de Charlotte. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatisticamente significativos e ▼ representa perdas estatisticamente significativas. Geo-Cat consegue melhorar todos recomendadores base em termos de diversidade, sem que ocorra perda da efetividade.

para o algoritmo *Item Attribute-KNN* sobre a métrica de *EPC*, porém, com uma diferença muito pequena (inferior à 1%). O *Item Attribute-KNN*, por si só, alcança bons resultados de novidade, pois apresenta POIs semelhantes aos POIs mais bem avaliados pelo usuário, porém, com uma acurácia muito baixa, inferior a todos os demais. Essa pequena diferença na diversidade e novidade é significativamente compensada na efetividade das recomendações, uma vez que o *Geo-Cat* é capaz de melhorar em até 56% a métrica *Precision* desse algoritmo. Esses resultados mostram que nossa abordagem é uma potencial estratégia a ser aplicada em domínios reais. Além disso, conseguimos abordar por completo nossa primeira

| Algoritmo | Las Vegas | | | | | |
|----------------------------|-----------|----------|----------|----------|----------|----------|
| | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 |
| ItemAttributeKNN | 0,0146 | 0,0060 | 0,8740 | 0,1835 | 0,1129 | 0,2548 |
| Geo-cat (ItemAttributeKNN) | 0,0228 ▲ | 0,0089 ▲ | 0,8625 ▼ | 0,4846 ▲ | 0,3878 ▲ | 0,4869 ▲ |
| MostPopular | 0,0228 | 0,0090 | 0,1062 | 0,7242 | 0,1109 | 0,2170 |
| Geo-cat (MostPopular) | 0,0391 ▲ | 0,0148 ▲ | 0,4255 ▲ | 0,7071 ▼ | 0,3354 ▲ | 0,4720 ▲ |
| UserKNN | 0,0501 | 0,0200 | 0,5276 | 0,6115 | 0,2227 | 0,2776 |
| Geo-cat (UserKNN) | 0,0506 ● | 0,0206 ● | 0,6241 ▲ | 0,6790 ▲ | 0,3921 ▲ | 0,4784 ▲ |
| WRMF | 0,0520 | 0,0204 | 0,4714 | 0,6013 | 0,2106 | 0,2807 |
| Geo-cat (WRMF) | 0,0521 ● | 0,0207 ● | 0,5885 ▲ | 0,6741 ▲ | 0,3844 ▲ | 0,4854 ▲ |
| USG | 0,0548 | 0,0237 | 0,5805 | 0,6070 | 0,2503 | 0,2740 |
| Geo-cat (USG) | 0,0497 ▼ | 0,0214 ● | 0,6543 ▲ | 0,6868 ▲ | 0,4216 ▲ | 0,5033 ▲ |
| GeoSoCa | 0,0234 | 0,0199 | 0,9599 | 0,7773 | 0,0492 | 0,0558 |
| Geo-cat (GeoSoCa) | 0,0321 ▲ | 0,0187 ▼ | 0,9330 ▼ | 0,7823 ▲ | 0,2536 ▲ | 0,2149 ▲ |
| GeoMF | 0,0340 | 0,0335 | 0,8917 | 0,7781 | 0,0434 | 0,0569 |
| Geo-cat (GeoMF) | 0,0477 ▲ | 0,0268 ▼ | 0,8701 ▼ | 0,7829 ▲ | 0,3260 ▲ | 0,2621 ▲ |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 |
| ItemAttributeKNN | 0,0138 | 0,0110 | 0,8784 | 0,2113 | 0,1640 | 0,3197 |
| Geo-cat (ItemAttributeKNN) | 0,0193 ▲ | 0,0150 ▲ | 0,8632 ▼ | 0,4333 ▲ | 0,4721 ▲ | 0,5873 ▲ |
| MostPopular | 0,0218 | 0,0167 | 0,1981 | 0,6984 | 0,1540 | 0,3564 |
| Geo-cat (MostPopular) | 0,0331 ▲ | 0,0254 ▲ | 0,4264 ▲ | 0,7048 ▲ | 0,3951 ▲ | 0,5970 ▲ |
| UserKNN | 0,0428 | 0,0338 | 0,5561 | 0,6293 | 0,2893 | 0,4180 |
| Geo-cat (UserKNN) | 0,0443 ● | 0,0357 ▲ | 0,6300 ▲ | 0,6742 ▲ | 0,4852 ▲ | 0,6170 ▲ |
| WRMF | 0,0458 | 0,0356 | 0,5159 | 0,6134 | 0,2742 | 0,4142 |
| Geo-cat (WRMF) | 0,0468 ● | 0,0366 ● | 0,5956 ▲ | 0,6653 ▲ | 0,4761 ▲ | 0,6202 ▲ |
| USG | 0,0460 | 0,0392 | 0,6006 | 0,6236 | 0,3248 | 0,4181 |
| Geo-cat (USG) | 0,0446 ● | 0,0383 ● | 0,6620 ▲ | 0,6795 ▲ | 0,5342 ▲ | 0,6329 ▲ |
| GeoSoCa | 0,0201 | 0,0322 | 0,9516 | 0,7780 | 0,0767 | 0,1074 |
| Geo-cat (GeoSoCa) | 0,0205 ▲ | 0,0298 ▼ | 0,9299 ▼ | 0,7857 ▲ | 0,2803 ▲ | 0,3095 ▲ |
| GeoMF | 0,0291 | 0,0551 | 0,9115 | 0,7774 | 0,0739 | 0,1102 |
| Geo-cat (GeoMF) | 0,0273 ▼ | 0,0513 ▼ | 0,8741 ▼ | 0,7860 ▲ | 0,3639 ▲ | 0,3912 ▲ |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 |
| ItemAttributeKNN | 0,0122 | 0,0186 | 0,8783 | 0,2558 | 0,2185 | 0,4068 |
| Geo-cat (ItemAttributeKNN) | 0,0155 ▲ | 0,0239 ▲ | 0,8687 ▼ | 0,3816 ▲ | 0,4650 ▲ | 0,6473 ▲ |
| MostPopular | 0,0224 | 0,0320 | 0,3068 | 0,6770 | 0,2216 | 0,5068 |
| Geo-cat (MostPopular) | 0,0271 ▲ | 0,0404 ▲ | 0,4375 ▲ | 0,7005 ▲ | 0,4006 ▲ | 0,7085 ▲ |
| UserKNN | 0,0375 | 0,0589 | 0,5902 | 0,6440 | 0,3561 | 0,5780 |
| Geo-cat (UserKNN) | 0,0385 ● | 0,0603 ● | 0,6389 ▲ | 0,6695 ▲ | 0,5230 ▲ | 0,7254 ▲ |
| WRMF | 0,0406 | 0,0634 | 0,5611 | 0,6292 | 0,3412 | 0,5702 |
| Geo-cat (WRMF) | 0,0407 ● | 0,0635 ● | 0,6085 ▲ | 0,6569 ▲ | 0,5104 ▲ | 0,7253 ▲ |
| USG | 0,0385 | 0,0645 | 0,6244 | 0,6396 | 0,3995 | 0,5760 |
| Geo-cat (USG) | 0,0386 ● | 0,0643 ● | 0,6633 ▲ | 0,6736 ▲ | 0,5863 ▲ | 0,7388 ▲ |
| GeoSoCa | 0,0163 | 0,0515 | 0,9498 | 0,7778 | 0,1159 | 0,1982 |
| Geo-cat (GeoSoCa) | 0,0178 ▲ | 0,0569 ▲ | 0,9291 ▼ | 0,7870 ▲ | 0,2675 ▲ | 0,3774 ▲ |
| GeoMF | 0,0244 | 0,0898 | 0,9108 | 0,7773 | 0,1160 | 0,1993 |
| Geo-cat (GeoMF) | 0,0223 ▼ | 0,0969 ▲ | 0,8787 ▼ | 0,7888 ▲ | 0,3333 ▲ | 0,4655 ▲ |

Tabela 2 – Pós-Processamento do Geo-Cat sobre as listas geradas pelos recomendadores base para a cidade de Las Vegas. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatisticamente significativos e ▼ representa perdas estatisticamente significativas. Geo-Cat consegue melhorar todos recomendadores base em termos de diversidade, sem que ocorra perda da efetividade.

hipótese de trabalho: (*H1*): *É possível explorar a diversidade categórica e geográfica, simultaneamente, sem perdas relevantes de efetividade*).

Para finalizar nossa avaliação, resta comparar o desempenho de nossa estratégia frente as outras de diversificação. Conforme já discutido, todas as estratégias a serem comparadas, assim como o próprio *Geo-Cat*, são estágios de pós-processamento e, portanto, dependem de um recomendador base. Assim, realizamos uma comparação do desempenho dos três algoritmos de recomendação específicos para POIs (*USG*, *GeoSoCa* e *GeoMF*)

