

## MACHINE LEARNING APLICADO A O&amp;G

**Fecha y Horario:**

7 al 11 de Octubre de 9:30 a 12:30 y de 14 a 17 hs

**Lugar:**

Sede Central del IAPG - Maipú 639, Ciudad de Buenos Aires

**Aranceles:**

Consultar a [cursos@iapg.org.ar](mailto:cursos@iapg.org.ar)

**Objetivos:**

Aprender a utilizar los modelos de aprendizaje automático más conocidos y aplicados en la industria petrolera. Esto lo lograremos estudiando sus fundamentos teóricos, cómo estimar sus parámetros y la forma de determinar la calidad de sus predicciones. Los ejemplos y ejercicios del curso se basan en datos reales y públicos relevantes para el sector.

Los participantes también aprenderán a evaluar la calidad de un set de datos y prepararlo para el entrenamiento del modelo elegido.

**A quién esta dirigido:**

Aplica a ingenieros, geocientistas y otros profesionales de la industria (de áreas como Finanzas, Negocios y RRHH) que posean conocimientos básicos previos en el lenguaje de programación Python y la librería Pandas para procesamiento de datos, así como adquisición de datos provenientes de bases de datos y APIs.

Si no cuenta con los conocimientos descriptos previamente, se recomienda:

- realizar el curso "Ciencia de datos en Python para O&G" ofrecido por el IAPG si no cuenta con conocimientos de Python.
- realizar el curso "Adquisición de datos con Python aplicado a O&G" ofrecido por el IAPG para comprender cómo adquirir y preprocesar datos para su uso en proyectos de aprendizaje automático.

**Introducción:**

Los modelos de aprendizaje automático (Machine Learning) han demostrado en los últimos años ser de gran utilidad para extraer conclusiones a partir de grandes cantidades de datos tanto dentro como fuera de la industria del O&G. Dichos modelos complementan a los tipos de algoritmos tradicionales ya usados en la industria.

El abanico de problemas que se pueden resolver con estos modelos contempla la producción, la perforación, desde el refinado hasta la caracterización de reservorios, sin estar limitados a estos ámbitos.

En este curso los participantes aprenderán a identificar un problema o pregunta que se pueda resolver mediante aprendizaje automático. Algunos de los modelos vistos son: decision trees, random forest, k-means, entre otros.

Además, podrán definir el problema de forma cuantitativa, verificar la utilidad, calidad y relevancia de datos disponibles, entrenar los modelos y determinar la validez de sus pronósticos, permitiendo elegir las mejores predicciones para el problema definido.

**Programa:****Módulo 1 (Introducción a ML):**

- Historia
- Problemas no supervisados
- Problemas supervisados
- Ejemplos en O&G

**Módulo 2 (Conceptos básicos y flujo general de ML):**

- Definición del problema y adquisición de datos
- Análisis exploratorio de los datos
- Preparación del dato
- Elección del modelo
- Entrenamiento del modelo
- Evaluación de la calidad y ajuste de hiperparámetros
- Predicción

**Módulo 3 (Preparación del dato):**

- Concepto de tidy dataset
- Feature engineering
- Manejo de datos de fechas
- Interpolación, datos nulos
- Variables dummies, etiquetas, otros casos

**Módulo 4 (Modelado):**

- División en training, test y validación
- Tipos de modelos
- Criterios de elección de modelos
- Decision trees
- Random forests
- K-means / Gaussian mixtures
- XGBoost

**Módulo 5 (Medición de calidad del modelo):**

- Accuracy, recall, precision, matriz de confusión
- Crossvalidation
- K-fold
- Error cuadrático medio
- Overfitting
- Ajuste de hiperparámetros y búsqueda eficiente

**Módulo 6 (Predicción de salida de modelo):**

- Elección de modelo o modelos ajustados
- Mezcla de modelos
- Reentrenamiento o cuando repetir todo el proceso

Todos los conceptos se ilustrarán de forma teórica y práctica.

**Importante**

Los participantes deben llevar sus laptops al curso. Es requisito técnico para la participación que la computadora de cada uno cuente con la instalación efectuada de Python y Jupyter Notebook. Los instructores podrán guiarlos vía mail previo al inicio del curso.

**Instructores:****Alfonso Barrios**

Alfonso Barrios es Ingeniero Geofísico, MS en Computación y Diplomado en Ciencia de Datos con más de 15 años de experiencia en software y data management aplicados a la industria petrolera. Hoy es Científico de Datos en Infolytics donde aplica sus conocimientos en proyectos de variadas industrias.

**Martín Gruber**

Martín Gruber es Analista de Sistemas graduado de ORT y cuenta con una diplomatura en Business Intelligence de la UTN. Se desempeña como Analista de Datos en Infolytics, trabajando con diversas bases de datos y herramientas de visualización tales como Power BI.

[Inscripción on line](#)[Descargar Formulario De Inscripción](#)[Calendario de Cursos](#)

para más Información solicitarla a [cursos@iapg.org.ar](mailto:cursos@iapg.org.ar)