

**INCLUIR ESTE PÁRRAFO EN EL NOTEBOOK QUE ENTREGA COMO SOLUCIÓN AL PARCIAL:**

“Al entregar la solución de este parcial, yo, **[nombre y apellido]** con código **[#]** me comprometo a no conversar durante el desarrollo de este examen con ninguna persona que no sea el profesor del curso, sobre aspectos relacionados con el parcial; tampoco utilizaré algún medio de comunicación por voz, texto o intercambio de archivos, para consultar o compartir con otros, información sobre el tema del parcial. Soy consciente y acepto las consecuencias que acarrearé para mi desempeño académico cometer fraude en este parcial”.

## **Inmobiliaria Los Alpes**

Inmobiliaria Los Alpes es una inmobiliaria con varios años de trayectoria la cual se especializa en el arrendamiento de inmuebles de corta estancia. La empresa ha ido creciendo poco a poco y en la actualidad enfrenta una serie de retos relacionados con la gran cantidad de competencia que hay en el sector, dificultando la obtención de nuevos inmuebles para su administración como arrendador, así como conseguir los arrendatarios idóneos para el uso de los mismos.

Como estrategia para mantener su competitividad en el mercado, la Inmobiliaria Los Alpes ha decidido enfocarse en determinar los inmuebles que podrían llegar a ser populares desde antes de su publicación. La popularidad de un inmueble está dada en términos de la cantidad de comentarios (*reviews*) que ha obtenido independientemente de su calificación. Esta actividad es importante dado que los inmuebles más populares son los que generan mayor rotación, mientras que para los de menor popularidad hay que hacer una mayor inversión en marketing para promocionarlos.

La inmobiliaria busca contratar con usted un proyecto de consultoría que tiene como objetivo validar el uso de técnicas de *machine learning* que apoyen la toma de decisiones del equipo de marketing respecto a los inmuebles para los que deben hacer inversiones en promoción dada su baja popularidad. Le han compartido dos conjuntos de datos: (1) el histórico de sus inmuebles con su respectiva popularidad y (2) los inmuebles que están próximos a publicarse. Adicionalmente, le han compartido el diccionario de datos correspondiente. Dentro del alcance del proyecto se deben considerar los siguientes entregables:

1. **[25%]** Un análisis exploratorio y de calidad de datos a partir de los datos históricos suministrados. En la medida de lo posible, incluya técnicas de análisis univariado, bivariado, gráficas y no gráficas. Realice la limpieza y preparación correspondiente para que puedan ser consumidos por los algoritmos de *machine learning* que va a utilizar. Adjunte la evidencia de como quedan los datos preparados, explique y justifique el proceso implementado.
2. **[50%]** Muestre los resultados tras el entrenamiento y selección del mejor modelo obtenido a partir de los datos suministrados. Utilice las técnicas de transformación

de datos, selección de atributos y regularización que considere convenientes. Explique cómo este modelo responde al objetivo planteado y si los resultados serían, de acuerdo a su criterio, suficientes para su uso por parte de la inmobiliaria. Recuerde argumentar sus decisiones con base en la interpretación cuantitativa y cualitativa de los resultados.

3. **[25%]** Utilice el mejor modelo obtenido para estimar la popularidad de los inmuebles próximos a publicarse. A partir de este resultado, apoye a la inmobiliaria en la definición del presupuesto de marketing que debe aprovisionar para promocionar los inmuebles menos populares sabiendo que la plataforma en la que se publican cobra por anuncio el 2% del precio del inmueble. Por ejemplo, si un inmueble tiene como precio \$200, la plataforma cobrará por publicar un anuncio para el inmueble un total de \$4. Usted es el encargado, a partir de los resultados de las predicciones, de definir el límite superior de comentarios para considerar un inmueble no popular. Explique esta decisión claramente así como el total de presupuesto que se debe aprovisionar acompañado del total de inmuebles a promocionar.

#### Condiciones de entrega:

- La entrega debe hacerse por BloqueNeón antes de la hora estipulada. Adjunte el notebook **ejecutado** con el desarrollo de cada uno de los puntos planteados de forma clara y ordenada. Además del notebook en formato IPYNB, adjunte la versión HTML o PDF del mismo. Estos archivos se pueden generar con la opción *File > Export As* en Jupyter Lab. Si está trabajando con Google Colab, siga alguna de las instrucciones planteadas en [este hilo](#) de StackOverflow.
- El nombre de los archivos a entregar debe tener como prefijo su login uniandes.
- No olvide incluir el código de honor.

**¡Éxitos!**