

“Creación de un Chatbot basado en el libro de recetas de cocina de Nitza Villapol”

Curso de Posgrado Análisis de Redes Complejas

Dr. Yudivián Almeida Cruz

Autores: Roberto Martí Cedeño

Rafael Martínez Martínez

Facultad Matemática-Computación

Universidad de La Habana

26 de junio de 2023



Objetivo:

Implementar un chatbot que genere respuestas de recetas de cocina y además proponer ingredientes que sean los sustitutos en el caso de no contar con un ingrediente determinado para elaborar una receta.

Propuesta de Implementación:

Se desarrolló un programa en el lenguaje Python, usando las librerías: networkx, pandas-ods-reader y matplotlib, capaz de resolver diferentes consultas generadas por el usuario.

Las funcionalidades son el resultado de una base teórica sólida en el tema del análisis de redes complejas y precisamente en el trabajo con los grafos.

Para intentar resolver el problema inicial se diseña un grafo bipartito en el cual se representan los ingredientes y las recetas como los nodos del grafo. Los ingredientes se agrupan en una partición y las recetas en otra partición y las aristas se van insertando con la estructura <ingrediente, receta>, es decir, los ingredientes se conectan a las recetas, nunca ingredientes con ingredientes ni recetas con recetas.

Categorías:

Las categorías muestran los grupos globales en el que se encuentran ubicadas las recetas. Constituyen los grupos más grandes y en el nivel jerárquico del grafo están ubicados en la cima.

Ejemplo:

- 1 - Salsas, Aliños y Mayonesa
- 2 - Vegetales y Frutas

Subcategorías:

Son grupos más pequeños y de mayor especialización para las recetas. Las subcategorías forman parte de categorías más generales.

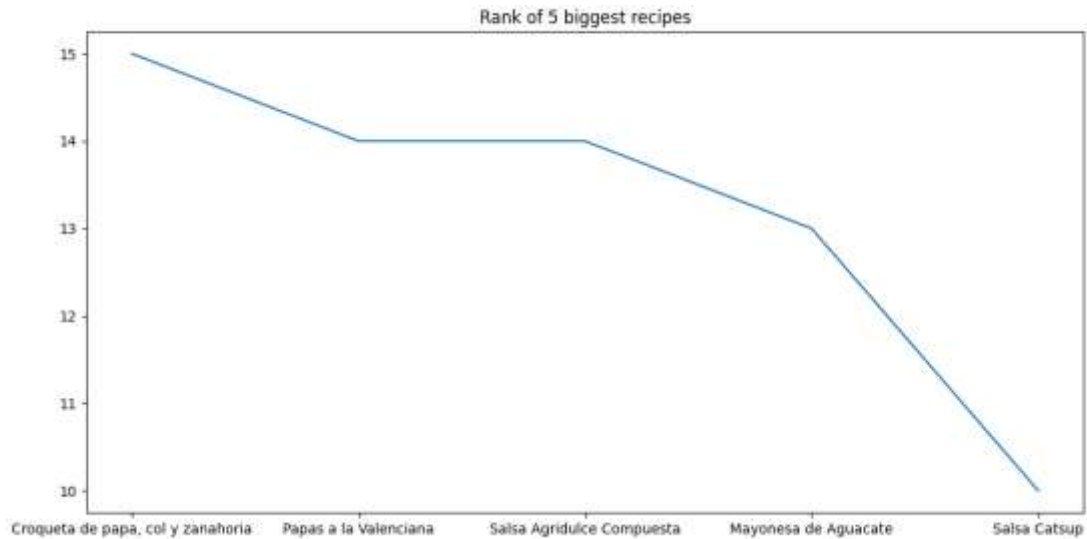
Ejemplo:

- 1 - Subcategoría: Mayonesa, Salsas
- 2 - Subcategoría: Papa

Recetas:

Las recetas constituyen el conglomerado de ingredientes que bajo las cantidades adecuadas y con el procesamiento de los mismos, dígame: picados, hervidos, en trocitos, en forma de pasta, en forma

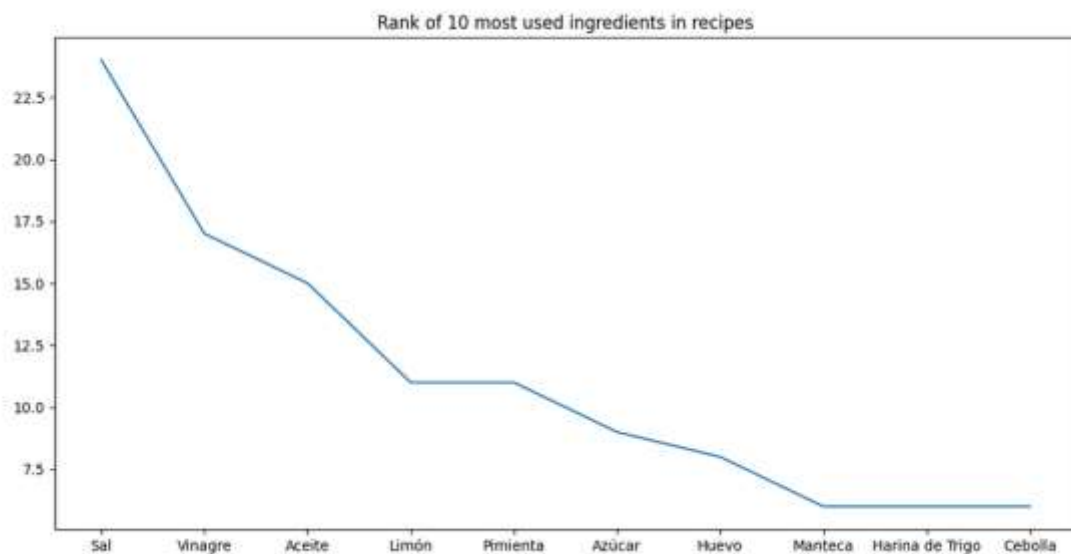
de jugos, etc, y bajo el tiempo de cocción seleccionado, dan a lugar un plato de la cocina típica cubana.