| Algoritmo | Phoenix | | | | | |
|----------------------------|----------|----------|----------|----------|----------|----------|
| | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 |
| ItemAttributeKNN | 0,0064 | 0,0021 | 0,9255 | 0,1852 | 0,0241 | 0,2573 |
| Geo-cat (ItemAttributeKNN) | 0,0121 ▲ | 0,0039 ▲ | 0,9214 ▼ | 0,4735 ▲ | 0,1710 ▲ | 0,5340 ▲ |
| MostPopular | 0,0141 | 0,0046 | 0,3041 | 0,4682 | 0,0108 | 0,3405 |
| Geo-cat (MostPopular) | 0,0188 ▲ | 0,0064 ▲ | 0,5495 ▲ | 0,5404 ▲ | 0,1445 ▲ | 0,5706 ▲ |
| UserKNN | 0,0386 | 0,0140 | 0,6907 | 0,4762 | 0,1180 | 0,3486 |
| Geo-cat (UserKNN) | 0,0367 ● | 0,0135 ● | 0,7732 ▲ | 0,5757 ▲ | 0,2862 ▲ | 0,5327 ▲ |
| WRMF | 0,0319 | 0,0111 | 0,5934 | 0,4734 | 0,0792 | 0,3521 |
| Geo-cat (WRMF) | 0,0331 ● | 0,0117 ● | 0,7093 ▲ | 0,5560 ▲ | 0,2551 ▲ | 0,5433 ▲ |
| USG | 0,0460 | 0,0182 | 0,7407 | 0,4695 | 0,1537 | 0,3533 |
| Geo-cat (USG) | 0,0380 ▼ | 0,0153 ▼ | 0,7989 ▲ | 0,5926 ▲ | 0,3261 ▲ | 0,5491 ▲ |
| GeoSoCa | 0,0199 | 0,0065 | 0,9456 | 0,7725 | 0,0143 | 0,0741 |
| Geo-cat (GeoSoCa) | 0,0234 ▲ | 0,0090 ▲ | 0,8952 ▼ | 0,7785 ▲ | 0,0746 ▲ | 0,2409 ▲ |
| GeoMF | 0,0170 | 0,0071 | 0,9390 | 0,7739 | 0,0127 | 0,0773 |
| Geo-cat (GeoMF) | 0,0195 ▲ | 0,0068 ▼ | 0,9339 ▼ | 0,7813 ▲ | 0,1282 ▲ | 0,3811 ▲ |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 |
| ItemAttributeKNN | 0,0056 | 0,0035 | 0,9275 | 0,2125 | 0,0415 | 0,3282 |
| Geo-cat (ItemAttributeKNN) | 0,0095 ▲ | 0,0058 ▲ | 0,9234 ▼ | 0,3990 ▲ | 0,1928 ▲ | 0,6076 ▲ |
| MostPopular | 0,0134 | 0,0085 | 0,3780 | 0,4776 | 0,0254 | 0,4474 |
| Geo-cat (MostPopular) | 0,0166 ▲ | 0,0110 ▲ | 0,5316 ▲ | 0,5121 ▲ | 0,1713 ▲ | 0,6663 ▲ |
| UserKNN | 0,0340 | 0,0240 | 0,7129 | 0,5012 | 0,1771 | 0,4789 |
| Geo-cat (UserKNN) | 0,0338 ● | 0,0246 ● | 0,7698 ▲ | 0,5660 ▲ | 0,3664 ▲ | 0,6533 ▲ |
| WRMF | 0,0286 | 0,0197 | 0,6303 | 0,4908 | 0,1276 | 0,4727 |
| Geo-cat (WRMF) | 0,0293 ● | 0,0206 ● | 0,7058 ▲ | 0,5545 ▲ | 0,3232 ▲ | 0,6596 ▲ |
| USG | 0,0371 | 0,0287 | 0,7495 | 0,4946 | 0,2220 | 0,4836 |
| Geo-cat (USG) | 0,0356 ● | 0,0278 ● | 0,7927 ▲ | 0,5772 ▲ | 0,4301 ▲ | 0,6671 ▲ |
| GeoSoCa | 0,0171 | 0,0112 | 0,9492 | 0,7752 | 0,0244 | 0,1500 |
| Geo-cat (GeoSoCa) | 0,0180 ▲ | 0,0108 ▼ | 0,8990 ▼ | 0,7843 ▲ | 0,0827 ▲ | 0,3180 ▲ |
| GeoMF | 0,0192 | 0,0093 | 0,9411 | 0,7744 | 0,0230 | 0,1443 |
| Geo-cat (GeoMF) | 0,0185 ▼ | 0,0122 ▲ | 0,9314 ▼ | 0,7867 ▲ | 0,1352 ▲ | 0,4760 ▲ |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 |
| ItemAttributeKNN | 0,0052 | 0,0066 | 0,9275 | 0,2520 | 0,0657 | 0,4175 |
| Geo-cat (ItemAttributeKNN) | 0,0070 ▲ | 0,0086 ▲ | 0,9242 ▼ | 0,3415 ▲ | 0,1842 ▲ | 0,6412 ▲ |
| MostPopular | 0,0121 | 0,0149 | 0,4446 | 0,4181 | 0,0640 | 0,5630 |
| Geo-cat (MostPopular) | 0,0136 ▲ | 0,0176 ▲ | 0,5420 ▲ | 0,4807 ▲ | 0,1812 ▲ | 0,7270 ▲ |
| UserKNN | 0,0286 | 0,0405 | 0,7354 | 0,5277 | 0,2419 | 0,6105 |
| Geo-cat (UserKNN) | 0,0295 ● | 0,0422 ● | 0,7718 ▲ | 0,5637 ▲ | 0,4062 ▲ | 0,7398 ▲ |
| WRMF | 0,0246 | 0,0338 | 0,6669 | 0,5190 | 0,1885 | 0,6007 |
| Geo-cat (WRMF) | 0,0258 ▲ | 0,0357 ▲ | 0,7176 ▲ | 0,5554 ▲ | 0,3515 ▲ | 0,7412 ▲ |
| USG | 0,0303 | 0,0464 | 0,7628 | 0,5221 | 0,2971 | 0,6149 |
| Geo-cat (USG) | 0,0305 ● | 0,0465 ● | 0,7916 ▲ | 0,5678 ▲ | 0,4853 ▲ | 0,7524 ▲ |
| GeoSoCa | 0,0144 | 0,0183 | 0,9391 | 0,7739 | 0,0419 | 0,2499 |
| Geo-cat (GeoSoCa) | 0,0192 ▲ | 0,0241 ▲ | 0,9295 ▼ | 0,7884 ▲ | 0,1270 ▲ | 0,5343 ▲ |
| GeoMF | 0,0183 | 0,0199 | 0,9296 | 0,7745 | 0,0391 | 0,2533 |
| Geo-cat (GeoMF) | 0,0192 ▲ | 0,0241 ▲ | 0,9295 ● | 0,7884 ▲ | 0,1270 ▲ | 0,5343 ▲ |

Tabela 3 – Pós-Processamento do Geo-Cat sobre as listas geradas pelos recomendadores base para a cidade de Phoenix. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatisticamente significativos e ▼ representa perdas estatisticamente significativas. Geo-Cat consegue melhorar todos recomendadores base em termos de diversidade, sem que ocorra perda da efetividade.

utilizando as tabelas acima. Primeiramente, observa-se que o método *USG* é superado pelo *GeoMF* apenas na métrica *Recall* na cidade de *Las Vegas*, conforme pode ser observado na Tabela 2. Porém, em todos os outros casos nas métricas *Precision* e *Recall*, o *USG* apresenta resultados muito superiores aos que são apresentados pelos métodos *GeoMF* e *GeoSoCa*, demonstrando a superioridade do *USG* frente aos outros dois métodos de recomendação de POIs, no que tange aos resultados de efetividade. Contudo, é importante destacar também um comparativo sobre as métricas de diversidade e novidade entre esses recomendadores. Isso posto, observa-se que o *USG* sempre apresenta resultados abaixo

dos outros dois para as métricas *EPC* e *ILD*. Entretanto, o *USG* é muito superior aos outros dois nas métricas de *PRg* e *Genre Coverage*, em todos os casos analisados. A fim de confirmar qual é o melhor SR de POIs frente as múltiplas dimensões de qualidade, apresentamos na Tabela 4 o resultado do cálculo da MAUT entre os métodos *USG*, *GeoMF* e *GeoSoCa*. Observa-se portanto que, o *USG* é comprovadamente superior aos outros dois métodos em todos os casos analisados. Dessa forma, utilizamos o *USG* como SR base para todas os algoritmos de diversificação, cujos resultados são apresentados na próxima seção.

| Charlotte | | | | | | | |
|-----------|---------|--------|--------|--------|--------|--------|---------------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 | MAUT |
| GeoMF | 0,4103 | 0,5203 | 1,0000 | 0,9597 | 0,0000 | 0,0517 | 0,4884 |
| GeoSoCa | 0,0000 | 0,0000 | 0,7298 | 1,0000 | 0,0428 | 0,0000 | 0,2943 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,6640 |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 | MAUT |
| GeoMF | 0,4250 | 0,5587 | 1,0000 | 0,9752 | 0,0000 | 0,0379 | 0,4975 |
| GeoSoCa | 0,0000 | 0,0000 | 0,8550 | 1,0000 | 0,0504 | 0,0000 | 0,3163 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,6640 |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 | MAUT |
| GeoMF | 0,4583 | 0,5731 | 1,0000 | 0,9880 | 0,0000 | 0,0501 | 0,5095 |
| GeoSoCa | 0,0000 | 0,0000 | 0,8697 | 1,0000 | 0,0489 | 0,0000 | 0,3185 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,6640 |
| Las Vegas | | | | | | | |
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 | MAUT |
| GeoMF | 0,3376 | 1,0000 | 0,8202 | 1,0000 | 0,0000 | 0,0050 | 0,5250 |
| GeoSoCa | 0,0000 | 0,0000 | 1,0000 | 0,9953 | 0,0280 | 0,0000 | 0,3359 |
| USG | 1,0000 | 0,2794 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,5444 |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 | MAUT |
| GeoMF | 0,3475 | 1,0000 | 0,8858 | 0,9961 | 0,0000 | 0,0090 | 0,5376 |
| GeoSoCa | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,0112 | 0,0000 | 0,3339 |
| USG | 1,0000 | 0,3057 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,5487 |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 | MAUT |
| GeoMF | 0,3649 | 1,0000 | 0,8801 | 0,9964 | 0,0004 | 0,0029 | 0,5386 |
| GeoSoCa | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,0000 | 0,0000 | 0,3320 |
| USG | 1,0000 | 0,3394 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,5543 |
| Phoenix | | | | | | | |
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 | MAUT |
| GeoMF | 0,0000 | 0,0513 | 0,9678 | 1,0000 | 0,0000 | 0,0115 | 0,3371 |
| GeoSoCa | 0,1000 | 0,0000 | 1,0000 | 0,9954 | 0,0113 | 0,0000 | 0,3497 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,6640 |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 | MAUT |
| GeoMF | 0,1050 | 0,0000 | 0,9594 | 0,9971 | 0,0000 | 0,0000 | 0,3422 |
| GeoSoCa | 0,0000 | 0,0979 | 1,0000 | 1,0000 | 0,0070 | 0,0168 | 0,3522 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,6640 |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 | MAUT |
| GeoMF | 0,2453 | 0,0468 | 0,9461 | 1,0000 | 0,0000 | 0,0093 | 0,3731 |
| GeoSoCa | 0,0000 | 0,0000 | 1,0000 | 0,9976 | 0,0109 | 0,0000 | 0,3334 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,0000 | 1,0000 | 1,0000 | 0,6640 |

Tabela 4 – Aplicação da MAUT para selecionar o melhor SR base entre os SsR *USG*, *GeoMF* e *GeoSoCa*. Conforme podemos observar, o *USG* é aquele que apresenta o melhor resultado sob as várias dimensões de qualidade.

5.3.3 Avaliação do Geo-Cat com os Baselines

A segunda análise consiste em comparar o *Geo-Cat* com outras estratégias de diversificação de POIs existentes na literatura. Nestas análises, consideramos como SR base o *USG* por dois motivos: (1) ser específico para o cenário de POIs e, (2) por ser considerado

o SR que apresentou os melhores resultados comparados com todos os outros avaliados neste trabalho. Os parâmetros considerados em todas as estratégias foram aqueles descritos na Seção 5.3.1. Novamente, comparamos os resultados utilizando o teste de *Wilcoxon* com 95% de confiança. Assim, os resultados das comparações realizadas para as cidades de *Charlotte*, *Las Vegas* e *Phoenix* são apresentados nas tabelas 5, 6 e 7, respectivamente. Além disso, resumimos os resultados nos gráficos apresentados nas Figuras 11, 12 e 10 que apresentam diferenças percentuais de cada estratégia aplicada sobre o USG. Para cada estratégia, calculamos o quanto ela ganhou ou perdeu em relação aos resultados do SR base (*USG*), e plotamos essa variação nos gráficos.

