**Dominio escogido: Juegos de PS4**

1. **Recopilación de ítems:**

Se usó la librería BeautifulSoup, la cual resultó ser muy ineficiente para el Crawling, ya que tiene tiempos muy altos de ejecución al recorrer todas las páginas, pero, fue la más accesible de usar por su poca complejidad a la hora de programar. Las dos páginas usadas fueron cambiadas a las propuestas en el taller con aprobación del profesor.

<https://www.bestbuy.com/site/playstation-4-ps4/playstation-4-ps4-video-games/pcmcat296300050018.c?id=pcmcat296300050018&intl=nosplash>

<https://www.flipkart.com/q/ps4-games>

Adicionalmente, para la comparación de los títulos se usó la librería FuzzyWuzzy

conda install -c conda-forge fuzzywuzzy

1. **Control de usuarios:**

El control de usuarios se hace con la base de datos sqlite, importando la librería sqlite3 de python apoyado de interfaces gráficas de wxPython, para usarlas se debe instalar la librería wxPython con el siguiente comando en el prompt de anaconda como administrador

pip install -U wxPython

usando las funciones definidas en el archivo Funciones.py dentro de la carpeta principal.

1. **Explorador de ítems y registro de calificación:**

El explorador de ítems y los registro y cambios se hace igual que en el punto anterior,se hace con la base de datos sqlite, importando la librería sqlite3 de python apoyado de interfaces gráficas de wxPython, usando las funciones definidas en el archivo Funciones.py dentro de la carpeta principal.

1. **Recomendación:** Para probar el sistema el sistema de recomendación se hace necesario instalar las librerías Surprise de python en el prompt de conda como administrador con el comando: $ conda install -c conda-forge scikit-surprise

Se desarrollaron dos sistemas de recomendación; un sistema de recomendación basado en contenido y otro de tipo colaborativo.

El sistema de recomendación basado en contenido toma en consideración tres características de los juegos, dos categóricas, ‘game mode’ y ‘publisher’, y como variable numérica toma precio. Ambas características categóricas se estandarizan desde el web scraping, para luego ser convertidas en características discretas con la función ‘get\_dummies’ de pandas; por otro lado como una de las páginas es de Estados Unidos y la otra de India, se estandarizan los precios llevándolos todos a dólar. Una vez se disponen todos los valores numéricos en la tabla se procede a calcular el perfil del usuario, el cual será mostrado al usuario por medio de la interfaz gráfica para que conozca sus preferencias; luego se calcula el perfil normalizado de usuario, con el cual se calculará la similitud coseno con cada uno de los juegos que el usuario no ha calificado, dicho valor es normalizado con respecto a las notas, para así obtener una nota estimada con la que el usuario calificaría dicho juego si lo juega. Los juegos con sus respectivos puntajes se almacenan en una lista la cual se almacena de manera descendente, y así conocemos los juegos que más se adecúan al perfil del usuario.

El sistema de recomendación colaborativo es de tipo usuario-usuario, se realizó de esta manera porque para este caso de estudio en concreto se tendrán más juegos que usuarios, por lo cual resulta más simple computacionalmente hablando realizarlo de esta manera. Este sistema de recomendación se realiza aplicando la técnica SVD de factorización de matrices, y al igual que el sistema de recomendación anterior, devuelve los juegos que más le puedan interesar al usuario.

Finalmente para el sistema híbrido, se usaron los dos sistemas de recomendación previamente explicados para trabajar en conjunto, y asignándole un peso de 0.4 y 0.6 al sistema de recomendación por contenido y al colaborativo respectivamente, la asignación de los pesos se hacen de esta manera debido a que se considera un poco más relevante la recomendación del sistema de recomendación colaborativo ya que el recomendador por contenido puede ser un poco sesgado con los resultados debido a las condiciones de nuestras variables categóricas.

1. **Cold Start Problem:**

Para el Cold Start problem de usuarios, se toman los 10 elementos con la nota ponderada del sistema más alta y se muestran al nuevo usuario.

Para el Cold Start problem de ítems se le muestran al usuario los juegos añadidos más recientemente al sistema, de manera que este pueda calificarlos si ya los ha jugado.