

INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO



**SISTEMA DE RECOMENDACIÓN DE HOTELES  
SIMILARES**

TESIS

QUE PARA OBTENER EL TÍTULO DE  
MAESTRO EN CIENCIA DE DATOS

PRESENTA

FELIPE GERARD VALDÉS

ASESOR:

MÉXICO, D.F.

2015

“Con fundamento en el artículo 21 y 27 de la Ley Federal del Derecho de Autor y como titular de los derechos moral y patrimonial de la obra titulada **“MÉTODOS NUMÉRICOS PARA AJUSTAR MODELOS DE REGRESION LOGÍSTICA REGULARIZADA”**, otorgo de manera gratuita y permanente al Instituto Tecnológico Autónomo de México y a la Biblioteca Raúl Bailleres Jr., autorización para que fijen la obra en cualquier medio, incluido el electrónico, y la divulguen entre sus usuarios, profesores, estudiantes o terceras personas, sin que pueda percibir por tal divulgación una contraprestación”

**Felipe Gerard Valdés**

---

FECHA

---

FIRMA

# Índice general

<b>Prólogo</b>	<b>IV</b>
<b>1. Análisis previo</b>	<b>2</b>
1.1. Análisis de la competencia . . . . .	2
1.2. Análisis del sistema actual . . . . .	6
1.3. Diagnóstico . . . . .	6
1.4. Wish List . . . . .	6
<b>2. Modelo propuesto</b>	<b>7</b>
2.1. Puntos a resolver . . . . .	7
2.2. Distancia . . . . .	7
2.3. Precio . . . . .	7
2.4. Similitud de hoteles . . . . .	7
2.5. Distancia revisitada . . . . .	7
2.6. Integrando todo . . . . .	7
<b>3. Obtención de la información</b>	<b>8</b>
3.1. Formato necesario para el modelo . . . . .	8
3.2. Calidad de la información . . . . .	8
3.3. Sistema de respaldo . . . . .	8
<b>4. Desempeño</b>	<b>9</b>

# Índice de figuras

# Índice de tablas

# Prólogo

# Capítulo 1

## Análisis previo

Antes de desarrollar el modelo de recomendación de hoteles similares, es conveniente analizar el *status quo*. Comenzaremos con un análisis de la competencia del que trataremos de sacar información valiosa de qué hay o no hay en el mercado y tomar ideas. Posteriormente haremos un análisis del sistema de Best Day que pretendemos mejorar con el fin de entender en dónde hay áreas de oportunidad.

### 1.1. Análisis de la competencia

En esta sección analizaremos la manera en que la competencia de Best Day recomienda hoteles. Los ámbitos que nos interesan son

1. La posición dentro de la página de las recomendaciones.
2. La forma en la que se eligen las recomendaciones.

Enfatizaremos en los principales competidores, de Best Day, es decir: Booking, Price Travel, Expedia y Despegar. También veremos si podemos aprender algo de TripAdvisor, que aunque no es un competidor directo, vive de recomendar hoteles y de su comunidad de comentarios. Las empresas competidoras están ordenadas de menor a mayor relevancia del sistema de recomendación que tienen.

## Price Travel (pricetravel.com)

Pusimos a Price Travel en primer lugar en la lista porque no tienen un sistema de recomendación de hoteles similares. Cuando se busca un destino, se presentan los hoteles en el Sort Order determinado por su sistema, pero ya que se entra a la página de un hotel en particular (independientemente de cómo se llegó a esa página), no se recomienda hoteles similares ni cercanos. De hecho, la única indicación de otros hoteles es en la pestaña de mapa, en la que se pueden ver todos los cercanos en un mapa interactivo. Dado que no aprendimos nada de este competidor, procederemos al siguiente.

## Booking (booking.com)

Booking es un competidor muy importante de Best Day, aunque también tiene un mercado más grande y mayor cantidad de opciones en destinos fuera de Latinoamérica. Antes de entrar en los puntos de interés general, cabe mencionar que algo que hace booking que no hacen los demás es que cuando se busca un hotel en la página con la barra inteligente, en lugar de ir directamente a su página, se muestra en el primer lugar de una lista, dando la posibilidad de elegir algún otro hotel. Ahora sí pasemos al tema en cuestión:

1. Las recomendaciones se encuentran del lado izquierdo, abajo de la barra de búsqueda y fechas. Es una posición de prominencia media-baja, puesto que no es visible inmediatamente sin bajar en la página.
2. En esta página hay dos tipos de recomendaciones, según las visitas del cliente:
  - **Hoteles populares por destino:** Cuando se entra a la página (no al hotel en particular) por primera vez y se selecciona un hotel, aparece una sección de hoteles populares en el mismo destino del hotel buscado. El tema es que siempre aparece la misma selección para cada destino, independientemente del hotel que se haya elegido, pues la popularidad de los otros hoteles no tiene relación con el hotel elegido.
  - **Búsquedas recientes:** La información de los hoteles visitados se va almacenando en una *cookie*, de modo que para los hoteles subsecuentes, en lugar de los hoteles populares por destino, aparece



un listado de unos pocos hoteles recientemente buscados por el usuario.