Primeiramente, analisando os resultados referentes às métricas de efetividade, todas as estratégias de pós-processamento apresentaram pequenas variações negativas, na maioria das vezes sem significância estatística, demonstrando que grande parte dessas estratégias são capazes de gerar recomendações mais diversificadas sem afetar a eficácia delas. Por outro lado, do ponto de vista das métricas de diversidade e novidade (i.e. *EPC*, *ILD*, *PR_g* e *GC*), é possível observar que nenhuma das estratégias de pós-processamento apresenta os melhores resultados sobre todas as métricas consideradas. Por exemplo, o LD é o método que apresenta os melhores resultados quando observamos a métrica de *ILD*. Esse método utiliza a dissimilaridade entre os POIs para aumentar a diversidade das recomendações, ou seja, recomenda POIs o mais diferente possível entre si. Apesar de essa estratégia melhorar o *ILD*, este não é capaz de apresentar melhorias expressivas em outras métricas, sobretudo em diversificação geográfica (*PR_g@N*). Outro exemplo é o método Geo-Div, que em sua proposta original visa à diversificação geográfica dos POIs. Apesar de conseguir apresentar os melhores resultados em *PGr*, para as demais métricas apresenta resultados bem inferiores aos outros métodos analisados. Esse é um fato importante de se destacar, uma vez que na maioria dos trabalhos apresentados na literatura, apenas um ou duas dessas métricas são consideradas nas avaliações. Contudo, observando os resultados referentes ao Geo-Cat, temos que seus resultados estão sempre entre os melhores obtidos para cada uma das métricas avaliadas, apresentando-se como a melhor solução dentre todas as estratégias.

| Charlotte | | | | | | |
|-----------|----------|----------|----------|----------|----------|----------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 |
| USG | 0,0634 | 0,0246 | 0,4770 | 0,6268 | 0,2164 | 0,2988 |
| PM2 | 0,0588 ● | 0,0228 ● | 0,6326 ▲ | 0,4693 ▼ | 0,2398 ▲ | 0,3658 ▲ |
| LD | 0,0638 ● | 0,0250 ● | 0,4884 ● | 0,7944 ▲ | 0,2053 ▼ | 0,3034 ● |
| Geo-Div | 0,0589 ● | 0,0233 ● | 0,6679 ▲ | 0,5655 ▼ | 0,4499 ▲ | 0,2732 ▼ |
| Binom | 0,0633 ● | 0,0246 ● | 0,4768 ● | 0,6924 ▲ | 0,2108 ● | 0,3075 ▲ |
| Geo-Cat | 0,0551 ▼ | 0,0218 ● | 0,6635 ▲ | 0,6509 ▼ | 0,4189 ▲ | 0,4791 ▲ |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 |
| USG | 0,0544 | 0,0410 | 0,5265 | 0,5891 | 0,3181 | 0,4486 |
| PM2 | 0,0528 ● | 0,0398 ● | 0,6335 ▲ | 0,5164 ▼ | 0,3259 ● | 0,5157 ▲ |
| LD | 0,0520 ● | 0,0396 ● | 0,5372 ● | 0,7688 ▲ | 0,3167 ● | 0,4612 ▲ |
| Geo-Div | 0,0502 ● | 0,0377 ● | 0,6788 ▲ | 0,5823 ● | 0,6025 ▲ | 0,4068 ▼ |
| Binom | 0,0546 ● | 0,0417 ● | 0,5247 ● | 0,6316 ▲ | 0,3168 ● | 0,4597 ▲ |
| Geo-Cat | 0,0529 ● | 0,0405 ● | 0,6553 ▲ | 0,6331 ▲ | 0,5405 ▲ | 0,6109 ▲ |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 |
| USG | 0,0462 | 0,0682 | 0,5746 | 0,5768 | 0,4162 | 0,5839 |
| PM2 | 0,0467 ● | 0,0684 ● | 0,6418 ● | 0,5595 ▼ | 0,4164 ● | 0,6401 ▲ |
| LD | 0,0418 ● | 0,0616 ▼ | 0,5790 ● | 0,7300 ▲ | 0,4216 ● | 0,6090 ▲ |
| Geo-Div | 0,0435 ● | 0,0642 ● | 0,6825 ▲ | 0,5936 ▲ | 0,6481 ▲ | 0,5690 ▼ |
| Binom | 0,0466 ● | 0,0692 ● | 0,5750 ● | 0,5931 ▲ | 0,4175 ● | 0,5961 ▲ |
| Geo-Cat | 0,0464 ● | 0,0681 ● | 0,6534 ▲ | 0,6156 ▲ | 0,5959 ▲ | 0,7223 ▲ |

Tabela 5 – Comparativo entre as *Baselines* e o Geo-Cat para a cidade de Charlotte. A avaliação estatística foi feita utilizando-se o teste de *Wilcoxon* com 95% de intervalo de confiança. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatísticos e ▼ representa perdas estatísticas.

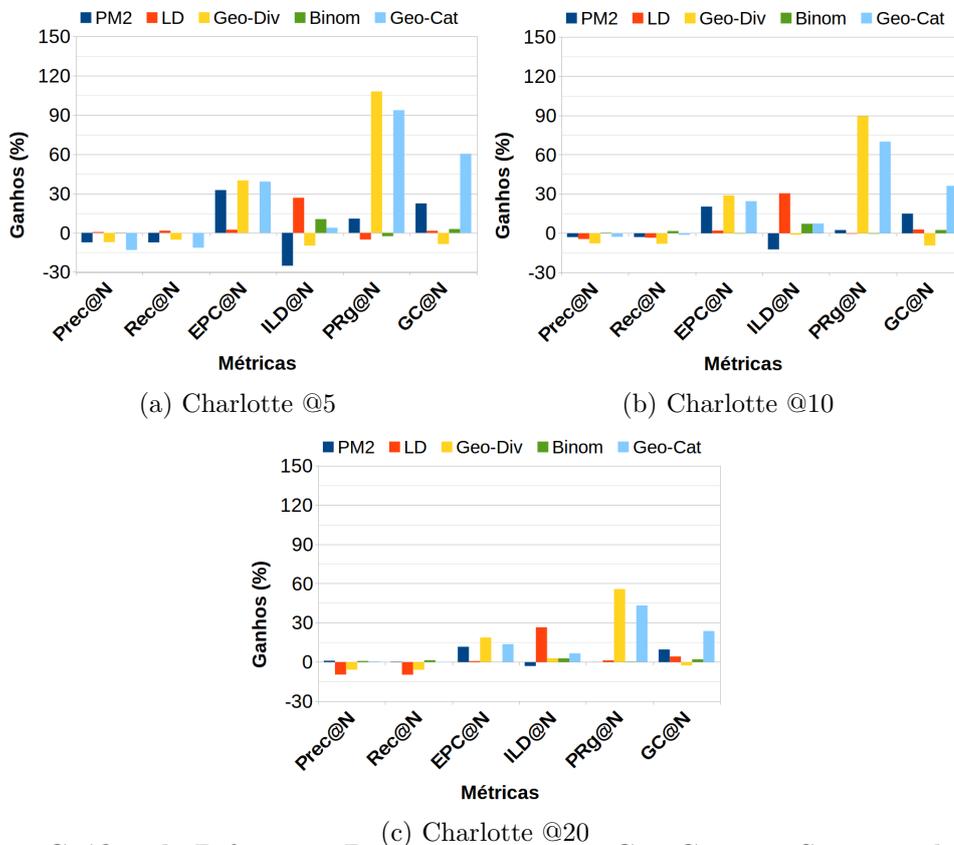


Figura 10 – Gráfico de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage da cidade de Charlotte.

| Las Vegas | | | | | | |
|-----------|----------|----------|----------|----------|----------|----------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 |
| USG | 0,0548 | 0,0237 | 0,5805 | 0,6070 | 0,2503 | 0,2740 |
| PM2 | 0,0502 ● | 0,0220 ● | 0,6512 ▲ | 0,5487 ▼ | 0,2613 ▲ | 0,3848 ▲ |
| LD | 0,0518 ● | 0,0220 ● | 0,5843 ● | 0,7809 ▲ | 0,2516 ● | 0,2875 ▲ |
| Geo-Div | 0,0500 ▼ | 0,0219 ● | 0,6290 ▲ | 0,6382 ▲ | 0,4579 ▲ | 0,2666 ▼ |
| Binom | 0,0537 ● | 0,0233 ● | 0,5806 ● | 0,6917 ▲ | 0,2514 ● | 0,2875 ▲ |
| Geo-Cat | 0,0497 ▼ | 0,0214 ● | 0,6543 ▲ | 0,6868 ▲ | 0,4216 ▲ | 0,5033 ▲ |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 |
| USG | 0,0460 | 0,0392 | 0,6006 | 0,6236 | 0,3248 | 0,4181 |
| PM2 | 0,0443 ● | 0,0384 ● | 0,6527 ▲ | 0,5928 ▼ | 0,3386 ▲ | 0,5247 ▲ |
| LD | 0,0420 ▼ | 0,0355 ▼ | 0,6029 ● | 0,7731 ▲ | 0,3330 ▲ | 0,4441 ▲ |
| Geo-Div | 0,0409 ▼ | 0,0342 ▼ | 0,6489 ▲ | 0,6578 ▲ | 0,6127 ▲ | 0,4102 ▼ |
| Binom | 0,0455 ● | 0,0385 ● | 0,5981 ● | 0,6677 ▲ | 0,3254 ● | 0,4360 ▲ |
| Geo-Cat | 0,0446 ● | 0,0383 ● | 0,6620 ▲ | 0,6795 ▲ | 0,5342 ▲ | 0,6329 ▲ |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 |
| USG | 0,0385 | 0,0645 | 0,6244 | 0,6396 | 0,3995 | 0,5760 |
| PM2 | 0,0383 ● | 0,0639 ● | 0,6572 ▲ | 0,6322 ▼ | 0,4079 ▲ | 0,6434 ▲ |
| LD | 0,0328 ▼ | 0,0539 ▼ | 0,6237 ● | 0,7684 ▲ | 0,4231 ▲ | 0,6136 ▲ |
| Geo-Div | 0,0362 ● | 0,0610 ● | 0,6665 ▲ | 0,6639 ▲ | 0,6484 ▲ | 0,5750 ● |
| Binom | 0,0376 ● | 0,0629 ● | 0,6238 ● | 0,6574 ▲ | 0,4007 ● | 0,5944 ▲ |
| Geo-Cat | 0,0386 ● | 0,0643 ● | 0,6633 ▲ | 0,6736 ▲ | 0,5863 ▲ | 0,7388 ▲ |

Tabela 6 – Comparativo entre as *Base Lines* e o Geo-Cat para a cidade de Las Vegas. A avaliação estatística foi feita utilizando-se o teste de *Wilcoxon* com 95% de intervalo de confiança. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatísticos e ▼ representa perdas estatísticas.

