

Propuesta de Aplicación Web para el costeo Gastronómico

TORRES, David †*, VILLEGAS, Rodrigo y VARGAS, Elizabeth

Recibido Febrero 15, 2017; Aceptado Junio 28, 2017

Resumen

El Web Scraping es una técnica que permite la extracción de contenido de varios sitios web, recopilando información que resulte de interés para el usuario; a fin de ser presentada de forma ordenada para su posterior utilización y análisis. El presente proyecto tiene como finalidad crear una herramienta que permita reducir hasta en una tercera parte el tiempo empleado en el proceso de costeo que realizan estudiantes de gastronomía, chefs profesionales y personas en general dedicadas a la comercialización de alimentos preparados. Lo anterior, debido a que estas personas están enfocadas en las actividades propias de la preparación, y disponen de poco tiempo para la solicitud de presupuestos y/o búsqueda de ingredientes. Para esto se desarrolló una herramienta web donde, una vez extraída la información y almacenada en una base de datos, ésta puede ser consultada a fin de realizar el presupuesto correspondiente a una receta; llegando a reducir considerablemente el tiempo empleado en dicho proceso.

Web scrapping, text mining, web data extraction, costeo gastronómico

Abstract

Web data extraction (or web scraping) is a technique that allows the extraction of content from various websites, collecting information that is of interest to the user; In order to be presented in an orderly manner for later use and analysis. The purpose of this project is to create a tool to reduce the time spent in the process of costing by students of gastronomy, professional chefs and people in general engaged in the commercialization of prepared foods. The above, because these people are focused on the activities of the preparation, and have a short time for the request for budgets and/or search for ingredients. A web tool was developed for this, where, once the information has been extracted and stored in a database, it can be consulted in order to make the budget corresponding to a recipe, reducing considerably the time spent in this process.

Web scrapping, text mining, web data extraction, gastronomic cost

Citación: TORRES, David, VILLEGAS, Rodrigo y VARGAS, Elizabeth. Propuesta de Aplicación Web para el costeo Gastronómico. Revista de Tecnologías Computacionales. 2017, 1-2:46-52.

*Correspondencia al Autor (Correo Electrónico: datorres@itesi.edu.mx)

† Investigador contribuyendo como primer autor.

Introducción

La planificación es el punto inicial del proceso de producción de un servicio de banquetes y eventos especiales de catering. Un gerente de banquetes debe partir del diseño del menú para poder controlar todas las actividades de producción y servicio. Además, debe reunirse periódicamente con el personal a su cargo para discutir los puntos clave de cada evento concertado con los clients.

Dentro de éste proceso, conocer con anticipación el número de asistentes al evento, facilita la adquisición de los ingredientes en las cantidades adecuadas para la elaboración de los diferentes platillos. Además, es responsabilidad del chef tener todo bajo control, para estar en posibilidad de proporcionar el servicio atendiendo a los requerimientos del cliente; desde la orden del banquete o evento, hasta la prestación del servicio y su análisis posterior para retroalimentación.

Una práctica común en éste tipo de negocios es la estandarización de las recetas, de forma que no solo se preste atención a la calidad y una excelente presentación en los platillos, sino que también se tenga un adecuado control sobre las porciones; con lo cual se puede tener un mayor control sobre la producción y los costos del evento. No obstante, el proceso de costeo frecuentemente representa un problema debido al poco tiempo del que se dispone, y al escaso conocimiento sobre el proceso de elaboración de presupuestos.

De no tener especial cuidado en el proceso de costeo, se pueden tener repercusiones en la atención al cliente, provocando inconformidades, costeos inexactos, pérdidas monetarias; y en algunos casos la pérdida del cliente.

Es por lo anterior que el presente proyecto tiene como objetivo desarrollar una herramienta para el cálculo de los ingredientes que permita reducir el tiempo hasta en una tercera parte del tiempo actual que se invierte al realizar un costeo.

En el presente artículo se explica de forma general el desarrollo de la herramienta mencionada en los siguientes apartados:

El Estado del Arte muestra una visión general de la técnica Web Scraping y su uso en diferentes proyectos, en diferentes contextos. La Metodología Empleada explica paso a paso el procedimiento seguido para la elaboración de la herramienta. En Resultados Experimentales se mencionan los resultados de las pruebas realizadas con personal relacionado con la gastronomía; y por último, en Conclusiones y Trabajo Futuro, se mencionan las características que se determinaron como necesarias de agregar a la herramienta; así como el estado en que se concluyó la herramienta desarrollada.

