

# Tarea de aula: SVD sistema de recomendación

Métodos numéricos

## SVD - Visualización de Sistema de recomendación

Se proporcionará un conjunto de datos que consiste en calificaciones de películas dadas por un grupo de usuarios. Este conjunto de datos puede ser una matriz de calificaciones donde las filas representan a los usuarios y las columnas representan las películas. Las entradas de esta matriz son las calificaciones dadas por cada usuario a las películas que ha visto. El objetivo es utilizar SVD para descomponer esta matriz en componentes que revelen patrones latentes relacionados con preferencias de películas.

### Pasos de la Tarea:

- **Exploración de Datos:** Inspeccionar el conjunto de datos para entender su estructura, como el número de usuarios, el número de películas, y la distribución de las calificaciones.
- **Preparación de Datos:** Preparar los datos para la descomposición, tratando los valores faltantes adecuadamente, pues es común que no todos los usuarios hayan calificado todas las películas. (Cuidado con trabajar con todos los datos)
- **Aplicación de SVD:** Utilizar SVD para descomponer la matriz de calificaciones en tres matrices:  $U$ ,  $\Sigma$  y  $VT$ , donde  $U$  representa las características latentes de los usuarios,  $\Sigma$  es una matriz diagonal con los valores singulares (que representan la importancia de cada característica latente), y  $VT$  representa las características latentes de las películas.

- **Interpretación de Resultados:** Analizar las matrices resultantes para interpretar los patrones latentes.
- **Realiza una reducción de dimensionalidad:** Aporta el valor de información mostrada ó información perdida y representa a los usuarios y los productos (películas) en un espacio tridimensional para interpretar su similitud (cercanía espacial)

### Rúbrica

#### 1. Exploración de Datos (20 puntos)

- **Insuficiente (0-5 puntos):** El estudiante muestra poco o ningún esfuerzo en explorar el conjunto de datos. Falta información básica como el número de usuarios, películas, y la distribución general de las calificaciones.
- **Básico (6-10 puntos):** El estudiante identifica el número de usuarios y películas, pero la exploración de la distribución de calificaciones es superficial.
- **Satisfactorio (11-15 puntos):** El estudiante realiza una exploración adecuada del conjunto de datos, identificando el número de usuarios, películas, y realiza un análisis básico de la distribución de las calificaciones.
- **Excelente (16-20 puntos):** El estudiante proporciona un análisis detallado del conjunto de datos, incluyendo estadísticas descriptivas completas y visualizaciones que aportan insights claros sobre la estructura y distribución de las calificaciones.

#### 2. Aplicación de SVD (20 puntos)

- **Insuficiente (0-5 puntos):** La descomposición SVD es incorrecta o incompleta. Falta una comprensión fundamental de cómo aplicar SVD al conjunto de datos.
- **Básico (6-10 puntos):** SVD se aplica correctamente, pero hay poco o ningún detalle sobre el proceso o la interpretación de las matrices resultantes.
- **Satisfactorio (11-15 puntos):** La descomposición SVD se realiza correctamente, con una explicación adecuada del proceso y una interpretación básica de las matrices  $U$ ,  $\Sigma$ , y  $V^T$ .
- **Excelente (16-20 puntos):** La aplicación de SVD es técnicamente sólida, con una explicación detallada del proceso y una interpretación profunda de las matrices resultantes, destacando insights relevantes.

#### 3. Interpretación de Resultados (20 puntos)

- **Insuficiente (0-5 puntos):** La interpretación de los resultados es inadecuada o incorrecta. Falta comprensión de los patrones latentes.
- **Básico (6-10 puntos):** Se ofrece una interpretación básica de los resultados, pero falta profundidad o detalle.
- **Satisfactorio (11-15 puntos):** Los resultados se interpretan correctamente, identificando algunos patrones latentes y proporcionando una explicación razonable.

- **Excelente (16-20 puntos):** Se proporciona una interpretación detallada y bien razonada de los resultados, destacando patrones latentes significativos y ofreciendo insights valiosos sobre las preferencias de los usuarios y las características de las películas.

#### 4. Reducción de Dimensionalidad y Visualización (20 puntos)

- **Insuficiente (0-5 puntos):** La reducción de dimensionalidad es incorrecta o no se realiza. Las visualizaciones son inadecuadas o no se proporcionan.
- **Básico (6-10 puntos):** Se realiza una reducción de dimensionalidad básica sin una explicación clara o justificación. Las visualizaciones son básicas y ofrecen poco valor adicional.
- **Satisfactorio (11-15 puntos):** La reducción de dimensionalidad se realiza correctamente,

#### Recursos:

- Archivo csv con los datos con los ratings de películas
- Plantilla código Octave