Al parecer en Booking sólo están interesados en lo que el cliente ha buscado y utilizan los hoteles populares por destino para no dejar vacía la sección para los clientes nuevos. El problema con los hoteles populares por destino es que no varían según el hotel buscado (siempre y cuando estén en el mismo destino), mientras que el de las búsquedas recientes es que no promueven que el cliente explore la página y encuentre la mejor opción para sus necesidades. Entraremos en detalles más adelante.

## **Despegar (despegar.com)**

Despegar es un competidor más directo y con un mercado más parecido al de Best Day que Booking, aunque está más enfocado en vender boletos de avión que cuartos de hotel. En este caso hay más de una estrategia y en más de un lugar de la página de cada hotel, por lo que las describiremos por separado. Cabe mencionar que no usan todas las estrategias para todos los hoteles, por lo que sospechamos que tienen información incompleta y sólo muestran lo que tienen disponible.

### **■ Búsquedas recientes**

1. Se encuentra en la izquierda abajo del mapa. Igual que en booking.com, se tiene que bajar en la página para verlo, por lo que su relevancia es media-baja.
2. Son los hoteles buscados recientemente por el usuario. Curiosamente, no siempre nos aparecieron al hacer búsquedas, incluso después de haber visto varios hoteles.

### **■ Hoteles cercanos**

1. Se encuentra directamente debajo de las búsquedas recientes.
2. Se listan los hoteles más cercanos al hotel en buscado, sin ningún criterio de precio, categoría o estilo del hotel.

### **■ Hoteles populares por destino**

1. Se encuentra directamente abajo de los hoteles cercanos.

2. Para cada destino muestra una lista de los hoteles más populares (presumiblemente los más vendidos). La lista es la misma para todos los hoteles en un destino dado y de hecho contiene al hotel buscado en su posición correspondiente (es decir ni se omite ni aparece hasta arriba).

■ **Otros usuarios también buscaron...**

1. Aparece en la parte inferior de la página, abajo de prácticamente todo. Su relevancia es muy baja y de hecho se omitió en los primeros análisis de la competencia porque no lo vimos.
2. En la parte inferior de la página también aparece una sección de hoteles que fueron visitados por los demás usuarios que han visto el hotel en cuestión. Este análisis no es complejo, pero puede ser pesado computacionalmente si no se utilizan métodos apropiados de *Market Basket Analysis* como el algoritmo *A Priori*, por ejemplo.

En Despegar tuvieron muchas ideas para recomendar, pero todos los criterios están separados y no colaboran entre sí. Es como tener una tabla ordenada por varios campos distintos. Los hoteles recomendados con base en lo que han visto otros clientes creemos que están desperdiciados en donde están, puesto que sólo los clientes que vayan hasta abajo de la página los verán. El sistema funciona razonablemente bien y permite movilidad hacia arriba en la categoría del hotel. El problema es que es claramente afectado por la popularidad de los hoteles, de modo que tiende a atorarse en aglomerados de hoteles populares y también a subir de categoría. El resultado es que después de algunos pocos *clicks*, casi invariablemente se llega a los hoteles más caros de la zona, sin posibilidad de explorar otros.

## **Expedia (expedia.com)**

En Expedia se le da una prioridad bajísima a las recomendaciones, puesto que están posicionadas en un lugar casi imposible de ver. Sin embargo, tienen un sistema diferente a los demás y por eso lo pusimos en esta posición en la lista de sistemas competidores.

1. Todas las recomendaciones se encuentran en el centro, pero casi hasta abajo, después de toda la información del hotel. La prioridad es muy

baja.

2. Tienen dos tipos de recomendaciones:

- **Búsquedas recientes:** Es igual que en las demás páginas. Aparecen los últimos cuatro hoteles buscados distintos del actual (en caso de repetir una búsqueda).
- **Hoteles recomendados:** Utilizan un método desconocido para recomendar hoteles similares que permite movilidad progresiva entre categorías. No estamos seguros si tiene alguna restricción de precios.

El sistema de recomendación de Expedia nos gustó bastante y de hecho creemos que lo están desaprovechando. Permite explorar los hoteles disponibles con hacia arriba y hacia abajo en la escala de precio y categoría y llegar a opciones completamente nuevas. Es probable que esté acotado por destino, aunque eso no es tan grave. Algo que no nos gustó tanto es que tal vez es algo rígido y después de buscar un poco se puede atorar. A diferencia del algoritmo de Booking, esto no necesariamente lleva a los hoteles más caros de la zona.

## Trip Advisor (tripadvisor.com)

¿Qué hacen (casi) todos? >Hoteles recientes >Barra de búsqueda inteligente (la de Best Day tiene fallas) ¿Qué hacen algunos? >Booking presenta muchos hoteles

### 1.2. Análisis del sistema actual

### 1.3. Diagnóstico

### 1.4. Wish List

## Capítulo 2

### Modelo propuesto

- 2.1. Puntos a resolver
- 2.2. Distancia
- 2.3. Precio
- 2.4. Similitud de hoteles
- 2.5. Distancia revisitada
- 2.6. Integrando todo

## Capítulo 3

### Obtención de la información

- 3.1. Formato necesario para el modelo
- 3.2. Calidad de la información
- 3.3. Sistema de respaldo

## Capítulo 4

### Desempeño