

Inteligencia Artificial y Aprendizaje Automático  
**Actividad Semana 10: Reducción de Dimensionalidad (SVD)**

Maestría en Inteligencia Artificial Aplicada  
Prof. Luis Eduardo Falcón Morales

Tecnológico de Monterrey

Nombre(s): \_\_\_\_\_ Matrícula(s): \_\_\_\_\_

Esta Tarea se deberá resolver en equipos, de acuerdo a como fueron integrados al inicio del curso.

Se trabajará con el conjunto de datos llamado "Restaurant & consumer", para la generación de un sistema de recomendación. Estos sistemas están dentro de los llamados métodos no supervisados y que se basan en técnicas de reducción de dimensionalidad.

El archivo a descargar se llama **RCdata.zip** y se encuentra en la siguiente liga del repositorio de la UCI:

<https://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data>

En particular se trabajará con los siguientes archivos, los cuales contienen la información mínima que requerimos para esta actividad: **rating\_final.csv** y **geoplaces2.csv**.

El objetivo es diseñar tres sistemas de recomendación basados en la evaluación de los usuarios a los restaurantes. Dicha evaluación será la general, la referente al servicio y la referente a la comida.

1. Descarga los archivos y genera la matriz de utilidad en la cual los renglones son los usuarios (user\_ID), las columnas los nombres de los restaurantes (name) y las entradas la evaluaciones generales dadas por los usuarios (rating). En particular indica el tamaño de dicha matriz y despliega sus primeros 10 renglones.
2. Con el objetivo de obtener la información con respecto a los vectores latentes de los restaurantes, obtener dicha factorización SVD utilizando todas las componentes posibles, es decir, usando todos sus valores singulares. En particular indica la cantidad de valores singulares obtenidos.
3. Usando la métrica de correlación de Pearson, obtener las 10 mejores recomendaciones de este modelo no supervisado con base a la información de alguien que evaluó bien de manera general al "Subway". Despliega también el valor numérico de la correlación de Pearson de cada una de estas 10 recomendaciones. NOTA: No debe estar incluido "Subway" dentro de las recomendaciones.
4. Disminuye ahora la dimensión de la factorización con un número de componentes tal, que las recomendaciones con esta menor cantidad de información coincidan en al menos 9 de las 10 recomendaciones dadas con todos los valores singulares. Indica la cantidad de valores singulares encontrado y la cantidad de información que contienen.

5. Repite el ejercicio con el restaurante ""tacos de barbacoa enfrente del Tec". Incluye tus conclusiones.
6. Repite los ejercicios 1 a 4 generando ahora la matriz de utilidad en la cual los renglones son los usuarios (user\_ID), las columnas los nombres de los restaurantes (name) y las entradas la evaluación dada a la comida (food\_rating). Incluye tus conclusiones.
7. Repite los ejercicios 1 a 4 generando ahora la matriz de utilidad en la cual los renglones son los usuarios (user\_ID), las columnas los nombres de los restaurantes (name) y las entradas la evaluación dada al servicio (service\_rating). Incluye tus conclusiones.
8. Incluye tus conclusiones finales. En particular indica la coincidencia o no de estos tres modelos.