

Sistema de recomendación de hoteles similares

TESIS

Felipe Gerard Valdés



INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO

Febrero de 2016

Agenda

1 Planteamiento

2 Solución

3 Implementación

4 Desempeño

Agenda

1 Planteamiento

2 Solución

3 Implementación

4 Desempeño

- Modelo implementado en **Best Day** (agencia de viajes online)
- **Situación inicial:**
 - El sistema de recomendación es estático
 - Criterio: Destino + Estrellas
 - Los clientes potenciales no encuentran el hotel que buscan fácilmente
 - Por lo tanto, pueden abandonar la página o no encontrar la mejor opción para ellos
 - Esto impacta los KPIs de la empresa
- **Solución propuesta:** Recomendar con un criterio dinámico que tome en cuenta las principales características del hotel seleccionado

Análisis de la competencia

- Competidores principales: Price Travel, Booking, Despegar, Expedia, Trip Advisor
- Diversas formas de recomendar:
 - Hoteles recientemente vistos: No son recomendaciones genuinas
 - Hoteles cercanos / en el mismo destino: Estilo y precios pueden ser muy distintos
 - Recomendaciones con métodos desconocidos (p.ej. “a otros usuarios les gustó...”): Calidad variable; en general son demasiado rígidos
- **Conclusión:** La competencia deja mucho que desear, pero hay mucho qué aprender

Consideraciones

- Evitar mostrar hoteles demasiado caros
- Tomar en cuenta el perfil del hotel
- Incorporar información geográfica más detallada que el destino
- Diseñar un criterio que se adapte a zonas con diferentes densidades de hoteles
- Evitar que el sistema se estanque en las mismas recomendaciones (efecto laberinto)

Agenda

1 Planteamiento

2 Solución

3 Implementación

4 Desempeño

Plan de ataque

- 1 **Precio:** Filtrar hoteles demasiado caros
- 2 **Similitud:** Construir un *criterio integral de similitud*
 - Cantidad de servicios
 - Perfil similar
- 3 **Distancia:** Recomendar hoteles geográficamente cercanos

- Poner hoteles demasiado caros no promueve el interés del cliente
 - Generalmente más barato es mejor (siempre y cuando tenga la categoría suficiente)
- Porcentaje máximo arriba del hotel original
- Hoteles por encima del precio establecido son invisibles

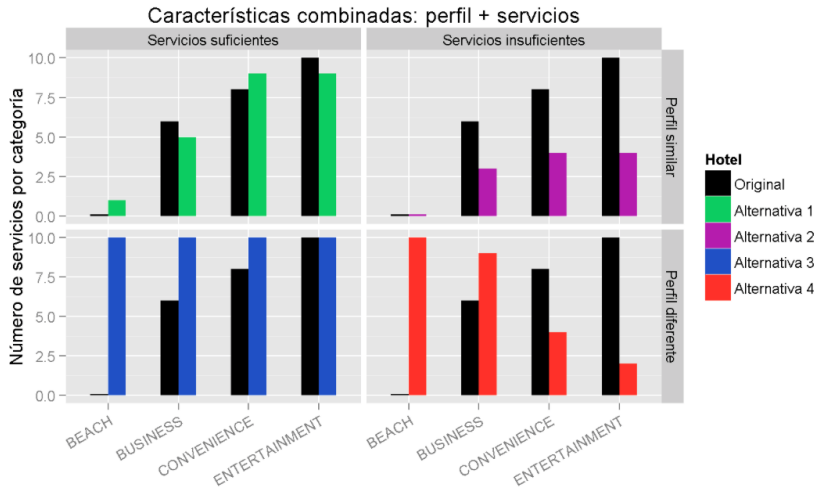
■ Caracterización

- No bastan las estrellas
- Usar información de servicios
- Usarla directamente es ruidoso porque no toma en cuenta sustitutos
- **Solución:** Agrupar los servicios en categorías
- **Método:** Manual (criterio de negocio) + aglomerado jerárquico

Similitud (P2)

- La similitud se puede dividir en dos aspectos básicos
- **Servicios:**
 - Al menos ciertos servicios (estrellas)
 - Los servicios faltantes restan similitud
 - Los servicios extras no se penalizan
- **Perfil:**
 - Hoteles del mismo estilo
 - Misma proporción de cada categoría
 - Ignorar la cantidad de servicios
 - Podría dar hoteles más sencillos o más lujosos

Similitud (P2)

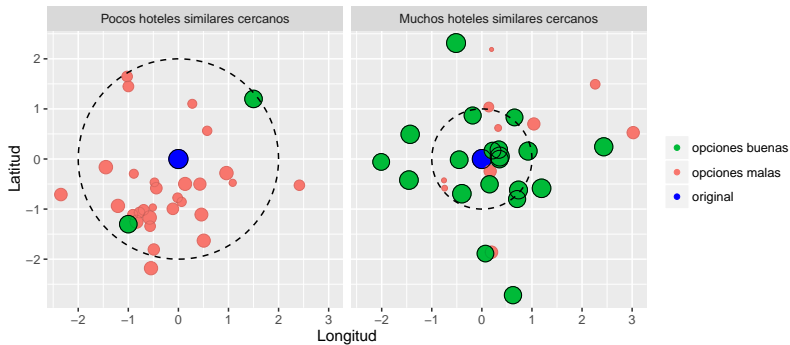


Similitud (P3)