Estado del Arte

El Web Scraping es una técnica que permite extraer y leer el contenido de uno o varios sitios web mediante diferentes herramientas, como motores de búsqueda; la cual permite analizar dicha información. Un ejemplo de ello es la comparación de precios que se obtienen de sitios web de diferentes negocios, tales como hoteles y agencias de viajes; en donde toda la información que cumpla con ciertos criterios se presenta en una sola página, evitando que se invierta tiempo en buscar en todos los sitios uno por uno.

Al hacer uso de web Scraping se automatiza el proceso de búsqueda de información, evitando que las personas se enfrenten a sitios que son difíciles de comprender por contar con una estructura compleja.

De ésta manera, una vez extraídos los datos, éstos pueden ser reutilizados mejorando la estructura en la que se presentan, facilitando su lectura y comprensión.

En [1] los autores muestran algunas metodologías para lograr obtener resultados de búsquedas que se hacen por parte de los votantes, para saber por quién votar y quien es su candidato preferido. Esta cantidad de información se obtiene de diferentes sitios, en los cuales las herramientas que ofrece el Web Scraping se adaptan en dos fases: la extracción de información específica y el análisis de los datos.

El web Scraping se utiliza en varios ámbitos, uno de ellos es parte importante en una sociedad: las elecciones. Es por eso que se encuentra interesante como a partir de diferentes medios de comunicación, algunos partidos intervienen para dar a conocer sus distintos temas para la sociedad. Con algo tan simple como esto, los políticos comunican de forma dinámica sus campañas, y no solo aburren a los votantes con sus largos discursos, logrando obtener apoyo de los jóvenes que utilizan las redes sociales.

En [2] los autores presentan DEiXTo, una herramienta orientada a la extracción de datos en la web, en aquellas páginas que contienen información de interés e importancia, y que en ocasiones no cuentan con la presentación de datos como se necesita para su entendimiento. De ésta manera, DEiXTo se encarga de reunir la información para presentarla en una mejor estructura de fácil comprensión para sus lectores DEiXTo hace la reutilización de contenido web, lo cual favorece a cualquiera que necesite de cierta información y le de otros usos de interés para los propósitos de los usuarios.

Para ésto, los desarrolladores establecen reglas en la extracción, que es lo que hace posible obtener los datos con mayor precisión, aunque sea de diferentes sitios.

En [3] se utiliza el web scraping y el conocimiento obtenido para realizar pruebas sobre BOCYL (Boletín Oficial de Castilla y León). BOCYL es el periódico oficial en el que se publican documentos que sean solicitados de manera gratuita.

El objetivo consiste en aplicar la extracción de información del boletín para observar los resultados y alojar la información en una base de datos relacional. Para cada etapa de aprendizaje se tuvieron que realizar mediciones de qué tan factible era la extracción, tanto del boletín como de otros sitios web; y se dan a conocer algunos inconvenientes de aplicar esta técnica, siendo uno de ellos el que algunos sitios están protegidos del scraping y no se tienen los permisos.

También se describen las diferentes pruebas que se utilizan y se dan a conocer las herramientas, así como las librerías y equipo que se utilizó para realizarlas. En el proyecto se desarrollan distintas tareas que se llevan a cabo para su desarrollo desde la planificación, análisis, diseño, implementación y pruebas, a fin de detectar errores y realizar casos de uso del proyecto.

Como resultado de este proyecto se tiene la obtención y almacenamiento de la información del BOCyL en una base de datos que permite reutilizar la información para cualquier actividad diferente, además de plantear diferentes usos que se pueden dar a la información.

Metodología empleada

En éste apartado se detallan los pasos que se siguieron en el proyecto, siendo ésta una investigación descriptiva. En ella se incluyen los siguientes tipos de estudios: encuestas y casos de desarrollo. También se precisan los métodos y procedimientos (diseño de la investigación) que se utilizaron durante el desarrollo de la investigación.

Análisis y selección del método de costeo

El costeo es el proceso de investigación de precios y su acumulado, necesario para establecer el presupuesto total de un proyecto. A su vez, los costos son todas las salidas (egresos) de mercancía que en su momento fue comprada con la intención de procesarse y venderse. Para la determinación de un sistema de costos en cualquier empresa específica, es necesario tomar en cuenta varios factores, entre estos están:

- El tipo de mercado que abastece y su posición dentro de él.
- La naturaleza de sus procesos fabriles.
- El grado de complejidad en las etapas de elaboración.
- El surtido de artículo que se fabrica.

