**Procesamiento Natural del Lenguaje (NPL)**

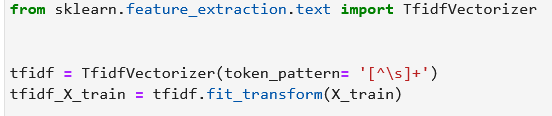
Actividad 3

**Contexto**

Al día siguiente, presentamos los resultados anteriores al equipo y obtenemos el VoBo para continuar. Efectivamente, lo obtenido hasta el momento hace sentido, y lo importante es que ya se cuenta con un corpus de comentarios limpio y se tiene un vocabulario resumido. Alguien del equipo expresa que ya se puede aplicar, entonces, un modelo para clasificación, por lo que -con todo la razón- argumentamos que primero se debe obtener una representación vectorial del corpus, y justamente es lo que se pondrá en foco a continuación.

**Consignas**

Representación vectorial: en esta parte, se debe aplicar un modelo de representación vectorial. Se recomienda el uso de TfidfVectorizer por sobre una representación BoW debido a que facilita la inclusión de los emojis, tal como se expone en el siguiente ejemplo:



El modelo para la vectorización debe ajustarse primero con los datos X de entrenamiento y, posteriormente, realizar la transformación con los datos X de testeo:



**Formato de entrega**

****

Se debe entregar esta parte en formato *notebook* de Jupyter. Se escribe el código de cada parte solicitada en una celda “code”, y se mantienen las salidas dentro del mismo documento. Los comentarios y respuestas textuales se las puede escribir en una celda tipo “markdown”.

**Cómo cargar la actividad**

**¡Felicitaciones por llegar al final de esta actividad!**

Para concluir tu entrega sigue estos pasos:

1. Haz clic en el botón “Enviar tareas” para comenzar el proceso de carga de tu archivo con las respuestas.

2. Adjunta el archivo que contiene tus respuestas a través de la interfaz proporcionada.

3. Una vez que hayas adjuntado el archivo, confirma tu entrega haciendo clic en el botón correspondiente.

Recuerda completar todos los puntos y tener cuenta las condiciones de entrega.

**¡Éxitos!**