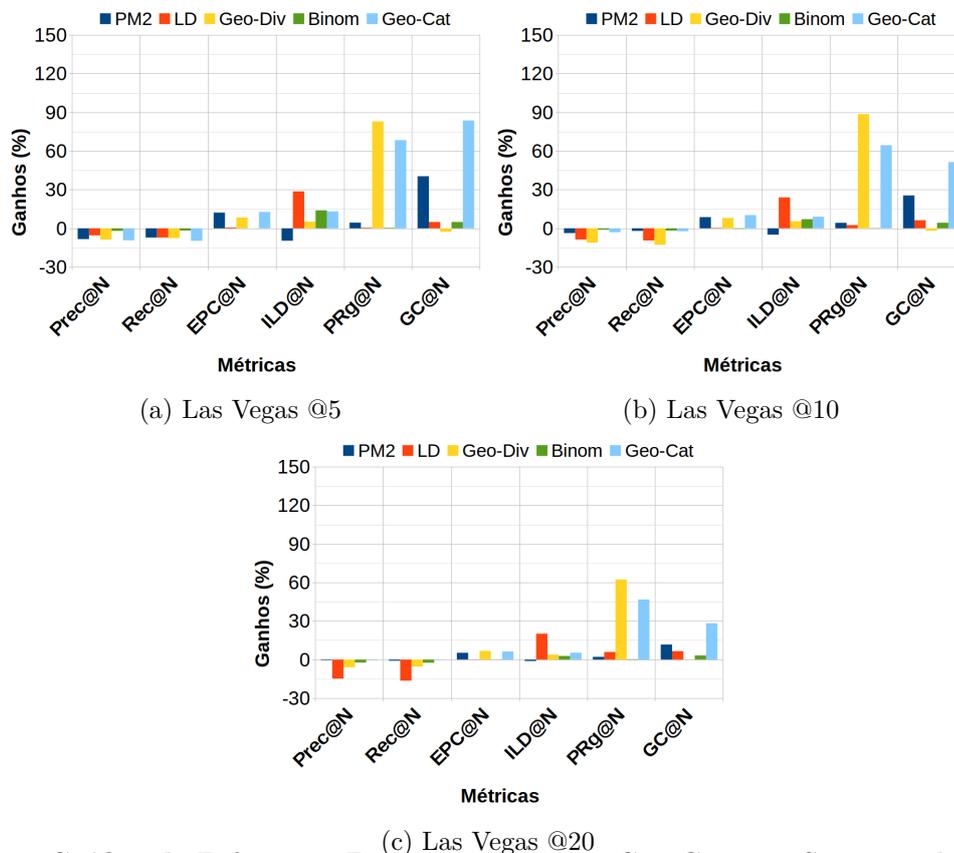


Figura 11 – Gráfico de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage da cidade de Las Vegas.

| Phoenix | | | | | | |
|-----------|----------|----------|----------|----------|----------|----------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 |
| USG | 0,0460 | 0,0182 | 0,7407 | 0,4695 | 0,1537 | 0,3533 |
| PM2 | 0,0402 ▼ | 0,0158 ▼ | 0,7515 ▲ | 0,4113 ▼ | 0,1531 ● | 0,4248 ▲ |
| LD | 0,0441 ● | 0,0176 ● | 0,7486 ● | 0,7438 ▲ | 0,1566 ▲ | 0,3594 ▲ |
| Geo-Div | 0,0353 ▼ | 0,0138 ▼ | 0,8124 ▲ | 0,5437 ▲ | 0,3741 ▲ | 0,3224 ▼ |
| Binom | 0,0456 ● | 0,0182 ● | 0,7453 ● | 0,5938 ▲ | 0,1551 ● | 0,3661 ▲ |
| Geo-Cat | 0,0380 ▼ | 0,0153 ▼ | 0,7989 ▲ | 0,5926 ▲ | 0,3261 ▲ | 0,5491 ▲ |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 |
| USG | 0,0371 | 0,0287 | 0,7495 | 0,4946 | 0,2220 | 0,4836 |
| PM2 | 0,0358 ● | 0,0279 ● | 0,7610 ▲ | 0,4671 ▼ | 0,2228 ● | 0,5573 ▲ |
| LD | 0,0345 ▼ | 0,0267 ▼ | 0,7569 ● | 0,7278 ▲ | 0,2321 ▲ | 0,5029 ▲ |
| Geo-Div | 0,0299 ▼ | 0,0226 ▼ | 0,8173 ▲ | 0,5670 ▲ | 0,5129 ▲ | 0,4653 ▼ |
| Binom | 0,0369 ● | 0,0286 ● | 0,7533 ● | 0,5564 ▲ | 0,2251 ▲ | 0,5034 ▲ |
| Geo-Cat | 0,0356 ● | 0,0278 ● | 0,7927 ▲ | 0,5772 ▲ | 0,4301 ▲ | 0,6671 ▲ |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 |
| USG | 0,0303 | 0,0464 | 0,7628 | 0,5221 | 0,2971 | 0,6149 |
| PM2 | 0,0305 ● | 0,0467 ● | 0,7712 ▲ | 0,5159 ▼ | 0,2997 ▲ | 0,6652 ▲ |
| LD | 0,0272 ▼ | 0,0409 ▼ | 0,7663 ● | 0,7045 ▲ | 0,3148 ▲ | 0,6437 ▲ |
| Geo-Div | 0,0278 ▼ | 0,0430 ▼ | 0,8141 ▲ | 0,5642 ▲ | 0,5403 ▲ | 0,6108 ▼ |
| Binom | 0,0301 ● | 0,0460 ● | 0,7646 ● | 0,5425 ▲ | 0,2989 ● | 0,6336 ▲ |
| Geo-Cat | 0,0305 ● | 0,0465 ● | 0,7916 ▲ | 0,5678 ▲ | 0,4853 ▲ | 0,7524 ▲ |

Tabela 7 – Comparativo entre as *Base Lines* e o Geo-Cat a para cidade de Phoenix. A avaliação estatística foi feita utilizando-se o teste de *Wilcoxon* com 95% de intervalo de confiança. O símbolo ● representa empates estatísticos, ▲ representa ganhos estatísticos e ▼ representa perdas estatísticas.

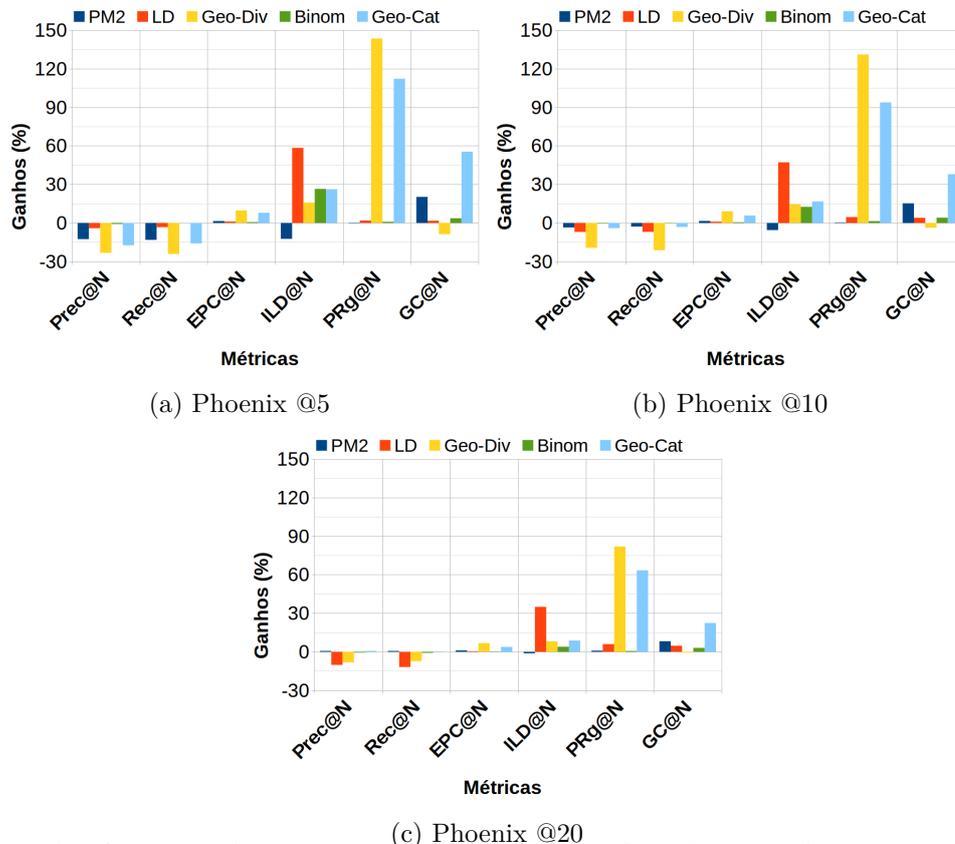


Figura 12 – Gráfico de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage da cidade de Phoenix

Para confirmar as conclusões obtidas no parágrafo anterior, novamente utilizamos a métrica extraída a partir da MAUT, com o objetivo de avaliar cada um dos métodos, considerando simultaneamente as várias métricas analisadas. Assim como na definição do melhor SR base, também consideramos que todas as métricas de avaliação são igualmente importantes e possuem pesos iguais ($w = 0,1667$ na Equação. 5.3). Conforme apresentado na Seção 5.2, o valor da MAUT está entre 0 e 1: quanto mais próximo de 1, melhor é o comportamento de um determinado método frente às múltiplas métricas avaliadas. Os resultados referentes à essa análise estão apresentados nas tabelas 8, 9 e 10 e sumarizados nos gráficos da Figura 13.

| Charlotte | | | | | | | |
|-----------|---------|--------|--------|--------|--------|--------|--------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 | MAUT |
| Geo-Cat | 0,0000 | 0,0000 | 1,0000 | 0,5947 | 0,8251 | 1,0000 | 0,5677 |
| Binom | 0,7843 | 0,8261 | 0,0014 | 0,6158 | 0,0053 | 0,0883 | 0,3853 |
| Geo-Div | 0,0588 | 0,2174 | 0,6572 | 0,3854 | 1,0000 | 0,0000 | 0,3849 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,2511 | 0,0000 | 0,0313 | 0,3789 |
| PM2 | 0,0980 | 0,2609 | 0,9580 | 0,0000 | 0,0530 | 0,4994 | 0,3103 |
| LD | 0,4118 | 0,2609 | 0,0515 | 1,0000 | 0,0063 | 0,0883 | 0,3019 |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 | MAUT |
| Geo-Cat | 0,6136 | 0,7000 | 0,8475 | 0,4624 | 0,7831 | 1,0000 | 0,7315 |
| Binom | 1,0000 | 1,0000 | 0,0000 | 0,4564 | 0,0003 | 0,2592 | 0,4508 |
| PM2 | 0,5909 | 0,5250 | 0,7060 | 0,0000 | 0,0322 | 0,5336 | 0,3964 |
| USG | 0,9545 | 0,8250 | 0,0117 | 0,2880 | 0,0049 | 0,2048 | 0,3800 |
| Geo-Div | 0,0000 | 0,0000 | 1,0000 | 0,2611 | 1,0000 | 0,0000 | 0,3753 |
| LD | 0,4091 | 0,4750 | 0,0811 | 1,0000 | 0,0000 | 0,2665 | 0,3705 |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 | MAUT |
| Geo-Cat | 0,9388 | 0,8553 | 0,7303 | 0,3290 | 0,7749 | 1,0000 | 0,7683 |
| PM2 | 1,0000 | 0,8947 | 0,6228 | 0,0000 | 0,0009 | 0,4638 | 0,4950 |
| Geo-Div | 0,3469 | 0,3421 | 1,0000 | 0,2000 | 1,0000 | 0,0000 | 0,4796 |
| Binom | 0,9796 | 1,0000 | 0,0037 | 0,1971 | 0,0056 | 0,1768 | 0,3922 |
| USG | 0,8980 | 0,8684 | 0,0000 | 0,1015 | 0,0000 | 0,0972 | 0,3262 |
| LD | 0,0000 | 0,0000 | 0,0408 | 1,0000 | 0,0233 | 0,2609 | 0,2199 |

Tabela 8 – Aplicação da MAUT como ferramenta para selecionar o melhor algoritmo de diversificação para a cidade de Charlotte.