Ingredientes:

Los ingredientes son todos los tipos de alimentos necesarios para armar una receta de cocina. Pueden ser frutas, vegetales, viandas, carnes, bebidas, jugos, refrescos entre otros.

Es de especial interés conocer todos los tipos de ingredientes en el grafo y aquellos los cuales más participan en recetas de cocina.



Degree:

Conocido como Centralidad de Grado calcula este índice para cada nodo del grafo y se normaliza por el mayor posible grado ($n-1$) para los grafos de n nodos. Colores más oscuros representan nodos con baja centralidad, colores más claros con más alta.

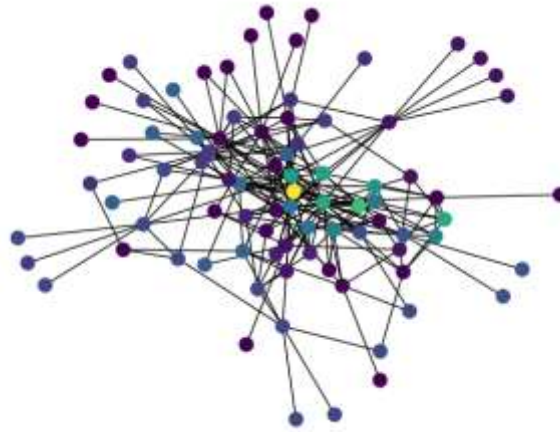


Figura 1. Centralidad de Grado

Closeness:

Conocida como Centralidad de Cercanía. Se calcula como $1/(\text{promedio de las distancia a los otros nodos})$. Es normalizada por $n-1/N-1$ donde n es el número de nodos de la componente que contiene al nodo y N el número de nodos del grafo.

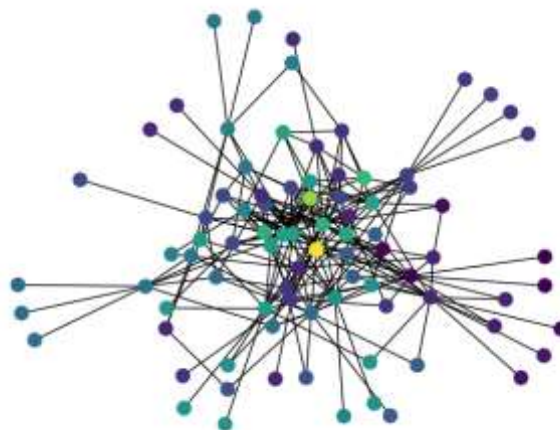


Figura 2. Centralidad de Cercanía

Eigenvector:

Conocida como Centralidad del Vector Propio. La idea fundamental es cuán central es un nodo en función de cuán central son sus vecinos.

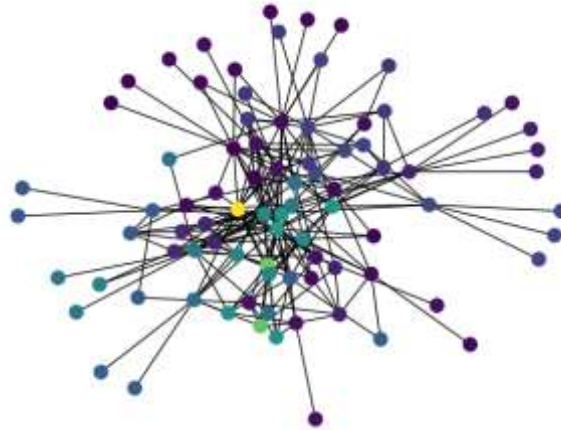


Figura 3. Centralidad del vector propio.

Betweenness:

Conocida como Centralidad de Intermediación. Se calcula en base a la cantidad de caminos mínimos del grafo que contienen a un nodo. Se normaliza por $2/((n-1)(n-2))$ si el grafo es no dirigido y por $1/((n-1)(n-2))$ si es dirigido con n número de nodos. Se les asocia con los cuellos de botella.

