

Recomendação de Ingredientes com Base na Frequência de Utilização em Determinada Culinária

1st Mateus Henrique Pereira

dept. de computação

Centro Federal de Educação Tecnológica de Minas Gerais.

Divinópolis, Brasil

mateushenriqpereira02@gmail.com

Resumo—Este trabalho propõe um método para a recomendação de ingredientes com base na análise de dados relacionados a sua utilização em receitas culinárias. foi aplicado a teoria dos grafos para identificar e analisar padrões e interconexões entre ingredientes, permitindo a recomendação de novos ingredientes com maior probabilidade de complementar ou substituir outros em uma receita específica. A abordagem utiliza técnicas de processamento de dados e modelagem de grafos para capturar as relações entre ingredientes frequentemente utilizados em conjunto. Ao estabelecer essas relações, o sistema proposto é capaz de sugerir ingredientes compatíveis, ampliando as possibilidades de criação culinária e facilitando a descoberta de combinações inovadoras e harmoniosas.

I. INTRODUÇÃO

A culinária contemporânea é um conjunto de práticas gastronômicas e culturais, refletindo a riqueza dos ingredientes e tradições de diferentes regiões do mundo. Com a chegada da tecnologia e o acesso a conjuntos de dados abrangentes, tornou-se possível explorar e modelar as relações entre os ingredientes utilizados em diversas culinárias.

Neste estudo, exploramos um conjunto de dados que descreve a interconexão entre ingredientes presentes na culinária de um país específico. Essas conexões não apenas capturam a frequência com que os ingredientes são combinados em pratos culinários, mas também oferecem uma visão das associações robustas e recorrentes entre elementos-chave da culinária regional.

O foco principal deste trabalho é investigar e modelar esses dados interconectados por meio da análise de grafos. Foi possível realizar a construção do grafo a partir de arestas ponderadas, possibilitando compreender e quantificar as relações entre os ingredientes, destacando as conexões mais proeminentes que indicam associações consistentes na preparação de pratos culinários.

A abordagem de análise de grafos nos permite identificar os ingredientes mais fortemente conectados, aproveitando a força dessas conexões para oferecer recomendações de outros ingredientes frequentemente e robustamente utilizados em conjunto. Essa metodologia não apenas enriquece a compreensão dos elementos fundamentais da culinária de um país específico, mas também proporciona informações valiosas para chefs, entusiastas da culinária e profissionais da indústria alimentícia na criação e refinamento de novas receitas e experiências gastronômicas.

Este estudo visa contribuir significativamente para o entendimento das dinâmicas culinárias por meio da análise de grafos, evidenciando as interações fundamentais entre ingredientes e abrindo novos horizontes para a descoberta e exploração na culinária.

II. TRABALHOS CORRELATOS

Ao explorar o campo da recomendação de alimentos e planejamento de dietas, diversos estudos anteriores oferecem perspectivas valiosas sobre abordagens, modelos e metodologias empregadas nesse domínio.

Uma das primeiras incursões no desenvolvimento de sistemas computacionais para auxiliar no planejamento de dietas foi apresentada por (HAMMOND, 1986). O sistema proposto, conhecido como CHEF, buscava simular o processo cognitivo e as decisões humanas envolvidas na elaboração de receitas da culinária Szechwan.

Alguns dos principais modelos utilizados atualmente para recomendação de receitas e ingredientes incluem a Filtragem Colaborativa, a Filtragem Baseada em Conteúdo e os Modelos Híbridos. No que diz respeito a um dos modelos mais tradicionais de recomendação, são as técnicas baseadas em Filtragem Colaborativa, que constituem-se em diversas aplicações de recomendação, buscando similaridades em hábitos dos usuários para prever suas decisões futuras. Elas foram amplamente utilizadas em (FREYNE; BERKOVSKY, 2010), onde os autores analisaram técnicas variadas e uma versão híbrida das mesmas para a recomendação de alimentos. Em (OHATA; NISHIHARA; YAMANISHI, 2019) [2], é apresentado um sistema de recomendação de receitas que inova ao utilizar os dados das refeições dos usuários ao longo do dia para oferecer opções alimentares, incluindo pratos de restaurantes, visando alcançar uma dieta equilibrada e personalizada.

De Souza, Oliveira, e Moreira (2018) [4] focaram no desenvolvimento de um corpus de avaliações de hotéis em português brasileiro. Embora seu escopo seja diferente, a metodologia de coleta de dados e análise de informações específicas para criar um corpus pode ser comparável ao processo de coleta e preparação de dados em estudos de recomendação de alimentos.

Analisando o conteúdo presente nas pesquisas, em geral, a base de dados relaciona as receitas para fornecer índices de pratos e diferentes opções alimentares. No entanto, ao

contrário de muitas pesquisas anteriores que se apoiam em métodos convencionais de recomendação ou regras predefinidas, esta pesquisa adota uma perspectiva singular ao utilizar a análise de grafos. Isso nos permite mapear a relação de cada elemento, habilitando o uso de vários algoritmos para responder a diferentes perguntas. O diferencial deste trabalho está na disposição dos ingredientes, possibilitando diferentes opções para substituir um ingrediente ausente em específico.

