

Business Intelligence

- Reto 2 - Sistemas de Recomendación -

1 Objetivos y Entregables

1.1 Objetivo General

Desarrollar un sistema que da recomendaciones efectivas que hacen que el cliente compra más (y de más valor) en comparación con las compras que realiza sin la presencia del sistema.

El sistema desarrollado debe obedecer la especificación de la clase **Recommender** que viene al lado de esta tarea.

Entregables:

1. la entrega en gradescope (.zip o .py)
2. un PDF de hasta 2 páginas que explica la solución y justificando las decisiones técnicas que tomaron para su implementación.

2 Avisos Técnicos

Deben subir sus soluciones en la tarea “Reto 2” en el correspondiente curso en gradescope: <https://www.gradescope.com/courses/662684>. El código de acceso es: 7DZ7NZ

El proceso para hacer entregas dentro del sistema se explica solo. El nombre que se coloca para la entrega se usa únicamente para el leaderboard, así que allí pueden usar un nombre ficticio que quieren que aparezca.

Para calificar una entrega, el sistema hace lo siguiente:

1. en el caso de que se entregó un archivo **.zip**, se desempaca el archivo y se busca si existe un archivo **requirements.txt**. De ser así, se instalan los paquetes definidos en ese archivo según los estándares de **pip**.

Después se carga el archivo **recommender.py** que tiene que obedecer el formato especificado en el archivo compartido. El sistema ahora crea un objeto de la clase **Recommender**.

2. llama **train** con la base de datos. Este proceso tiene que terminar dentro de un minuto máximo. Sin eso, la entrega no se vale.
3. en una secuencia, llama **get_recommendations** para cada uno de una serie de carritos de compras. La función tiene que realizar las recomendaciones dentro de 100ms para ser usado. Una recomendación es una lista de números de items.

El sistema presenta las recomendaciones al cliente ficticio, quién decide si va a agregar uno o más de los items recomendados a su carrito.

4. Al final, se calcula el aumento en ganancias generado por las recomendaciones y se lo usa para dar una puntuación.

3 Rúbrica de Evaluación

Se valdrá la última solución entregada en el momento del cierre del sistema (22 de noviembre, medianoche). Para la calificación del trabajo, se usará una base de datos *nueva*. Así que es importante que no sobreajusten su solución para la base presentada durante del periodo de desarrollo. Se valora la mejor entrega por grupo (en caso de que entreguen varios).

Algunas observaciones:

1. la rúbrica es monótona. Es decir, para conseguir un 100% uno necesita todo lo que se requiere para 80% más lo que es particular para 100%. Y lo mismo para las otras categorías.
2. afirmaciones como “la solución evidencia una profunda comprensión de los temas visto en clase y los explota bien” no implican que hay que usar soluciones visto en clase. Es posible argumentar por qué no se usó y por qué la solución elegida parece más apropiada.

Concepto	100%	80%	50%	20%
Código original y limpio	el código está conciso y fácil de entender para un tercero.	no se usa bibliotecas externas otras que numpy, pandas, y scipy.	100% del código entregado es trabajo propio y no generado o copiado	al menos 50% del código entregado es trabajo propio y no generado o copiado
Solución original y coherente	la solución tiene un aspecto innovador que sugiere posible superioridad a soluciones más básicas	la solución evidencia una profunda comprensión de los temas visto en clase y los explota bien	la solución se justifica de manera razonable y coherente	la solución es propia del grupo y justificada de manera mínima
Rendimiento del Sistema	comparable o mejor a una solución que usa de manera óptima las técnicas de itemset mining visto en clase	comparable con una solución que usa de manera avanzada las técnicas de itemset mining visto en clase	comparable con una solución que usa técnicas básicas de itemset mining	el sistema es funcional y no produce errores. Mejor puntaje que la solución trivial del archivo original.
Presentación	hacen una interpretación elocuente de los resultados e pueden argumentar limitaciones de su análisis y como se podría mejorarlo	todas las decisiones y afirmaciones están bien fundamentadas (el <i>por qué</i>)	en la presentación y las preguntas se evidencia que los integrantes comprenden su propia solución (el <i>cómo</i>)	la presentación respeta el tiempo y evidencia lo que se ha hecho