Podemos observar novamente que há muitas diferenças entre os métodos avaliados. Nas Tabelas 8, 9 e 10 que apresentam os resultados da MAUT, observa-se que individualmente cada método tende a ser melhor em métricas específicas. Por exemplo, o método LD é o melhor (valor = 1) na coluna ILD. Entretanto, o mesmo método não apresenta bons resultados sempre em outras colunas das tabelas. De forma análoga, o método Geo-Div é sempre o melhor resultado na coluna PRg, porém, em quase todas as situações é o pior (valor = 0) em *Genre Coverage* (GC). Nessas tabelas, deve-se considerar o maior valor obtido na última coluna para definir qual é o melhor método. Por essa análise, é possível observar que o Geo-Cat é aquele que apresenta os melhores resultados quando consideramos todos as métricas simultaneamente.

| Las Vegas | | | | | | | |
|-----------|---------|--------|--------|--------|--------|--------|--------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 | MAUT |
| Geo-Cat | 0,0000 | 0,0000 | 1,0000 | 0,5947 | 0,8251 | 1,0000 | 0,5677 |
| Binom | 0,7843 | 0,8261 | 0,0014 | 0,6158 | 0,0053 | 0,0883 | 0,3853 |
| Geo-Div | 0,0588 | 0,2174 | 0,6572 | 0,3854 | 1,0000 | 0,0000 | 0,3849 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,2511 | 0,0000 | 0,0313 | 0,3789 |
| PM2 | 0,0980 | 0,2609 | 0,9580 | 0,0000 | 0,0530 | 0,4994 | 0,3103 |
| LD | 0,4118 | 0,2609 | 0,0515 | 1,0000 | 0,0063 | 0,0883 | 0,3019 |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 | MAUT |
| Geo-Cat | 0,7255 | 0,8200 | 1,0000 | 0,4809 | 0,7273 | 1,0000 | 0,7891 |
| PM2 | 0,6667 | 0,8400 | 0,8545 | 0,0000 | 0,0479 | 0,5141 | 0,4853 |
| Binom | 0,9020 | 0,8600 | 0,0000 | 0,4154 | 0,0021 | 0,1159 | 0,3810 |
| USG | 1,0000 | 1,0000 | 0,0391 | 0,1708 | 0,0000 | 0,0355 | 0,3727 |
| Geo-Div | 0,0000 | 0,0000 | 0,7950 | 0,3605 | 1,0000 | 0,0000 | 0,3578 |
| LD | 0,2157 | 0,2600 | 0,0751 | 1,0000 | 0,0285 | 0,1522 | 0,2874 |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 | MAUT |
| Geo-Cat | 1,0000 | 0,9811 | 0,9252 | 0,3040 | 0,7505 | 1,0000 | 0,8235 |
| Geo-Div | 0,5862 | 0,6698 | 1,0000 | 0,2327 | 1,0000 | 0,0000 | 0,5791 |
| PM2 | 0,9483 | 0,9434 | 0,7827 | 0,0000 | 0,0337 | 0,4176 | 0,5189 |
| USG | 0,9828 | 1,0000 | 0,0164 | 0,0543 | 0,0000 | 0,0061 | 0,3419 |
| Binom | 0,8276 | 0,8491 | 0,0023 | 0,1850 | 0,0048 | 0,1184 | 0,3299 |
| LD | 0,0000 | 0,0000 | 0,0000 | 1,0000 | 0,0948 | 0,2357 | 0,2209 |

Tabela 9 – Aplicação da MAUT como ferramenta para selecionar o melhor algoritmo de diversificação para a cidade de Las Vegas.

| Phoenix | | | | | | | |
|-----------|---------|--------|--------|--------|--------|--------|--------|
| Algoritmo | Prec@5 | Rec@5 | EPC@5 | ILD@5 | PRg@5 | GC@5 | MAUT |
| Geo-Cat | 0,2523 | 0,3409 | 0,8117 | 0,5453 | 0,7828 | 1,0000 | 0,6197 |
| LD | 0,8224 | 0,8636 | 0,1102 | 1,0000 | 0,0158 | 0,1632 | 0,4939 |
| Binom | 0,9626 | 1,0000 | 0,0642 | 0,5489 | 0,0090 | 0,1928 | 0,4611 |
| Geo-Div | 0,0000 | 0,0000 | 1,0000 | 0,3982 | 1,0000 | 0,0000 | 0,3981 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,1750 | 0,0027 | 0,1363 | 0,3841 |
| PM2 | 0,4579 | 0,4545 | 0,1506 | 0,0000 | 0,0000 | 0,4517 | 0,2515 |
| Algoritmo | Prec@10 | Rec@10 | EPC@10 | ILD@10 | PRg@10 | GC@10 | MAUT |
| Geo-Cat | 0,7917 | 0,8525 | 0,6372 | 0,4223 | 0,7154 | 1,0000 | 0,7336 |
| LD | 0,6389 | 0,6721 | 0,1091 | 1,0000 | 0,0347 | 0,1863 | 0,4384 |
| Binom | 0,9722 | 0,9836 | 0,0560 | 0,3425 | 0,0107 | 0,1888 | 0,4239 |
| Geo-Div | 0,0000 | 0,0000 | 1,0000 | 0,3832 | 1,0000 | 0,0000 | 0,3956 |
| PM2 | 0,8194 | 0,8689 | 0,1696 | 0,0000 | 0,0028 | 0,4559 | 0,3845 |
| USG | 1,0000 | 1,0000 | 0,0000 | 0,1055 | 0,0000 | 0,0907 | 0,3646 |
| Algoritmo | Prec@20 | Rec@20 | EPC@20 | ILD@20 | PRg@20 | GC@20 | MAUT |
| Geo-Cat | 1,0000 | 0,9655 | 0,5614 | 0,2752 | 0,7738 | 1,0000 | 0,7596 |
| Geo-Div | 0,1818 | 0,3621 | 1,0000 | 0,2561 | 1,0000 | 0,0000 | 0,4648 |
| PM2 | 1,0000 | 1,0000 | 0,1637 | 0,0000 | 0,0107 | 0,3842 | 0,4247 |
| Binom | 0,8788 | 0,8793 | 0,0351 | 0,1410 | 0,0074 | 0,1610 | 0,3490 |
| USG | 0,9394 | 0,9483 | 0,0000 | 0,0329 | 0,0000 | 0,0290 | 0,3236 |
| LD | 0,0000 | 0,0000 | 0,0682 | 1,0000 | 0,0728 | 0,2323 | 0,2280 |

Tabela 10 – Aplicação da MAUT como ferramenta para selecionar o melhor algoritmo de diversificação para a cidade de Phoenix.

Dessa forma, com base nos resultados apresentados nessa seção, bem como naqueles apresentados no final da Seção 5.3.2, quando comparamos os diversos SsR específicos para POIs, consolidamos a segunda hipótese deste trabalho: *(H2): É possível avaliar SsR considerando múltiplas perspectivas de qualidade, tais como acurácia, diversidade categórica e diversidade geográfica.* Trata-se da segunda contribuição do nosso trabalho, uma vez que por meio dessa adaptação da MAUT conseguimos avaliar a qualidade dos recomendadores, considerando diversas perspectivas simultaneamente. Apenas para ilustrar

a importância desta contribuição, no artigo em que o GeoMF (66) é apresentado, os autores o comparam ao USG (51) e, com base em apenas duas métricas de avaliação relacionadas à efetividade de recomendadores, argumentam que trata-se de uma proposta melhor que o USG. Entretanto, ao avaliarmos sob diferentes dimensões de qualidade simultaneamente pela MAUT, vimos que o USG é bem melhor que o GeoMF.

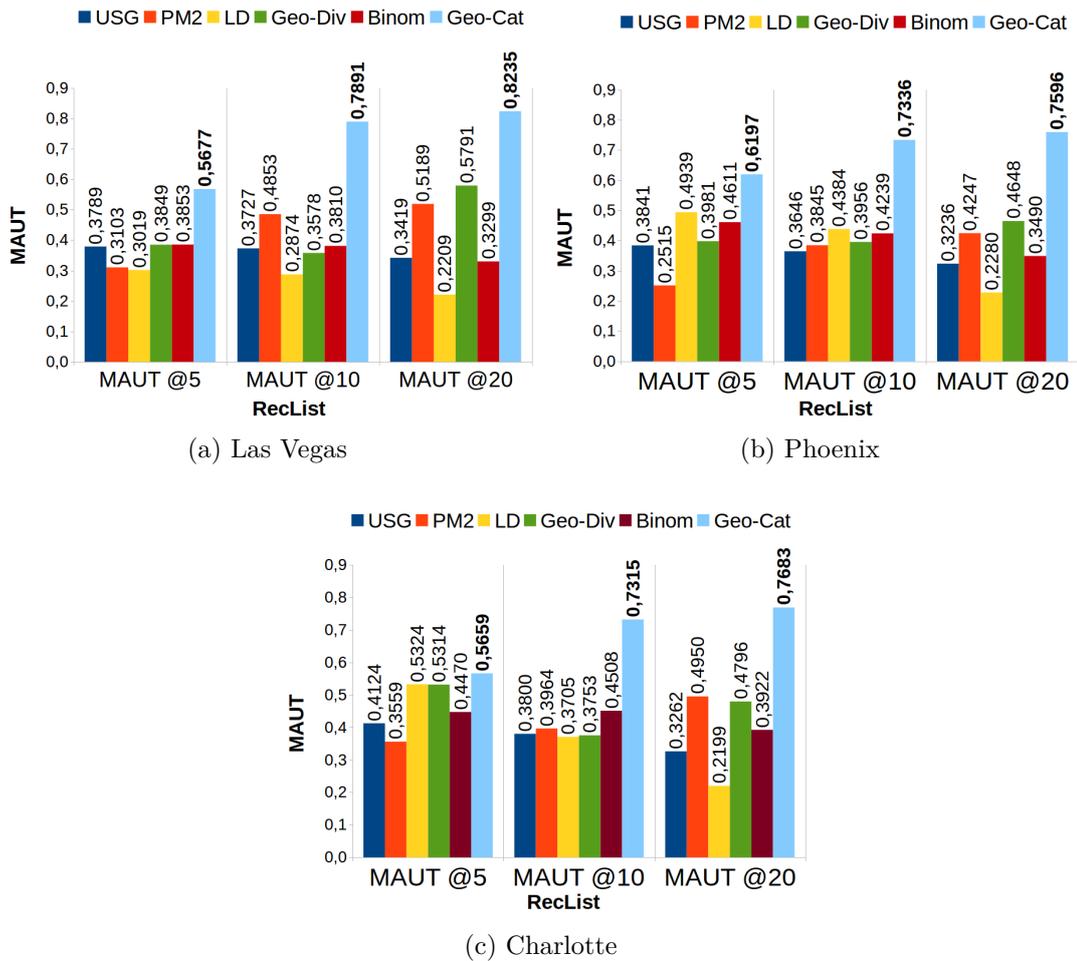


Figura 13 – Gráficos de Diferenças Percentuais entre o Geo-Cat e os Sistemas de Recomendação base nas Métricas Precision, Recall, EPC, ILD, PRg e Genre Coverage

6 Conclusão & Trabalhos Futuros

Este trabalho se propôs a analisar os aspectos de diversidades categórica e geográfica em SsR de POIs, a fim de propor e avaliar novas estratégias de diversificação. A principal justificativa para a elaboração deste trabalho está relacionada à importância de que os SsR de POIs devam se concentrar não apenas na efetividade da tarefa de correspondência entre POIs e usuários, mas também na satisfação dos usuários (7, 9, 10, 11). A efetividade está relacionada à relevância dos POIs recomendados, enquanto que a satisfação está relacionada à diversidade desses POIs (geográfica e categoricamente).