III. METODOLOGIA

A metodologia proposta para este estudo começa com a etapa de coleta e preparação dos dados. Isso envolve a obtenção de conjuntos de dados que descrevem a relação entre ingredientes na culinária de um país específico. Esses dados passam por uma fase de limpeza e formatação para garantir a consistência, removendo informações duplicadas ou irrelevantes, preparando-os para análise.

Após essa etapa inicial, segue-se para a análise exploratória dos dados. Aqui, busca-se compreender a estrutura e a natureza das informações contidas nos conjuntos de dados. Isso ajuda a identificar padrões iniciais e tendências nos dados.

Em seguida, parte-se para a modelagem de grafos, utilizando ferramentas como a biblioteca NetworkX em Python. O objetivo é criar um grafo que represente as conexões entre os ingredientes. Esse grafo é construído atribuindo pesos às arestas, que indica a frequência com que os ingredientes são utilizados em conjunto, oferecendo uma representação visual das interações entre os elementos culinários.

Com o grafo modelado, aplicam-se algoritmos de análise de grafos para identificar os nós mais fortemente conectados, ou seja, os ingredientes frequentemente combinados na culinária desse país. Além disso, esses algoritmos permitem mapear padrões de associação entre os ingredientes, revelando informações sobre suas interações.

Uma etapa crítica consiste na implementação do algoritmo de recomendação. Esse algoritmo é desenvolvido para receber um ingrediente como entrada e sugerir os principais ingredientes que normalmente são combinados com ele em receitas culinárias. Esse processo pode envolver a aplicação de algoritmos de recomendação específicos para a análise dos dados do grafo.

Por fim, os resultados são avaliados e visualizados. Isso inclui a análise da eficácia do algoritmo de recomendação, utilizando métricas de desempenho adequadas, e a representação visual dos resultados por meio de gráficos ou outras representações visuais para facilitar a interpretação.

IV. RESULTADOS

Com base nos dados mencionados, esta seção tem como objetivo analisar os resultados obtidos na pesquisa. Antes de tudo, é importante ressaltar que o trabalho apresentado neste artigo possui código aberto [7], e a base de dados utilizada para realizá-lo foi adquirida no Kaggle [6], a qual foi organizada nos seguintes formatos:

Tabela I
TABELA COM 50 INGREDIENTES

Label	Id
milk	1
vegetable	2
tomatoes	3
...	...

Tabela II
TABELA DE RELAÇÕES ENTRE INGREDIENTES

Source	Target	Weight
garlic	purple onion	399
milk	vegetable oil	219
milk	tomatoes	78
natural yogurt	oil	1
...

Utilizando a linguagem Python e a biblioteca Pandas para a leitura de arquivos, foi realizado uma seleção aleatória de apenas 50 ingredientes para filtragem, conforme demonstrado na Tabela 1. As colunas 'Source' e 'Target' representam os nós que denotam a relação entre os ingredientes, enquanto 'Weight' refere-se às arestas (Tabela 2), indicando a quantidade de vezes em que esses ingredientes foram utilizados em conjunto. Para otimização e análise dos dados, uma filtragem nos dados originais por meio do notebook do Kaggle foi necessária, resultando em 768 combinações distintas de ingredientes. Utilizamos a biblioteca NetworkX para realizar os cálculos e construir o grafo a partir dessa base de dados.

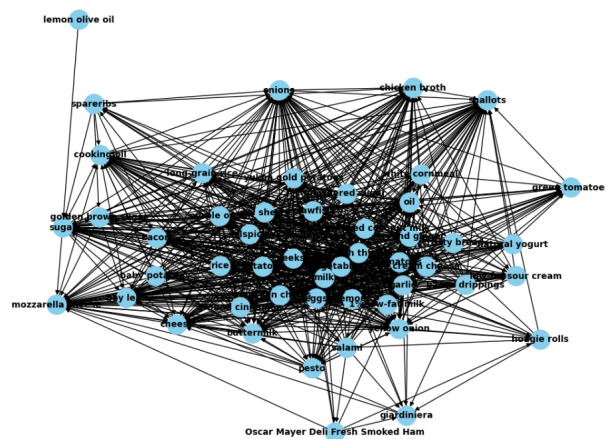


Figura 1. Grafo das relações entre os ingredientes

Este grafo é composto por nós que representam os ingredientes e arestas que denotam a frequência com que esses ingredientes aparecem juntos. Essa representação gráfica oferece uma representação visual da interconexão entre os ingredientes e a intensidade de suas relações.