Para la determinación del tipo de costeo, se realizaron encuestas al personal experto en el área de gastronomía para determinar el método más utilizado, más práctico y el que más se le facilita de acuerdo a sus características; resultando seleccionado el *sistema de costos por órdenes específicas*.

Proceso de Estandarización

La estandarización consiste en aquellos procesos aplicados a recetas que han sido probadas y comprobadas varias veces y adaptadas a las necesidades.

Por lo tanto, conociendo el tipo de costeo, es necesario identificar el mejor método de estandarización. Los métodos de estandarización más conocidos son:

- Método sistemático
- Método Factorial
- Método Porcentual

Al igual que en la etapa anterior, se buscó determinar cuál método es el más utilizado por chefs profesionales y estudiantes de gastronomía; resultando seleccionado el método porcentual. El método porcentual es ampliamente recomendable para la producción de gran escala, en la que el tamaño de los lotes puede variar mucho. Una vez que se establece el porcentaje del ingrediente, éste permanece constante para todas las adaptaciones futuras. Los aumentos y disminuciones a las recetas se obtienen multiplicando el porcentaje de cada ingrediente por el peso total deseado. Al verificar los ingredientes para que la receta resulte bien balanceada, se dispone del porcentaje de cada ingrediente.

Adicional a esto, se preguntó abiertamente a las personas encuestadas ¿cuál lugar es la primera opción al momento de comprar?, dando como primera opción las tiendas de autoservicio por la distribución y acomodo de los productos, lo que se ve reflejado en menor inversión de tiempo en la compra; y como segunda opción una central de abastos, en caso de compras al mayoreo, además de contar con mayor variedad de productos.

Configuración de Web Scraper

Para el proyecto se utiliza el Web scraper integrado en las herramientas para desarrolladores de Google Chrome; en donde se encuentran todas las herramientas con las que se debe iniciar a crear los nuevos mapas del sitio y de las páginas de las que se extraerán los datos.

Se debe introducir un nombre para identificar el apartado del que se desea obtener información, e introducir la URL de la página principal (Figura 1).

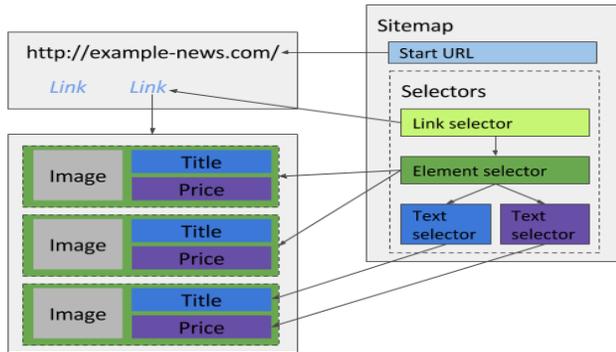


Figura 1 Ejemplo de mapa del sitio para el web scraper

Existen selectores de contenido, de texto y tipo link; los cuales ayudan en la navegación entre las páginas del sitio y extraen los datos. Algunos selectores y conceptos básicos que requiere conocer son:

- Mapa del sitio. Consiste en especificar la dirección URL de inicio. Esta es la URL desde la que se iniciará la obtención de información.
- Selector de texto. Permite extraer el texto del elemento seleccionado y de todos sus elementos secundarios.
- Selector de enlace. Se utiliza para la selección de enlace y la navegación web.
- Selector de elementos. Es para la selección de información que contiene varios elementos de datos.

El web scraper ejecutará los selectores en el orden en que se organizan en la estructura de árbol. Primero tendrá que recorrer la página web con un selector o más selectores tipo link (enlace), el cual solo realizara un escaneo de las paginas disponibles en ese mismo nivel.

Si existe paginación es necesario utilizar estos selectores hasta llegar a los datos finales en la última sección. De esta manera se marca una ruta por la que se debe pasar. Al llegar en la última sección se podrán utilizar selectores tipo elemento que hacen que en toda esa página se puedan elegir todas las características como imágenes, precios, títulos. Esto hace que para los selectores tipo texto sea más fácil encontrar solo aquellos productos que fueron seleccionados anteriormente con mayor exactitud.