A combinação da efetividade com a diversidade contribui potencialmente para impulsionar a qualidade do SR. Entretanto, efetividade e diversidade representam conceitos opostos: os POIs que tendem a aumentar a efetividade dos SsR são aqueles mais populares entre os usuários, enquanto que para se prover uma maior diversidade, deve-se recomendar POIs menos convencionais. A partir disso, derivamos a primeira hipótese deste trabalho: *(H1): SsR de POIs devem explorar as diversidades categórica e geográfica, simultaneamente, sem perdas relevantes de efetividade..* Para abordar essa hipótese, propomos e avaliamos uma nova estratégia de pós-processamento, a ser aplicada em conjunto com SsR, a fim de prover diversidade em SsR de POIs sem perdas significativas de efetividade. Denominamos essa estratégia de *Geo-Cat*.

Outra questão importante levantada nesse trabalho é que o processo de avaliação de SsR ainda é um dos grandes desafios da área. Particularmente, não existe na literatura um consenso sobre quais métricas de avaliação são necessárias para consolidar um novo SR (23, 24, 25, 26, 16, 27, 28). Dessa maneira, pesquisadores se limitam, na maioria das vezes, a avaliar métricas focadas nas técnicas propostas. Por exemplo, trabalhos que propõem SsR focados em efetividade não avaliam a diversidade e vice-versa.

Dessa forma, com base nessas informações e também com base em nossa primeira hipótese, derivamos a segunda hipótese deste trabalho: *(H2): A utilidade dos SsR de POIs deve ser avaliada considerando múltiplas perspectivas de qualidade, tais como acurácia, diversidade categórica e diversidade geográfica..* Para abordar essa hipótese, adaptamos a Teoria da Utilidade Multiatributo MAUT, utilizada no domínio de Teoria dos Jogos (29) para manipular problemas em que os resultados são dependentes de dois ou mais atributos, materializando-a como uma métrica capaz de avaliar a eficácia de um método em função de diferentes dimensões de qualidade.

Para alcançar o objetivo principal do trabalho, é preciso contextualizar o cenário de recomendações ao qual este trabalho é aplicado. Assim, no Capítulo 2 apresentamos um referencial teórico, iniciando pela conceituação de SsR, apresentando as estratégias básicas,

bem como a divisão das mesmas em diferentes classes. Além disso, abordamos nesse capítulo o problema específico de recomendação de POIs, destacando as principais fontes de informação utilizadas por SsR de POIs: influência social, influência temporal e influência geográfica. Dessas informações, a influência geográfica assume destaque como o fator mais importante, por distinguir a recomendação de POIs das recomendações tradicionais (i.e., filmes, músicas). Finalizamos o referencial teórico apresentando as principais métricas utilizadas para avaliar a qualidade de SsR. Para melhorar a compreensão, a apresentação das métricas foi dividida em blocos: métricas de efetividade, tais como acurácia, precisão, revocação etc. e dimensões complementares de qualidade, tais como novidade, serendipidade, diversidade categórica e, para o caso especial de recomendação de POIs, diversidade geográfica.

De forma complementar ao Capítulo 2, no Capítulo 3 apresentamos os principais SsR de POIs da literatura. Nossa apresentação foi dividida de acordo com as informações utilizadas por esses SsR, seja a influência social, a influência temporal, a influência geográfica ou mesmo a combinação dessas informações. Em nossas análises, observamos que as estratégias que apresentam os melhores resultados são aquelas que utilizam a influência geográfica (sozinha ou combinada) com a influência social. Os destaques dessa análise foram os algoritmos GeoMF (66), GeoSoCa (71) e USG (51), e que foram, portanto, escolhidos na avaliação de nossa proposta. Além disso, nesse capítulo são apresentados trabalhos focados em ampliar a diversidade das listas de recomendações para aumentar a qualidade das recomendações. De maneira geral, a diversificação em recomendação de POIs pode ser alcançada de duas maneiras: categórica e/ou geográfica. Dessa forma, são apresentados métodos de diversificação tradicionais que focam na diversificação categórica dos itens. Essa diversificação pode ainda ser realizada com base nas categorias dos itens recomendados (i.e., *Binomial* (18)) ou com base no conteúdo dos itens (i.e., LD (16)). Em outras palavras, as recomendações podem ser diversificadas, sugerindo itens que pertençam a diferentes categorias e/ou recomendando itens que sejam diferentes entre si. Além disso, em recomendação de POIs os itens recomendados devem ser diversificados geograficamente. Nesse sentido, apresentamos o único trabalho até então que lida com o problema da diversificação geográfica (9). Esse trabalho propõe uma nova forma de distribuir os POIs recomendados geograficamente de maneira proporcional e também uma adaptação da métrica de proporcionalidade (PR) para avaliar essa distribuição. Contudo, o método Geo-Cat aborda os conceitos de diversificação categórica e geográfica simultaneamente e oferece listas de recomendações mais úteis para os usuários.

O estudo realizado e apresentado no Capítulo 3 embasou a proposta do método de diversificação de POIs *Geo-Cat* detalhado no Capítulo 4 desta dissertação. Especificamente, o *Geo-Cat* é um método de pós-processamento capaz de trabalhar sobre as recomendações oriundas de um SR base e, utilizando conceitos da MMR (20, 21, 22), maximizar a diversidade categórica e geográfica, mantendo a relevância dos POIs. Nesse sentido, a relevância dos POIs é mensurada pelo *score* atribuído pelo SR base. Por sua vez, a diversidade é a combinação linear de dois fatores: diversidade geográfica e diversidade categórica. Em suma, o

Geo-Cat faz a reordenação da lista de recomendações baseado no objetivo de maximizar as diversidades categórica e geográfica. Dessa maneira, a diversidade categórica é alcançada de duas maneiras: (1) aumentando a quantidade de categorias entre os POIs recomendados (i.e., restaurantes, museus); e (2) aumentando a diferença entre si dos POIs recomendados. Por outro lado, a diversificação geográfica é alcançada considerando que os POIs recomendados estejam localizados, proporcionalmente, em áreas de atividade do usuário. Dessa forma, esse capítulo abordou parcialmente a **HI**, apresentando a primeira contribuição deste trabalho.

Para avaliar nossa proposta de estratégia de diversificação geográfica categórica e, conseqüentemente, consolidar a primeira hipótese de nosso trabalho, no Capítulo 5 apresentamos uma vasta avaliação experimental. Primeiramente apresentamos nossa configuração experimental, detalhando as bases de dados, recomendadores base, estratégias de diversificação utilizadas como linhas de base e métricas de avaliação consideradas. Basicamente, como base de dados utilizamos dados referentes às cidades de *Charlotte*, *Las Vegas* e *Phoenix* nos Estados Unidos disponibilizadas por meio da *Yelp Challenge*. Realizamos uma caracterização dessas coleções, demonstrando o quão variadas elas são no que diz respeito a número de POIs, total de usuários e de *check-ins*, bem como a relação dessas características. Além disso, apresentamos como os POIs dessas cidades estão divididos entre diferentes categorias da própria *Yelp*. Os recomendadores básicos considerados foram: *most popular*, *user-knn*, *item attribute-knn* e *WRMF* como SsR tradicionais e os SsR específicos de POIs (*GeoSoCa*, *GeoMF* e *USG*). As métricas avaliadas foram: *Precision*, *Recall*, *EPC*, *ILD*, *PR_g* e *Genre Coverage*. Ainda nesse capítulo, apresentamos nossa segunda contribuição: uma adaptação da MAUT (29), para o cenário de recomendação de POIs. Definido o arcabouço e processo de avaliação, apresentamos os resultados divididos em dois estudos comparativos. Primeiramente comparamos os recomendadores bases com e sem o uso do *Geo-Cat*, demonstrando que a utilização de nossa estratégia fez com que a diversidade dos recomendadores básicos fossem muito melhores, sem apresentar perdas significativas em termos de efetividade, consolidando nossa primeira hipótese.

Em um segundo momento, comparamos o desempenho de nossa estratégia de diversificação frente a outras estratégias existentes na literatura: **PM2** (17), **LD** (18), **Binom** (18) e **Geo-Div** (9). Para tanto, foi necessário escolher o recomendador base que seria utilizado nessa etapa da avaliação. Assim, comparamos os SsR específicos para POIs frente as múltiplas dimensões de qualidade utilizando a MAUT, quando esta apontou para o *USG* como aquele que apresenta os melhores resultados, sendo, portanto, o escolhido para a segunda etapa de avaliação. Esses estudos sobre diversidade apontam que individualmente cada método tende a ser melhor em algumas métricas específicas. Por exemplo, o método LD apresenta bons resultados sob a métrica *ILD*; entretanto, os resultados alcançados em *PR_g* são ruins. Em suma, a aplicação da MAUT sobre os métodos de diversificação aponta que o *Geo-Cat* é o método de recomendação mais equilibrado e a melhor opção entre todos os outros. Sobretudo, é possível observar que o *Geo-Cat* apresenta os melhores resultados

ao avaliarmos todas as métricas simultaneamente. Em especial, a proposta de adaptação da *MAUT* representa um avanço significativo nas avaliações de SsR, uma vez que eles podem ser avaliados sob múltiplas perspectivas, evitando avaliações enviesadas por alguma métrica específica, consolidando nossa segunda hipótese de trabalho. Além disso, toda essa vasta avaliação experimental constitui-se na terceira contribuição dessa dissertação de mestrado.