Após organizar os dados de forma que seja possível ter uma análise mais visual do que está acontecendo, com o grafo já criado, utilizamos um algoritmo que encontra os vizinhos desse ingrediente no grafo (os ingredientes que estão

conectados a ele) e os ordena pelos pesos das conexões, do mais alto para o mais baixo. Para prosseguir com o código é necessário atribuir a variável 'ingredient_choice' um nome exatamente igual ao nome do ingrediente na tabela 1, para este exemplo foi usado o ingrediente 'cooking oil' como mostra a figura 2:

```
# Ingrediente de interesse  
ingredient_choice = 'cooking oil'
```

Figura 2. Variável de atribuição do Ingrediente desejado

Posteriormente, o algoritmo retorna os três principais ingredientes mais utilizados em conjunto com o ingrediente fornecido como mostra a figura 3:

```
Os 3 ingredientes mais utilizados com 'cooking oil':  
Ingrediente: garlic, Frequência: 191  
Ingrediente: onions, Frequência: 164  
Ingrediente: tomatoes, Frequência: 54
```

Figura 3. Resultado dos ingredientes mais utilizados em conjunto com o ingrediente 'cooking oil'

Ao aplicar algoritmos de recomendação, foi possível identificar os principais ingredientes frequentemente combinados com um ingrediente específico, como exemplificado na Figura 3 para o ingrediente 'cooking oil'.

Além disso, a análise revelou padrões consistentes de associação entre certos ingredientes, indicando preferências regionais ou tradicionais na culinária em estudo. Ingredientes como alho ('garlic'), cebola roxa ('purple onion'), e óleo vegetal ('vegetable oil') emergiram como os mais frequentemente associados, revelando-se como combinações fundamentais em diversas receitas que utilizam 'cooking oil'.

V. CONSIDERAÇÕES FINAIS

Este estudo explorou as relações entre ingredientes culinários, utilizando a teoria dos grafos para recomendar ingredientes com base em suas interações. A construção do grafo revelou conexões fortes entre ingredientes, permitindo a identificação de associações frequentes.

A aplicação do algoritmo de recomendação destacou a eficácia desse método na sugestão de ingredientes complementares. Além de ampliar a criatividade na culinária, esse enfoque oferece suporte valioso para pessoas que desejam substituir algum ingrediente ou até mesmo para profissionais na criação e aprimoramento de receitas.

Os resultados deste estudo não só ajudam a compreender a culinária, mas também indicam potenciais aplicações em outros domínios. A análise de grafos é uma ferramenta versátil para explorar padrões em diferentes conjuntos de dados, abrindo caminho para futuras pesquisas baseadas nesse estudo.

Em suma, este trabalho oferece uma abordagem inovadora na análise de dados inter-relacionados na culinária, apresentando possibilidades promissoras para futuras investigações em diversas áreas.

VI. TRABALHOS FUTUROS

Uma expansão do conjunto de dados pode fortalecer significativamente a análise proposta. A busca por informações mais abrangentes e diversificadas sobre distintas culinárias pode enriquecer a compreensão das interações entre ingredientes. Ao explorar dados mais variados, será possível examinar um leque mais amplo de ingredientes e suas relações, aprofundando nas relações entre ingredientes em cada culinária.

A introdução de métricas de avaliação para mensurar a precisão das recomendações de ingredientes geradas pode ser uma etapa crucial. A validação dos resultados por meio de comparações entre as recomendações do algoritmo e as preferências reais de chefs renomados, assim como a análise comparativa com receitas consagradas, traria robustez à metodologia. A recomendação poderia ser baseada na tabela nutricional de cada ingrediente, fornecendo um conteúdo mais abrangente e pessoal para pessoas que demandam de alguma restrição alimentícia.

Buscar parcerias com chefs, nutricionistas ou profissionais da indústria alimentícia para validar e aprimorar os achados é uma estratégia interessante. Experimentar novas abordagens e refinamentos na metodologia possibilita a descoberta de informações e uma compreensão mais aprofundada das dinâmicas dos ingredientes na culinária e além dela.

REFERÊNCIAS

- [1] HAMMOND, K. J. Chef: A model of case based planning. In: AAAI. 1986.
- [2] OHATA, T. NISHIHARA, Y. YAMANISHI, R. Foods recommendation system for meals out in nutrient balance. In: Proceedings of the 11th Workshop on Multimedia for Cooking and Eating Activities. New York, NY, USA: Association for Computing Machinery, 2019. Disponível em: Link
- [3] de Souza, J. G. R., de Paiva Oliveira, A., and Moreira, A. (2018). Development of a brazilian portuguese hotel's reviews corpus. In International Conference on Computational.
- [4] Oliveira, E. G., Britto, L. F. S., Pacifico, L. D. S., & Ludermir, T. B. Recipe Recommendation and Generation Based on Ingredient Substitution. Disponível em: Link
- [5] Kaggle - Conjunto de dados de ingredientes de receitas
- [6] Kaggle - Conjunto de dados de ingredientes de receitas em formato de grafo
- [7] Github - Recomendação de ingredientes