Unidade 01 Trabalho 02

Análise de Assortatividade em Grafos de Ingredientes da Culinária Brasileira

New Podcast Episode

Comida mais cara e alimento fake na prateleira



Today • 34 min 30 sec





• • •

Episode Description

Uma pesquisa do Datafolha mostrou que a inflação fez mais da metade dos brasileiros (58%) reduzir a quantidade de alimentos que costuma comprar. Entre os mais pobres, o percentual sobe para 67%. Quem não está comprando menos está fazendo substituições: 50% dos entrevistados, por exemplo, disseram ter trocado a marca do café que tomam.

Objetivo

Construir um grafo de co-ocorrência de ingredientes a partir de pelo menos 50 receitas populares da culinária brasileira, classificar os ingredientes por tipo (proteína, carboidrato, vegetal etc.) e analisar a **assortatividade** do grafo com base nesses tipos, discutindo se a gastronomia brasileira tende a valorizar combinações homogêneas (entre ingredientes do mesmo tipo) ou contrastantes (entre tipos diferentes).

Criação da Base de Dados

Uso de Modelos de Linguagem (LLM)

- Utilizar um modelo de linguagem (pode ser uma LLM como ChatGPT Gemini, ou outro) para:
 - Gerar a descrição dos ingredientes a partir de uma imagem da receita.
 - Extrair automaticamente os ingredientes descritos na receita.
 - Classificar cada ingrediente em uma das seguintes categorias:
 - Proteína
 - Carboidrato
 - Vegetal
 - Fruta
 - Laticínio
 - Gordura
 - Condimento
 - Outro

Organização da base de dados em arquivo .CSV

nome_receita	ingredientes	tipos_ingredientes		
Feijoada	feijão preto, carne de porco, arroz	leguminosa, proteína, carboidrato		

Análise da assortatividade

Construir um grafo:

- Nós = ingredientes
- Arestas = ligação entre ingredientes que aparecem na mesma receita
- Atributo dos nós = tipo de ingrediente

Calcular:

- Coeficiente de assortatividade por tipo usando networkx.attribute_assortativity_coefficient(G, "tipo")
- Visualizar o grafo com layout adequado e coloração por tipo (nxviz)

Análise da assortatividade

Discutir o valor obtido:

- O valor é próximo de 1? → Homofilia: ingredientes do mesmo tipo tendem a estar juntos.
- O valor é próximo de 0? → Combinação aleatória.
- O valor é negativo? → Heterofilia: há preferência por ingredientes de tipos diferentes.

Entregável

A entrega deverá ser feita em um repositório (submeter o link no sigaa) no GitHub, contendo:

- 1. README.md
 - Nome completo e identificação (ex: matrícula, curso)
 - Breve descrição da tarefa, incluindo um link para um vídeo de até 5min apresentando o trabalho.
 - Discussão escrita de até 500 palavras com os principais achados da análise
 - Inserção da imagem do grafo gerado (em Markdown)
- 2. Ingredientes.csv (arquivo com os dados organizados)
- 3. Ingredientes.ipynb (leitura do CSV, construção do grafo com NetworkX, atribuição dos tipos como atributos dos nós, cálculo da assortatividade por tipo, geração da visualização do grafo, Interpretação dos resultados.
- 4. /images/ pasta contendo imagens geradas durante o trabalho (grafo, layouts, etc.)
- 5. Ilm_info.md (arquivo com o nome do modelo de linguagem utilizado (ex: ChatGPT 4.0, Gemini Pro, Claude, etc.), prompt(s) utilizados para identificar ingredientes)

Nota

Trabalho Individual

2,5 pontos na Unidade 1

Prazo: 27 de abril

Obs.: a não entrega do vídeo irá impactar nota zero.

Week 01

17/03 and 19/03 - Planning

Week 02

24/03 - Course Presentation (Python) 26/03 - Network Elements I (Networkx) + (TA)

Week 03

31/03 - Network Elements II 02/04 - [Project #01] PPgEEC Co-Authorship Network (16/04)

Week 04

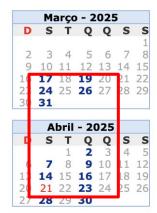
07/04 - [Project #01] 09/04 - [Project #01]

Week 05

14/04 - Homophily and Assortativity (TA) 16/04 - Paths, Distance, and Walks (TA)

Week 06

21/04 - no class 23/04 - Q&A about TA





Junho - 2025								
D	S	Т	Q	Q	S	S		
1	2	3	4	5	6	7		
8	9	10	11	12	13	14		
15	16	17	18	19	20	21		
22	23	24	25	26	27	28		
29	30							