Este trabalho não se propôs a esgotar todas as possibilidades de análise dos aspectos de diversidade que tangem os cenários de recomendações de POIs. Dessa maneira, sugerimos alguns estudos futuros que poderiam complementar a análise feita aqui, de modo a ratificar ou retificar as conclusões obtidas. Por exemplo, um estudo sobre o padrão de comportamento de usuários poderia ser realizado para tentar criar um SsR que gere diversificação personalizada. Em outras palavras, sugerimos que é possível equilibrar as diversificações geográfica e categórica de acordo com o perfil dos usuários. Ademais, um estudo mais detalhado das categorias de POIs pode ser realizado a fim de identificar correlação entre a satisfação dos usuários e a diversidade de categorias. Outros trabalhos futuros poderão considerar algoritmos mais elaborados para otimizar a reordenação da lista de POIs de um SR base, tais como meta-heurísticas bioinspiradas. Além disso, diferentes divisões e formatos de sub-regiões podem ser analisados a fim de ampliar ainda mais o estudo sobre diversidade geográfica dos POIs. Também seria interessante realizar estudos que apresentem a comparação entre os resultados dos SsR focados em diversidade com os resultados apresentados por SsR Multi-Objetivos. Um outro estudo interessante poderia propor estratégias de pré-processamento dos dados para maximizar a efetividade dos SsR base. Além disso, trabalhos futuros podem analisar como é o comportamento dos algoritmos de diversificação para diferentes configurações de RBs. Especificamente, métodos para tornar as recomendações mais pessoais, utilizando influências temporais e analisando melhor o comportamento de locomoção dos usuários, podem atingir esse objetivo. Esses novos trabalhos poderiam ser então utilizados em conjunto com o *Geo-Cat* para criar recomendações mais relevantes e mais diversificadas simultaneamente.

É igualmente importante destacar que, este trabalho apresenta algumas limitações de escopo, tais como: (1) Os experimentos realizados não foram feitos com turistas, ou seja, não fizemos testes de recomendação de POIs de uma cidade para usuários sem considerar seus históricos de *check-ins* na própria cidade; (2) Usuários com poucos *check-ins* foram eliminados da base de dados no pré-processamento, dessa forma, não foram analisados os comportamentos dos SsR para problemas de *cold-start* ou usuários com pouco histórico de informações; e (3) Não há informação na base de dados sobre o momento em que a recomendação é feita; dessa forma, não é possível avaliar se a geolocalização do usuário quando a recomendação é recebida afeta os resultados de satisfação e utilidade dos POIs recomendados.

Dessa forma, o problema da recomendação de POIs é ainda uma área com muitos aspectos que precisam ser melhor explorados. Contudo, o método de diversificação proposto

neste trabalho amplia as possibilidades de criar listas de recomendação com mais diversidades (geográfica e categórica), sem abrir mão da relevância dos POIs recomendados. Além disso, a adaptação da MAUT proposta de forma inédita para o cenário de recomendações possibilita avaliar os SsR sob diversos aspectos simultaneamente. Dessa forma, torna possível que novos métodos de recomendação propostos sejam avaliados de forma mais justa, independente do foco do SR.

Referências

- 1 COFFMAN, K. G.; ODLYZKO, A. M. Internet growth: is there a "moore's law" for data traffic? Kluwer Academic Publishers, Norwell, MA, USA, p. 47–93, 2002. Citado na página 21.
- 2 MAYER-SCHONBERGER, V.; CUKIER, K. *Big Data: Como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana*. Elsevier Brasil, 2014. ISBN 9788535273410. Disponível em: <<https://books.google.com.br/books?id=kJWoBQAAQBAJ>>. Citado na página 21.
- 3 HU, L.; SUN, A.; LIU, Y. Your neighbors affect your ratings: On geographical neighborhood influence to rating prediction. In: *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. New York, NY, USA: ACM, 2014. (SIGIR '14), p. 345–354. ISBN 978-1-4503-2257-7. Disponível em: <<http://doi.acm.org/10.1145/2600428.2609593>>. Citado 4 vezes nas páginas 21, 22, 32 e 44.
- 4 RICCI, F. et al. (Ed.). *Recommender Systems Handbook*. [S.l.]: Springer US, 2011. Citado 8 vezes nas páginas 21, 22, 27, 28, 29, 31, 36 e 59.
- 5 MOURAO, F. H. de J. *A hybrid recommendation method that combines forgotten items and non-content attributes*. Tese (Doutorado) — Universidade Federal de Minas Gerais, 2014. Citado 5 vezes nas páginas 22, 27, 29, 30 e 31.
- 6 S., S. K.; T., R.; S., P. Product recommendations over facebook: The roles of influencing factors to induce online shopping. *Asian Social Science*, Canadian Center of Science and Education, v. 11, n. 2, dec 2014. Citado na página 22.
- 7 GRIESNER, J.-B.; ABDESSALEM, T.; NAACKE, H. Poi recommendation: towards fused matrix factorization with geographical and temporal influences. In: ACM. *Proceedings of the 9th ACM Conference on Recommender Systems*. [S.l.], 2015. p. 301–304. Citado 3 vezes nas páginas 22, 23 e 74.
- 8 WANG, S. S.; STEFANONE, M. A. Showing off? human mobility and the interplay of traits, self-disclosure, and facebook check-ins. *Social Science Computer Review*, Sage Publications Sage CA: Los Angeles, CA, v. 31, n. 4, p. 437–457, 2013. Citado na página 22.
- 9 HAN, J.; YAMANA, H. Geographical diversification in POI recommendation. In: *Proceedings of the Eleventh ACM Conference on Recommender Systems - RecSys '17*. [S.l.]: ACM Press, 2017. Citado 17 vezes nas páginas 23, 25, 37, 40, 44, 45, 46, 47, 50, 53, 55, 57, 58, 59, 74, 75 e 76.
- 10 CELMA, Ò.; HERRERA, P. A new approach to evaluating novel recommendations. In: ACM. *Proceedings of the 2008 ACM conference on Recommender systems*. [S.l.], 2008. p. 179–186. Citado 2 vezes nas páginas 23 e 74.

- 11 SHANI, G.; GUNAWARDANA, A. Evaluating recommendation systems. *Recommender systems handbook*, Springer, p. 257–297, 2011. Citado 2 vezes nas páginas [23](#) e [74](#).
- 12 MCNEE, S. M.; RIEDL, J.; KONSTAN, J. A. Being accurate is not enough: How accuracy metrics have hurt recommender systems. In: *CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06*. [S.l.]: ACM Press, 2006. Citado 2 vezes nas páginas [23](#) e [35](#).
- 13 CHO, E.; MYERS, S. A.; LESKOVEC, J. Friendship and mobility: user movement in location-based social networks. In: ACM. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. [S.l.], 2011. Citado na página [23](#).
- 14 GE, M.; DELGADO-BATTENFELD, C.; JANNACH, D. Beyond accuracy: Evaluating recommender systems by coverage and serendipity. In: *Proceedings of the fourth ACM conference on Recommender systems - RecSys '10*. [S.l.]: ACM Press, 2010. Citado 2 vezes nas páginas [23](#) e [41](#).
- 15 KUNAVER, M.; POŽRL, T. Diversity in recommender systems—a survey. *Knowledge-Based Systems*, Elsevier, v. 123, p. 154–162, 2017. Citado 2 vezes nas páginas [23](#) e [46](#).
- 16 VARGAS, S.; CASTELLS, P. Rank and relevance in novelty and diversity metrics for recommender systems. In: ACM. *Proceedings of the fifth ACM conference on Recommender systems*. [S.l.], 2011. p. 109–116. Citado 11 vezes nas páginas [23](#), [24](#), [25](#), [36](#), [42](#), [45](#), [47](#), [48](#), [58](#), [74](#) e [75](#).
- 17 DANG, V.; CROFT, W. B. Diversity by proportionality: an election-based approach to search result diversification. In: ACM. *Proceedings of the international ACM SIGIR conference on Research and development in information retrieval*. [S.l.], 2012. Citado 5 vezes nas páginas [23](#), [37](#), [58](#), [59](#) e [76](#).
- 18 VARGAS, S. et al. Coverage, redundancy and size-awareness in genre diversity for recommender systems. In: ACM. *Proceedings of the 8th ACM Conference on Recommender systems*. [S.l.], 2014. p. 209–216. Citado 5 vezes nas páginas [23](#), [43](#), [58](#), [75](#) e [76](#).
- 19 ZHOU, T. et al. Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 107, n. 10, p. 4511–4515, 2010. Citado na página [23](#).
- 20 CARBONELL, J. G.; GOLDSTEIN, J. The use of mmr and diversity-based reranking for reordering documents and producing summaries. 1998. Citado 4 vezes nas páginas [24](#), [42](#), [44](#) e [75](#).
- 21 ZHANG, M.; HURLEY, N. Avoiding monotony: improving the diversity of recommendation lists. In: ACM. *Proceedings of the 2008 ACM conference on Recommender systems*. [S.l.], 2008. p. 123–130. Citado 3 vezes nas páginas [24](#), [42](#) e [75](#).
- 22 SANTOS, R. L.; MACDONALD, C.; OUNIS, I. Exploiting query reformulations for web search result diversification. In: ACM. *Proceedings of the 19th international conference on World wide web*. [S.l.], 2010. p. 881–890. Citado 3 vezes nas páginas [24](#), [42](#) e [75](#).

- 23 HERLOCKER, J. L. et al. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, Association for Computing Machinery (ACM), v. 22, n. 1, p. 5–53, jan 2004. Citado 5 vezes nas páginas 24, 33, 34, 35 e 74.
- 24 BOBADILLA, J. et al. Recommender systems survey. *Knowledge-based systems*, Elsevier, v. 46, p. 109–132, 2013. Citado 8 vezes nas páginas 24, 28, 30, 33, 34, 46, 59 e 74.
- 25 JANNACH, D. et al. *Recommender systems: an introduction*. [S.l.]: Cambridge University Press, 2010. Citado 4 vezes nas páginas 24, 28, 35 e 74.
- 26 BARJASTEH, I. et al. Cold-start item and user recommendation with decoupled completion and transduction. In: ACM. *Proceedings of the 9th ACM Conference on Recommender Systems*. [S.l.], 2015. p. 91–98. Citado 4 vezes nas páginas 24, 35, 40 e 74.
- 27 ZHANG, Y. C. et al. Auralist: introducing serendipity into music recommendation. In: ACM. *Proceedings of the fifth ACM international conference on Web search and data mining*. [S.l.], 2012. p. 13–22. Citado 3 vezes nas páginas 24, 36 e 74.
- 28 PARAMBATH, S. A. P.; USUNIER, N.; GRANDVALET, Y. A coverage-based approach to recommendation diversity on similarity graph. In: *Proceedings of the 10th ACM Conference on Recommender Systems - RecSys '16*. [S.l.]: ACM Press, 2016. Citado 6 vezes nas páginas 24, 36, 47, 48, 59 e 74.
- 29 NEIVA, S. B.; GOMES, L. F. A. M. A aplicação da teoria da utilidade multiatributo à escolha de um software de e-procurement. *Revista Tecnologia*, v. 28, n. 2, 2009. Citado 5 vezes nas páginas 24, 25, 59, 74 e 76.
- 30 RUSSELL, S. *Inteligência Artificial (Tradução da Segunda Edição)*. [S.l.]: Elsevier, 2004. ISBN 85-352-1177-2. Citado 3 vezes nas páginas 24, 59 e 60.
- 31 ISHIZAKA, A.; NEMERY, P. *Multi-criteria Decision Analysis: Methods and Software*. Wiley, 2013. ISBN 978-1-119-97407-9. Disponível em: <<https://www.amazon.com/Multi-criteria-Decision-Analysis-Methods-Software/dp/1119974070?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimboli05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=1119974070>>. Citado 3 vezes nas páginas 24, 25 e 59.
- 32 AGGARWAL, C. C. *Recommender Systems*. Springer-Verlag GmbH, 2016. ISBN 3319296574. Disponível em: <https://www.ebook.de/de/product/25601300/charu_c_aggarwal_recommender_systems.html>. Citado 6 vezes nas páginas 27, 28, 29, 30, 31 e 57.
- 33 ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, Institute of Electrical and Electronics Engineers (IEEE), v. 17, n. 6, p. 734–749, jun 2005. Citado 8 vezes nas páginas 27, 29, 30, 31, 32, 33, 34 e 35.
- 34 SILVA, N. de C. *Sistemas de Recomendação Não-Personalizados para Atrair Usuários Novos*. Dissertação (Mestrado) — Universidade Federal de Minas Gerais, 2018. Citado 3 vezes nas páginas 27, 31 e 32.

