

UNIVERSIDAD AUTONOMA DE NUEVO LEON

Facultad de ciencias físico matemáticas

Minería de datos

Equipo: 7

Almaguer Alanis Eduardo.....1741322

Vega Macias Oscar Saul1626997

Nasser Benavides Juan Pablo1753367

Grupo: 003

Frecuencia: viernes 6pm a 9pm

30/10/2020

Título de la base de datos

Nombre: Aplicaciones de Google Play Store

URL de la base: <https://www.kaggle.com/lava18/google-play-store-apps>

Descripción de los datos

Si bien muchos conjuntos de datos públicos (en Kaggle y similares) proporcionan datos de la App Store de Apple, no hay muchos conjuntos de datos equivalentes disponibles para las aplicaciones de Google Play Store en la web. Al profundizar, descubrí que la página de iTunes App Store implementa una estructura similar a un apéndice bien indexada para permitir un raspado web simple y fácil. Por otro lado, Google Play Store utiliza sofisticadas técnicas modernas (como la carga dinámica de páginas) utilizando JQuery, lo que hace que el scraping sea más desafiante.

Contenido

Cada aplicación (fila) tiene valores de categoría, clasificación, tamaño y más.

Este documento incluye 2 archivos tipo (.csv) y un tipo (.txt)

- Detalles de las aplicaciones de Google Play (googleplaystore.csv)
 - Aplicación
 - Categoría
 - Clasificación
 - Reseñas
 - Talla
 - Instala
 - Tipo
 - Precio
 - Calificación de contenido
 - Genero
 - Ultima actualización
 - Versión actual
 - Versión de Android
- Contiene las primeras 100 reseñas “mas relevantes” de cada aplicación (googleplaystore_user_reviews.csv)
 - Aplicación
 - Translated_review
 - Sentimiento
 - Sentiment_Polarity
 - Sentiment_subjectivity
- (license.txt)
 - Este trabajo tiene la licencia Creative Commons Attribution 3.0 Unported License. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by/3.0/>.

Justificación del uso de datos

Es un hecho que los padres le proporcionan el celular a sus hijos para su entretenimiento, por lo que vemos necesario y decidimos crear un apartado especial para niños dentro de la misma aplicación de Google Play Store para que sea un apartado de contenido apropiado y educativo para ellos, ya que al acceder a Google Play Store muestra todo el contenido sin distinción de edad.

Esta base de datos cuenta con lo necesario para llevar a cabo nuestro estudio como, por ejemplo:

- Categoría a la que pertenece la aplicación
- Grupo de edad al que está dirigida la aplicación:
 - Niños
 - Mayores de 21 años
 - Adultos

Planteamiento del problema

El brindar el servicio del celular a los niños sin la observación de un adulto nos deja vulnerables al hecho de no saber que aplicaciones están descargando y por ende el contenido de la misma, lo cual, lleva a que un menor tenga acceso a información no adecuada según su rango de edad.

Objetivo final

Crear un apartado especial para niños (Google Play Store Kids) dentro de la misma aplicación (Google Play Store) con la finalidad de limitar el contenido no apropiado para los niños.

Con los beneficios de:

- a) Tener una mayor confianza al brindar el servicio del celular a un menor.
- b) Descargar a la aplicación deseada más rápidamente.

Planeación de la herramienta a utilizar

1- Clasificación: Como nosotros sabemos, la clasificación ordena por clases tomando en cuenta las características de los elementos que contiene, por lo que nuestra principal herramienta será la misma, donde nuestro objetivo principal será clasificar las aplicaciones utilizando los datos (googleplaystore.csv), donde utilizaremos las columnas:

- Categoría a la que pertenece la aplicación:
 - Familia ✓
 - Juego ✓
 - Herramientas
 - Medico
 - Negocio
 - Educación..... ✓
 - Arte y diseño ✓
- Grupo de edad al que está dirigida la aplicación:
 - Niños ✓
 - Mayores de 21 años
 - Adultos