¡Felicidades! Has completado otro curso de la plataforma de entrenamiento. Ahora es el momento perfecto para probar tus habilidades y resolver un nuevo problema de machine learning. Realizarás este proyecto individualmente.

de las siguientes 48 horas. Después de eso, harás los cambios necesarios en tu trabajo y lo enviarás para una segunda revisión. Por lo general, este proceso se repetirá varias veces hasta que recibas el visto

Cuando termines, envía tu trabajo al revisor del proyecto. Recibirás feedback dentro

bueno de la revisión y se aprueben todas las correcciones. Todo eso es parte del trabajo.

Tu proyecto se considerará completado una vez que el revisor del proyecto lo apruebe.

### Descripción del proyecto Trabajas en la compañía de extracción de petróleo OilyGiant. Tu tarea es encontrar

los mejores lugares donde abrir 200 pozos nuevos de petróleo. Para completar esta tarea, tendrás que realizar los siguientes pasos:

región seleccionada: calidad de crudo y volumen de reservas.

• Leer los archivos con los parámetros recogidos de pozos petrolíferos en la

- Crear un modelo para predecir el volumen de reservas en pozos nuevos.
- Elegir los pozos petrolíferos que tienen los valores estimados más altos.
- Elegir la región con el beneficio total más alto para los pozos petrolíferos seleccionados.
- Tienes datos sobre muestras de crudo de tres regiones. Ya se conocen los parámetros de cada pozo petrolero de la región. Crea un modelo que ayude a elegir la región con el mayor margen de beneficio. Analiza los beneficios y riesgos potenciales utilizando la técnica bootstrapping.

## Condiciones:

- Solo se debe usar la regresión lineal para el entrenamiento del modelo.
- Al explorar la región, se lleva a cabo un estudio de 500 puntos con la selección de los mejores 200 puntos para el cálculo del beneficio.
- El presupuesto para el desarrollo de 200 pozos petroleros es de 100 millones de dólares. Un barril de materias primas genera 4.5 USD de ingresos. El ingreso de una
- unidad de producto es de 4500 dólares (el volumen de reservas está expresado en miles de barriles). • Después de la evaluación de riesgo, mantén solo las regiones con riesgo de
- seleccionar la región con el beneficio promedio más alto. Los datos son sintéticos: los detalles del contrato y las características del pozo no

pérdidas inferior al 2.5%. De las que se ajustan a los criterios, se debe

se publican.

### Los datos de exploración geológica de las tres regiones se almacenan en archivos:

Descripción de datos

geo\_data\_0.csv . Descarga el conjunto de datos

- geo\_data\_1.csv . Descarga el conjunto de datos
- geo\_data\_2.csv . Descarga el conjunto de datos
- id identificador único de pozo de petróleo
- f0, f1, f2 tres características de los puntos (su significado específico no es
- importante, pero las características en sí son significativas) product – volumen de reservas en el pozo de petróleo (miles de barriles).

# Instrucciones del proyecto

separadas.

2. Entrena y prueba el modelo para cada región en geo\_data\_0.csv:

validación en una proporción de 75:25

Descarga y prepara los datos. Explica el procedimiento.

- 1. Divide los datos en un conjunto de entrenamiento y un conjunto de
  - 2. Entrena el modelo y haz predicciones para el conjunto de validación.
  - 3. Guarda las predicciones y las respuestas correctas para el conjunto de validación.
  - 4. Muestra el volumen medio de reservas predicho y RMSE del modelo. 5. Analiza los resultados.
  - 6. Coloca todos los pasos previos en funciones, realiza y ejecuta los pasos 2.1-

2.5 para los archivos 'geo\_data\_1.csv' y 'geo\_data\_2.csv'.

3. Prepárate para el cálculo de ganancias:

1. Almacena todos los valores necesarios para los cálculos en variables

- 2. Dada la inversión de 100 millones por 200 pozos petrolíferos, de media un pozo petrolífero debe producir al menos un valor de 500,000 dólares en unidades para evitar pérdidas (esto es equivalente a 111.1 unidades). Compara esta cantidad con la cantidad media de reservas en cada región. 3. Presenta conclusiones sobre cómo preparar el paso para calcular el
- beneficio. 4. Escribe una función para calcular la ganancia de un conjunto de pozos de
- petróleo seleccionados y modela las predicciones: 1. Elige los 200 pozos con los valores de predicción más altos de cada una de
  - 2. Resume el volumen objetivo de reservas según dichas predicciones. Almacena las predicciones para los 200 pozos para cada una de las 3
  - regiones. 3. Calcula la ganancia potencial de los 200 pozos principales por región. Presenta tus conclusiones: propón una región para el desarrollo de pozos
- 1. Utilizando las predicciones que almacenaste en el paso 4.2, emplea la

petrolíferos y justifica tu elección.

5. Calcula riesgos y ganancias para cada región:

las 3 regiones (es decir, archivos 'csv').

- técnica del bootstrapping con 1000 muestras para hallar la distribución de los beneficios. 2. Encuentra el beneficio promedio, el intervalo de confianza del 95% y el riesgo de pérdidas. La pérdida es una ganancia negativa, calcúlala como
- una probabilidad y luego exprésala como un porcentaje. 3. Presenta tus conclusiones: propón una región para el desarrollo de pozos petrolíferos y justifica tu elección. ¿Coincide tu elección con la elección
- anterior en el punto 4.3? Evaluación del proyecto

#### Hemos definido los criterios de evaluación para el proyecto. Lee esto con atención antes de pasar al ejercicio.

Esto es lo que los revisores buscarán cuando evalúen tu proyecto:

 ¿Cómo preparaste los datos para el entrenamiento? ¿Seguiste todos los pasos de las instrucciones?

- ¿Consideraste todas las condiciones del negocio?
- ¿Cuáles son tus hallazgos sobre el estudio de tareas? ¿Aplicaste correctamente la técnica bootstrapping?
- ¿Sugeriste la mejor región para el desarrollo de pozos? ¿Justificaste tu elección?
- ¿Evitaste la duplicación de código? ¿Mantuviste la estructura del proyecto y el código limpio?

Ya tienes tus hojas informativas y los resúmenes de los capítulos, por lo que todo está listo para continuar con el proyecto.

¡Buena suerte!

¡Vamos allá!

O haz el proyecto en tu ordenador y súbelo cuando hayas terminado.

Subir y entregar



Siguiente