En algunas ocasiones, las páginas que se encuentran se desplazan para mostrar más elementos o se desprenden en otras secciones, por lo que se utilizan otros selectores que funcionan de manera similar al elemento selector. Así también, nos encontramos como ventanas emergentes las cuales tiene un trato diferente. Después de haber creado selectores para el mapa del sitio, se puede inspeccionar la estructura de árbol de selectores en el panel gráfico de selección.

Una vez configurada la herramienta, se realizaron pruebas con el sitio de Walmart, dando como resultado la información mostrada en la Figura 2.

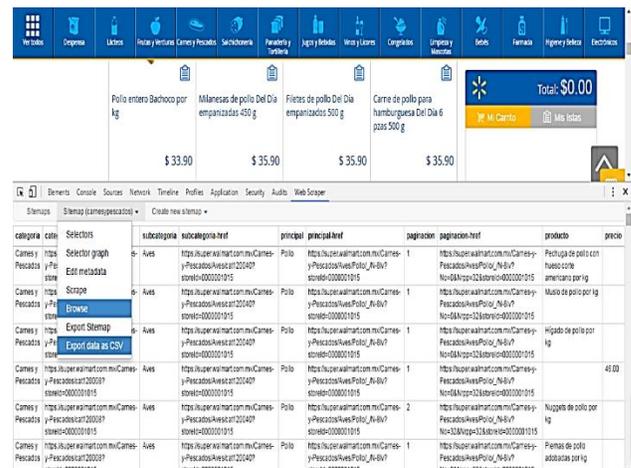


Figura. 2 Resultados de la obtención de datos del sitio web de Walmart

Creación de la Base de Datos

Se utilizaron 10 archivos, los cuales representan a los departamentos del sitio web donde se extrajo la información. Estos a su vez se unificaron en una sola base de datos creada en Navicat, quedando como se muestra en la Figura 3.

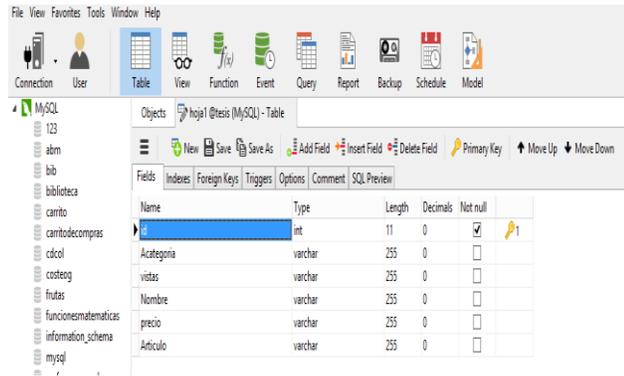


Figura 3 Base de datos creada

Desarrollo de la Plataforma

En la Figura. 4 se muestra la interfaz del prototipo buscador de productos, en donde se pueden observar los resultados que se obtuvieron al hacer la consulta por el usuario, a la base de datos.



Figura 4 Buscador Ajax

Por ultimo, en la Figura 5 se muestra la tabla que realiza las operaciones donde se cargan los productos seleccionados y se realiza el método de estandarización para el proceso de costeo.

Ingredientes	Cantidad en Receta	Porción	Precio	Cantidad Paquete	Contenido	Costo por unidad	Costo Por Grama	
queso cr	1	<input checked="" type="radio"/> Pieza <input type="radio"/> Taza <input type="radio"/> Cucharada	29.9	1	pt	29.9	29.9	
leche ev	1	<input checked="" type="radio"/> Pieza <input type="radio"/> Taza <input type="radio"/> Cucharada	12.3	1	pt	12.3	12.3	
leche cor	1	<input checked="" type="radio"/> Pieza <input type="radio"/> Taza <input type="radio"/> Cucharada	9.4	1	pt	9.4	9.4	
limones	1	<input checked="" type="radio"/> Pieza <input type="radio"/> Taza <input type="radio"/> Cucharada	12.5	1	pt	12.5	12.5	
galletas r	2	<input checked="" type="radio"/> Pieza <input type="radio"/> Taza <input type="radio"/> Cucharada	69	1	pt	69	138	
						C. Original	9	193.1
						C. Ingresada	9	193.1
<input type="button" value="Calcular"/>								
Total:						193.1		
Ver catalogo								

Figura 5 Interfaz para la obtención del costeo

Resultados Experimentales

Se realizó un muestreo por conveniencia, la cual es una técnica no probabilística donde los sujetos son seleccionados por la accesibilidad y proximidad de los sujetos para el investigador [4].