- Medida de similitud de servicios: Calidad mínima
- Medida de similitud de perfil: Proporciones de servicios
- Medida de similitud final:
 - Capturar ambos efectos
 - Promedio ponderado entre ambas medidas
 - Ponderación de acuerdo a la variabilidad de cada una (PCA)
- ¿Por qué no diferencia absoluta?
 - Es muy agresiva
 - No es ajustable

- Hoteles cercanos (usando coordenadas)
- El significado de “cerca” es variable
- Se necesita un criterio dinámico
- Cuidar que los hoteles distintos no afecten

Distancia



Modelo completo

- 1 Agrupar los servicios en categorías
- 2 Calcular la ponderación entre servicios y perfil
- 3 Definir una **geocerca** (radio de cercanía razonable) alrededor de cada hotel
 - Ignorar hoteles caros (precio de largo plazo)
 - Acumular una cantidad de similitud (mayor peso a hoteles parecidos, menor a los diferentes)
- 4 Recomendar dentro de la geocerca de acuerdo a la **similitud**

Agenda

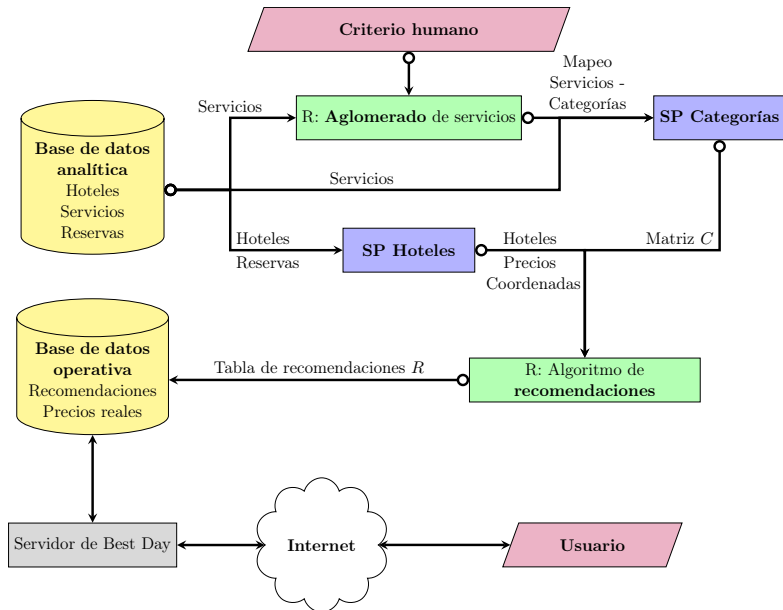
1 Planteamiento

2 Solución

3 Implementación

4 Desempeño

Arquitectura: SQL + R



Agenda

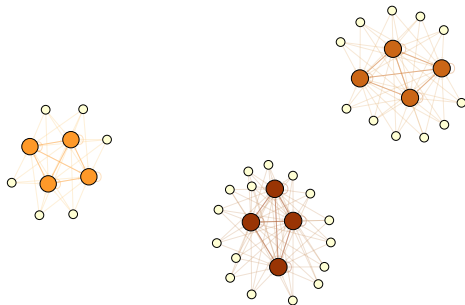
1 Planteamiento

2 Solución

3 Implementación

4 Desempeño

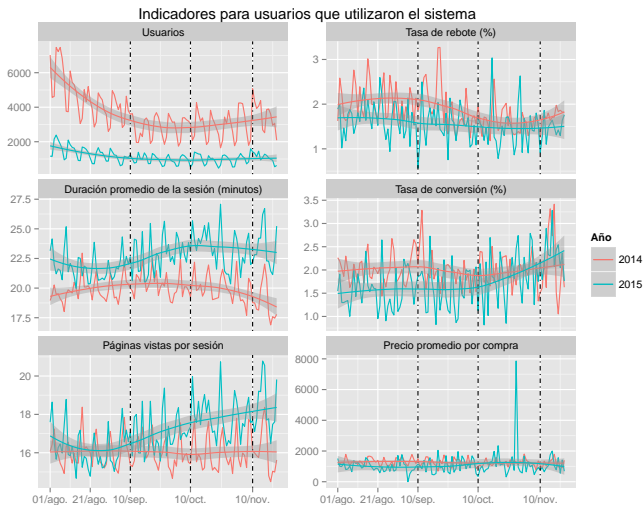
Desempeño teórico (antes)



Desempeño teórico (después)



Desempeño de la página



■ Teóricas:

- Mejores cualidades: **precio, distancia, similitud**
- La red de recomendaciones está mejor conectada
- Mayor **flexibilidad**: se puede cambiar de categoría gradualmente

■ Prácticas:

- Impacto importante en los clientes que empezaron a utilizar las recomendaciones
- **Mayor uso** una vez que se utiliza el sistema
- Con más tiempo la **conversión** empezó a aumentar
- No hay aumento de entradas (diseño y posición)

■ Trabajo futuro:

- Precio dinámico, filtro de precio en forma de “U”
- A/B testing
- Mayor periodo de observación