- 35 CANDILLIER, L. et al. State-of-the-art recommender systems. collaborative and social information retrieval and access. 2009. Citado na página 28.
- 36 SILVA, N. et al. Evaluating different strategies to mitigate the ramp-up problem in recommendation domains. In: *Proceedings of the 23rd Brazillian Symposium on Multimedia and the Web - WebMedia '17*. [S.l.]: ACM Press, 2017. Citado 2 vezes nas páginas 28 e 29.
- 37 AKSHITA, S.; SMITA, A. Recommender system: review. *International Journal of Computer Applications*, v. 71, n. 24, p. 38–42, 2013. Citado 4 vezes nas páginas 28, 29, 30 e 31.
- 38 KAWAI, M.; NOGAMI, S. A hybrid recommender system of collaborative and content based filtering. *International Information Institute (Tokyo). Information*, International Information Institute, v. 19, n. 6B, p. 2177, 2016. Citado na página 29.
- 39 METEREN, R. V.; SOMEREN, M. V. Using content-based filtering for recommendation. In: *Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop*. [S.l.: s.n.], 2000. p. 47–56. Citado 2 vezes nas páginas 29 e 30.
- 40 YANG, X. et al. A survey of collaborative filtering based social recommender systems. *Computer Communications*, Elsevier BV, v. 41, p. 1–10, mar 2014. Citado na página 30.
- 41 BELL, R. M.; KOREN, Y. Scalable collaborative filtering with jointly derived neighborhood interpolation weights. In: *Seventh IEEE International Conference on Data Mining (ICDM 2007)*. [S.l.]: IEEE, 2007. Citado na página 31.
- 42 KIM, H.-N. et al. Collaborative user modeling with user-generated tags for social recommender systems. *Expert Systems with Applications*, Elsevier BV, v. 38, n. 7, p. 8488–8496, jul 2011. Citado na página 31.
- 43 BILLSUS, D.; PAZZANI, M. J. User modeling for adaptive news access. *User modeling and user-adapted interaction*, Springer, v. 10, n. 2-3, p. 147–180, 2000. Citado na página 31.
- 44 MELVILLE, P.; MOONEY, R. J.; NAGARAJAN, R. Content-boosted collaborative filtering for improved recommendations. *Aaai/iaai*, v. 23, p. 187–192, 2002. Citado na página 32.
- 45 MOONEY, R. J.; ROY, L. Content-based book recommending using learning for text categorization. In: ACM. *Proceedings of the fifth ACM conference on Digital libraries*. [S.l.], 2000. p. 195–204. Citado na página 32.
- 46 CAMPOS, L. M. de et al. Combining content-based and collaborative recommendations: A hybrid approach based on bayesian networks. *International Journal of Approximate Reasoning*, Elsevier BV, v. 51, n. 7, p. 785–799, sep 2010. Citado na página 32.
- 47 CHOI, K. et al. A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis. *Electronic Commerce Research and Applications*, Elsevier BV, v. 11, n. 4, p. 309–317, jul 2012. Citado na página 32.

- 48 ZHAO, S.; KING, I.; LYU, M. R. A survey of point-of-interest recommendation in location-based social networks. *arXiv preprint arXiv:1607.00647*, 2016. Citado 2 vezes nas páginas 32 e 33.
- 49 YU, Y.; CHEN, X. A survey of point-of-interest recommendation in location-based social networks. In: *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2015. Citado 4 vezes nas páginas 32, 33, 38 e 44.
- 50 HUANG, Z.; CHEN, H.; ZENG, D. Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Trans. Inf. Syst.*, ACM, New York, NY, USA, v. 22, n. 1, p. 116–142, jan. 2004. ISSN 1046-8188. Disponível em: <<http://doi.acm.org/10.1145/963770.963775>>. Citado na página 32.
- 51 YE, M. et al. Exploiting geographical influence for collaborative point-of-interest recommendation. In: ACM. *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. [S.l.], 2011. p. 325–334. Citado 8 vezes nas páginas 32, 40, 44, 46, 53, 57, 73 e 75.
- 52 ZHAO, S.; LYU, M. R.; KING, I. *Point-of-Interest Recommendation in Location-Based Social Networks*. [S.l.]: Springer Singapore, 2018. Citado na página 32.
- 53 YE, M.; YIN, P.; LEE, W.-C. Location recommendation for location-based social networks. In: ACM. *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*. [S.l.], 2010. p. 458–461. Citado 3 vezes nas páginas 32, 39 e 44.
- 54 FERNÁNDEZ-TOBIAS, I. et al. Accuracy and diversity in cross-domain recommendations for cold-start users with positive-only feedback. In: *Proceedings of the 10th ACM Conference on Recommender Systems - RecSys '16*. [S.l.]: ACM Press, 2016. Citado 2 vezes nas páginas 32 e 40.
- 55 YUAN, Q. et al. Time-aware point-of-interest recommendation. In: *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '13*. [S.l.]: ACM Press, 2013. Citado 2 vezes nas páginas 33 e 40.
- 56 LIU, B. et al. Learning geographical preferences for point-of-interest recommendation. In: ACM. *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. [S.l.], 2013. p. 1043–1051. Citado 2 vezes nas páginas 33 e 44.
- 57 WU, W.; HE, L.; YANG, J. Evaluating recommender systems. In: IEEE. *Digital information management (icdim), 2012 seventh international conference on*. [S.l.], 2012. p. 56–61. Citado 2 vezes nas páginas 35 e 41.
- 58 RICCI, F.; ROKACH, L.; SHAPIRA, B. *Introduction to recommender systems handbook*. [S.l.]: Springer, 2011. Citado 2 vezes nas páginas 35 e 58.
- 59 DING, Y.; LI, X. Time weight collaborative filtering. In: *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*. New York, NY, USA: ACM, 2005. (CIKM '05), p. 485–492. ISBN 1-59593-140-6. Disponível em: <<http://doi.acm.org/10.1145/1099554.1099689>>. Citado na página 38.
- 60 GAO, H. et al. Exploring temporal effects for location recommendation on location-based social networks. In: *Proceedings of the 7th ACM conference on Recommender systems - RecSys '13*. [S.l.]: ACM Press, 2013. Citado na página 38.

- 61 WANG, H.; TERROVITIS, M.; MAMOULIS, N. Location recommendation in location-based social networks using user check-in data. In: *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. New York, NY, USA: ACM, 2013. (SIGSPATIAL'13), p. 374–383. ISBN 978-1-4503-2521-9. Disponível em: <<http://doi.acm.org/10.1145/2525314.2525357>>. Citado na página 39.
- 62 NOULAS, A. et al. A random walk around the city: New venue recommendation in location-based social networks. In: *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*. [S.l.]: IEEE, 2012. Citado na página 39.
- 63 YING, J. J.-C. et al. Mining user check-in behavior with a random walk for urban point-of-interest recommendations. *ACM Transactions on Intelligent Systems and Technology*, Association for Computing Machinery (ACM), v. 5, n. 3, p. 1–26, sep 2014. Citado na página 39.
- 64 LIU, Y. et al. Exploiting geographical neighborhood characteristics for location recommendation. In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management - CIKM '14*. [S.l.]: ACM Press, 2014. Citado na página 39.
- 65 LI, X. et al. Rank-geofm: A ranking based geographical factorization method for point of interest recommendation. In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15*. [S.l.]: ACM Press, 2015. Citado na página 39.
- 66 LIAN, D. et al. Geomf: joint geographical modeling and matrix factorization for point-of-interest recommendation. In: *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '14*. [S.l.]: ACM Press, 2014. Citado 4 vezes nas páginas 39, 57, 73 e 75.
- 67 YUAN, Q.; CONG, G.; SUN, A. Graph-based point-of-interest recommendation with geographical and temporal influences. In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. New York, NY, USA: ACM, 2014. (CIKM '14), p. 659–668. ISBN 978-1-4503-2598-1. Disponível em: <<http://doi.acm.org/10.1145/2661829.2661983>>. Citado na página 40.
- 68 CHENG, C. et al. Fused matrix factorization with geographical and social influence in location-based social networks. In: *Twenty-Sixth AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2012. Citado na página 40.
- 69 ZHANG, J.-D.; CHOW, C.-Y. iGSLR: personalized geo-social location recommendation. In: *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - SIGSPATIAL '13*. [S.l.]: ACM Press, 2013. Citado na página 40.
- 70 LI, H. et al. Point-of-interest recommendations: Learning potential check-ins from friends. In: *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2016. (KDD '16), p. 975–984. ISBN 978-1-4503-4232-2. Disponível em: <<http://doi.acm.org/10.1145/2939672.2939767>>. Citado na página 40.

- 71 ZHANG, J.-D.; CHOW, C.-Y. GeoSoCa. In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15*. [S.l.]: ACM Press, 2015. Citado 7 vezes nas páginas 40, 44, 46, 53, 54, 58 e 75.
- 72 LIU, Y. et al. An experimental evaluation of point-of-interest recommendation in location-based social networks. *Proc. VLDB Endow.*, VLDB Endowment, v. 10, n. 10, p. 1010–1021, jun. 2017. ISSN 2150-8097. Disponível em: <<https://doi.org/10.14778/3115404.3115407>>. Citado 3 vezes nas páginas 40, 57 e 58.
- 73 ZIEGLER, C.-N. et al. Improving recommendation lists through topic diversification. In: *Proceedings of the 14th international conference on World Wide Web - WWW '05*. [S.l.]: ACM Press, 2005. Citado na página 44.
- 74 BENOUARET, I.; LENNE, D. A package recommendation framework for trip planning activities. In: ACM. *Proceedings of the 10th ACM Conference on Recommender Systems*. [S.l.], 2016. p. 203–206. Citado na página 44.
- 75 ZHANG, J.-D.; CHOWMEMBER, C.-Y.; LI, Y. iGeoRec: A personalized and efficient geographical location recommendation framework. *IEEE Transactions on Services Computing*, Institute of Electrical and Electronics Engineers (IEEE), v. 8, n. 5, p. 701–714, sep 2015. Citado na página 44.
- 76 GANTNER, Z. et al. MyMediaLite: A free recommender system library. In: *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys 2011)*. [S.l.: s.n.], 2011. Citado na página 57.