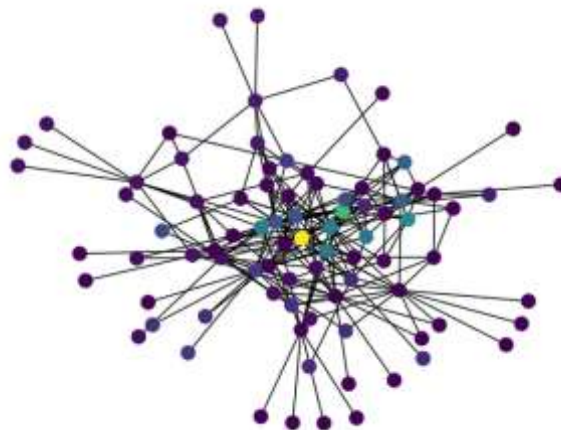


Figura 4. Centralidad de Intermediación

Pagerank:

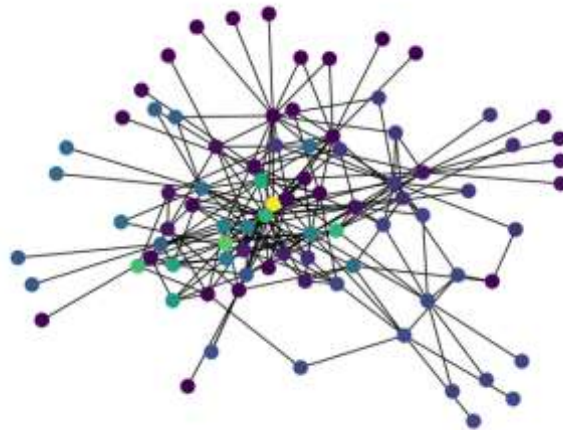


Figura 5. Pagerank

Closeness_vitality:

La medida de cuánto se incrementa el costo de la comunicación si no existe la arista/el nodo.

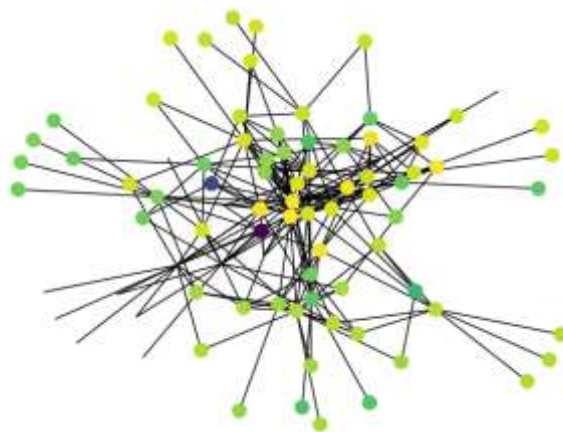


Figura 6. Closeness vitality en los nodos del grafo.

Edge_betweenness

Girvan_newman

Asyn_fluid

Asyn_lpa

K_clique

Kernighan_Lin

Ingr_similarity

Recipes_similarity

Las funcionalidades del chatbot que se implementaron se basaron mediante un sistema de comandos definidos de la siguiente manera:

<i>Flujo de ejecución del chatbot</i>	
<i>Load_graph</i>	<i>Comienza el chat. Crea un grafo desde un directorio con archivos .ods</i>
<i>exit</i>	<i>Termina el chat</i>
<i>Generales</i>	
<i>Categories</i>	<i>Muestra todas las categorías de recetas del grafo</i>
<i>Subcategories</i>	<i>Muestra todas las subcategorías agrupadas por categorías</i>
<i>Recipes</i>	<i>Imprime la cantidad total de recetas, el nombre de todas y plotea en un gráfico las 5 más grandes en cuanto a los diversos tipos de ingredientes.</i>
<i>Ingredients</i>	<i>Imprime la cantidad total de ingredientes, el nombre de ellos y plotea los 10 más repetidos en las recetas.</i>
<i>Índices de Centralidad</i>	
<i>Degree</i>	<i>Muestra la centralidad de Grado en el grafo</i>
<i>Closeness</i>	<i>Muestra la centralidad de Cercanía en el grafo</i>
<i>eigenvector</i>	<i>Muestra la centralidad de vector propio del grafo</i>
<i>betweenness</i>	<i>Muestra la centralidad de Intermediación</i>
<i>Pagerank</i>	<i>Muestra el índice pagerank en los nodos del grafo</i>
<i>Closeness_vitality (on node)</i>	<i>Muestra la vitalidad de cercanía en los nodos del grafo.</i>
<i>Edge_betweenness</i>	<i>Muestra la intermediación de aristas en el grafo</i>
<i>Detección de Comunidades</i>	
<i>(Algoritmos para detección de comunidades presentes en networkx)</i>	
<i>1- Girvan_newman</i>	
<i>2- Asyn_fluid</i>	
<i>3- Asyn_lpa</i>	
<i>4- K_clique</i>	
<i>5- Kernighan_Lin</i>	
<i>Similaridad en recetas e ingredientes</i>	
<i>Ingr_similarity X</i>	<i>Dado un ingrediente X en el grafo se muestran los tres ingredientes más similares computados.</i>
<i>Recipes_similarity X</i>	<i>Dado una receta X en el grafo se muestran las tres recetas más similares a X encontradas en el grafo.</i>

Recomendaciones

Bibliografía