Patrón de fondo

Descripción generada automáticamente con confianza media

Universidad Austral - Maestría en Ciencia de Datos

Caso #1

**Autor: Andres Montes de Oca**

Fecha: XX de Noviembre de 2022

Introduccion al

Data Mining

Explicación de los pasos del proceso de KDD, comparándolo con el ejemplo de la Cadena de Hoteles

1. Comprensión del dominio de negocio: En este paso comprendemos como es que funciona nuestro negocio, como generamos ingresos. Al ser una Cadena Hotelera, nuestro negocio se debería basar en conocer bien los gustos de los huéspedes, para ofrecerles servicios o actividades durante su estadía, pero así también para ofrecerles sus próximas vacaciones deseadas.
2. Selección de los Datos: En este paso seleccionamos las diversas fuentes de los distintos datos que vamos a tener como input. En nuestro ejemplo seria el cuestionario inicial que le hacemos a nuestros huéspedes al entrar, donde debería preguntársele cuales son sus gustos. Por otro lado el log de la tarjetas de servicios, con lo cual sabemos cuales son sus gustos de consumo del restaurant y del bar dentro del hotel, como así también cuales son los ammenities que mas le agradan. Por ultimo también tenemos el cuestionario de salida.
3. Pre-procesamiento de los Datos: Aquí se limpian un poco los datos, se completean/eliminan los nulos o duplicados. Se eliminan inconsistencias reduciendo el ruido en los datos, como así también los outliers. Aquí la verdad es que no veo como relacionarlo con el ejemplo, este paso me suena mas a algo muy técnico.
4. Minería de datos: Aquí es cuando analizamos los datos en busca de patrones y relaciones entre ellos para poder generar conocimiento. Al comprarlo con el ejemplo, para poder decidir sobre que publicidad enviarle en su canal privado de TV durante los primeros días de su estadía, utilizaríamos la data recolectada en el formulario de ingreso (suponiendo que es un husped nuevo que no conocemos). Luego, cuando ya tengamos mas data recolectada por la tarjeta de consumo de servicios y ammenities dentro del hotel, actualizamos las publicidades enviadas con las conclusiones que saquemos de estos nuevos datos. Otra comparación con el ejemplo seria la clusterizacion de mis clientes, teniendo en cuenta sus consumos internos, para la generación del mailing de marketing para las próximas vacaciones.
5. Interpretación y Evaluación de los modelos. Aquí es donde sacamos conclusiones sobre los resultados de como performaron nuestros modelos publicitarios durante la estadía del nuesped, como así también el éxito de nuestro mailing anual para las próximas vacaciones.