

Enunciado del proyecto final

Una empresa de publicidad llamada “the neighborhood” quiere hacer un estudio en redes sociales, para medir que tan efectivo sería un tipo de publicidad sobre una muestra de usuarios. Con este fin “the neighborhood” nos llama a nosotros para realizar este estudio de manera completa, eficiente y dinámica. Para la empresa no es suficiente entregar un reporte con los datos obtenidos, a ellos les gustaría obtener una herramienta que les permita realizar diferentes hipótesis y comprobarlas.

En la primera entrega de la aplicación se trabajará sobre Twitter, the neighborhood quiere obtener una muestra de usuarios y realizar lo siguiente:

*Poder añadir un usuario a la herramienta de estudio, con todas las características básicas que posee un perfil de Twitter normal como: -Numero de tweets -Numero de seguidores -Numero de seguidos -Numero de likes y una lista con sus tweets.

*Clasificar los usuarios en 3 categorías deportes, política y ciencia. Esta clasificación estará dada por un glosario de cada categoría, brindados por la compañía, que permitirá dar un puntaje a cada usuario en cada una de las categorías. El puntaje en una categoría depende de la cantidad de palabras que tengan en sus tweets que se encuentren en el glosario de dicha categoría.

*Conectar usuarios entre ellos a partir de la cantidad hashtags que tengan en común y la cantidad de @ que tengan en común. Además de eso, si existe una conexión esta también contendrá la afinidad entre las categorías de los 2 usuarios, la afinidad se determina como el valor absoluto de la resta entre las puntuaciones de los 2 usuarios.

Todo lo anterior con el fin de analizar posibles tendencias entre cierto tipo de usuarios entre las operaciones de investigación que le gustaría hacer a la compañía están:

*Estando en un usuario en particular a cuantos grados de libertad se encuentra:

- Los usuarios que superen cierto puntaje de la categoría de deportes
- Los usuarios que superen cierto puntaje de la categoría de política
- Los usuarios que supere cierto puntaje de la categoría de ciencia

* Empezando en un usuario en particular cual es el grupo de personar más grande que se puede formar que contenga el mayor numero de hashtags en común, o el mayor numero de @ en común

*A partir de dos usuarios determinar cual es el grupo de usuarios que poseen mayor afinidad en cualquiera de las tres categorías.

*Determinar cual es la persona más probable en tener hashtags o arrobas en común con un usuario en específico. Basados en la cantidad de hashtags y arrobas en común, y en la afinidad en las tres categorías entre los usuarios.