

# Prueba de conocimientos

Cargo: Data Scientist / ML Engineer

La intención general de la presente prueba es medir las capacidades técnicas de los candidatos para la elaboración de soluciones end-to-end. Le recomendamos prestar mucha atención a las sugerencias e indicaciones en general.

Instrucciones:

Se requiere que diseñe un algoritmo que permita realizar un análisis de sentimientos sobre los comentarios realizados por usuarios en una aplicación móvil, para luego disponibilizar su solución en un endpoint (API). Sabemos que existen diversos lenguajes de programación con los cuales se puede desarrollar la solución, pero le agradecemos limitar la solución a python3.

Secuencia de acciones

1. Inicie un repositorio en github para en adelante llevar el proyecto.
2. Acceda a los datos de los comentarios disponibles en la siguiente url y guarde los datos en la carpeta del proyecto  
[https://play.google.com/store/apps/details?id=com.clarocolombia.miclaro&hl=es\\_CO&gl=US&showAllReviews=true](https://play.google.com/store/apps/details?id=com.clarocolombia.miclaro&hl=es_CO&gl=US&showAllReviews=true)  
Plus: utilice técnicas de web scraping para automatizar el proceso.
3. Desarrolle un algoritmo que permita determinar la connotación positiva o negativa para cada uno de los comentarios extraídos de la url.  
Sugerencia: Cree funciones de apoyo a manera de pipeline, para permitir reutilizar su algoritmo más adelante.
4. Desarrolle un modelo sencillo para predecir el número de estrellas (calificación de 1 a 5) que pondría un usuario a partir del texto expuesto en el comentario  
Sugerencia: Use la información que se obtiene al aplicar el algoritmo desarrollado en el punto 3.
5. Cree una API con la que se pueda acceder a los resultados del punto 3; el resultado del algoritmo para los n-comentarios más recientes, donde n es un valor dado por el usuario de la API a modo de "Query parameters"  
Sugerencia: Utilice fastAPI, sabemos que puede desarrollarse en flask o django, pero con fastAPI lo podrá hacer muy rápidamente siguiendo los pasos disponibles en la documentación.
6. Cree un nuevo endpoint en la API que permita enviar un texto (puede ser contenido en un json) y cuya respuesta contenga al menos:
  - a. El número de estrellas resultante al aplicar el modelo desarrollado en el punto 4.

- b. La connotación positiva/negativa del comentario enviado, utilizando el modelo desarrollado en el punto 5.
- 7. (Opcional) Despliegue la API en un servidor cloud.  
Sugerencia: Si utilizó fastAPI para desarrollar la API lo natural es hacer el deploy en Heroku
- 8. Redacte una guía de uso rápido de su solución en el fichero readme.md en github  
Sugerencia: Incluya el enlace de acceso a la API en caso de haber realizado el paso 7.
- 9. Envíe enlace al repositorio github al correo [camilo.aguilar@easydata.com.co](mailto:camilo.aguilar@easydata.com.co) con el asunto: Prueba Data Scientist NOMBRE-APELLIDO, respetando el plazo que previamente se le ha indicado.

Sugerencias finales: Establezca una estrategia para llevar a cabo el proyecto evitando concentrar demasiado tiempo en una sola tarea, recuerde que se considera muy valioso el lograr la solución end-to-end.

¡Éxitos!