A las personas que participaron en el estudio se les otorgo una receta previamente elaborada, describiendo en ella los ingredientes necesarios, los pasos a seguir y la cantidad de personas para la cual está elaborada. Se les pidió realizar un costeo de forma manual, tomando el tiempo del proceso; dando como resultado un tiempo promedio de 10 minutos y 38 segundos. Posteriormente se les pidió realizar el mismo procedimiento con el prototipo costeo gastronómico.

Para demostrar la curva de aprendizaje Humano – Computadora se realizó tres veces el mismo procedimiento para comprobar si efectivamente se reduce el tiempo (grafico 1). Con los resultados obtenidos se puede observar que desde la primera prueba se reduce casi a la mitad el tiempo invertido en el procedimiento de costeo. Además de ello, se comprueba que, entre mayor interacción con el prototipo, se reduce aún más el tiempo empleado en el proceso de costeo.

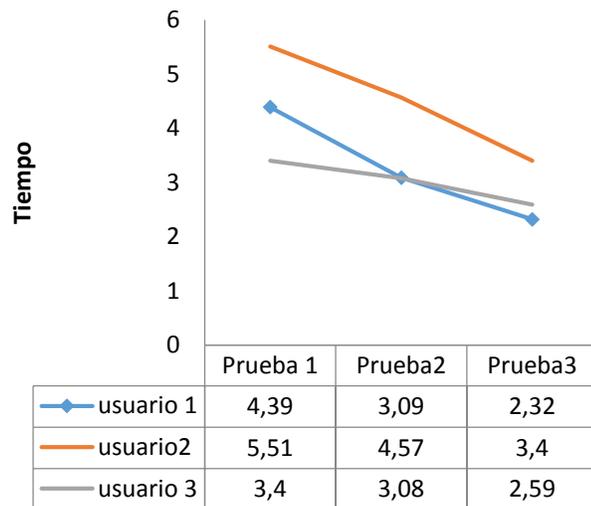


Gráfico 1 Curva Humano - Computadora

Conclusiones y Trabajo Futuro

Con la herramienta desarrollada, los participantes redujeron el tiempo de costeo en dos terceras partes del tiempo actual invertido con un proceso de estandarización de recetas para obtener calidad y el tamaño estándar de porciones. Los resultados comenzaron a ser notorios conforme hubo mayor interacción con la herramienta, debido a que entre más familiarizado se encuentre el usuario con la tecnología, más fácil será su uso y el tiempo estimado del proceso se reduce.

Lo anterior se vió reflejado en los resultados de la curva de aprendizaje, los cuáles descienden drásticamente de acuerdo al tiempo que se invierte al realizar un costeo. Sin embargo, también se reflejaron problemas e inconvenientes no previstos como la extracción de datos por web Scraping, debido a la implementación de HTML5 en las nuevas páginas web. Dentro de las recomendaciones para continuar con el Proyecto, se pueden mencionar las siguientes:

- Implantación de las conversiones de gramaje, puesto que se necesita aumentar

todo el glosario de unidades de medida del área gastronómica.

- Con la nueva versión del HTML5, es necesario revisar posibles problemas en la seguridad, en cuanto al proceso de web scraping.
- Búsqueda y extracción de datos de imágenes, en páginas que promocionen su información por medio de catálogos.
- Implantación de actualizaciones automáticas.
- Opción de alertas de ofertas en productos solicitados.

Referencias

- [1] A. Tadeo, E. Gómez, C. A. Berdejo, J. Montero, A. Calderón y R. Ibarra. *Metodologías para el análisis político utilizando Web Scraping*. Research in Computing Science. 2015.
- [2] F. Kokkoras, K. Ntonas y N. Bassiliades. *DEiXTO: a web data extraction suite*. Proceedings of the 6th Balkan Conference in Informatics, pp. 9-12. Thessaloniki, Grecia. 2013.
- [3] C. Hernández y M. A. Martínez. *Aplicación de Técnicas de Web Scraping al BOCyL*. Universidad de Valladolid. 2014.
- [4] J. H. MacMillan y S. Schumacher, *Research in Education: A Conceptual Introduction*. 5a. Ed. Addison Wesley Longman., 2001.
- [5] J. Muñoz et. al. *Temas de Diseño en Interacción Humano Computadora*. Iniciativa Latinoamericana de Libros de Texto Abiertos. 2014.