



2019 Data Scientist

CURSO INTERMEDIO



+54 9 11 63761130

| www.alexandria.com.ar

| C.A.B.A., Argentina

QUIENES SOMOS

Somos un equipo de especialistas en Data Mining y Business Intelligence. Contamos con Ingenieros, Economistas y Estadísticos.

Tenemos experiencia en la implementación de este tipo de herramientas en el sector bancario, seguros, retail, oil & gas y otras.

Nuestro objetivo principal es generar una cultura de datos en las empresas. Para ello, consideramos primordial capacitar al estrato gerencial en cuestiones de datos, donde las decisiones de implementaciones se hacen más accesibles y se potencian los resultados.

Queremos ayudar en el cómo plantear las necesidades, cómo enfocar los proyectos, y cómo puede aplicarse la tecnología en todas las actividades del mercado.

Para ello brindamos capacitaciones y consultoría especializada en Machine Learning.

APLICACIONES

Generación de hipótesis, preparación de datos, así como mejora de la calidad de estos. Análisis exploratorio con métodos estadísticos. Modelado, que implica algoritmos de aprendizaje supervisado y no supervisado.

DESCRIPCIÓN DEL CURSO

Este curso pretende brindar al participante herramientas para desarrollarse en el mundo del análisis de datos y sus diversas aplicaciones. Las herramientas utilizadas son de software libre, brindando modelos de ejemplo sobre conjuntos de datos.

Durante el curso, aprenderá las técnicas para aplicar en el análisis de datos y la visualización de la información generada, comprendiendo también cómo desempeñar exitosamente un rol de estas características en cualquier equipo de trabajo u organización.

Se proveerá al participante de conocimientos para que pueda optimizar sus procesos de toma de decisiones basados en la información a partir de los datos disponibles, que sea capaz de entender las potencialidades de aplicar técnicas de análisis de datos en la empresa y que pueda liderar proyectos pequeños referidos a la minería de datos.

DETALLES

Modalidad: In-Company

Duración óptima: 4 días, 32 hs.

Costos: (Incluye material didáctico del curso)

Prerrequisito: Instalación de entorno Anaconda.

TESTES
TESTES
TESTES
TESTES
TESTES
TESTES
TESTES
TESTES
TESTES
TESTES

PERFIL DEL PARTICIPANTE

Profesionales de todas las áreas con predisposición de utilizar herramientas de análisis de datos para generar proyectos, o comprender como se generan los proyectos relacionados con Analytics.

Directivos, conducción intermedia, técnico-profesionales que participen de la generación de información.

OBJETIVOS

Al finalizar el curso, el participante será capaz de discernir si se trata de un proyecto a ser resuelto por medio de un algoritmo de aprendizaje supervisado o no supervisado.

Analizar el problema para identificar la variable target, tener los conocimientos básicos para saber que tratamiento aplica.

Podrá también, visualizar variables, realizar análisis estadísticos básicos de variables, aplicar transformaciones, trabajar con distintos tipos de variables, y habrá aplicado al menos un algoritmo de aprendizaje supervisado por su cuenta.

Se le brindará al participante herramientas para poder interactuar con profesionales de Analytics, generar solicitudes y proponer proyectos que mejoren la performance de su propia área.



Programa

CURSO INTERMEDIO

MODULO 1

Preprocesamiento de datos.
Limpieza de Datos.
Integración y transformación de datos.
Reducción de la dimensionalidad.
Discretización de variables.
Implementación de conocimientos adquiridos.

MODULO 2

Introducción a estadística.
Análisis exploratorio de datos.
Características de la distribución de las variables.
Análisis de las relaciones entre las variables .
Transformación y análisis de las mejoras.
Exploración de resultados y aplicaciones prácticas.

MODULO 3

Random Forest:

Características.

Problemas de overfitting y underfitting.

Problemas de Generalización.

Estimación de error (train-test split, cross-validation).

Ajuste de parámetros e hiperparámetros.

Práctica.

MODULO 4

Regresión Logística:

Características.

Problemas de overfitting y underfitting.

Problemas de Generalización.

Estimación de error (train-test split, cross-validation).

Ajuste de parámetros e hiperparámetros.

Práctica.

MODULO 5

K-Nearest Neighbors:

Características.

Problemas de overfitting y underfitting.

Problemas de Generalización.

Estimación de error (train-test split, cross-validation).

Ajuste de parámetros e hiperparámetros.

Práctica.

PROGRAMA

MODULO 6

K-Means:

Características.

Problemas de overfitting y underfitting.

Problemas de Generalización.

Estimación de error (train-test split, cross-validation).

Ajuste de parámetros e hiperparámetros.

Práctica.

MODULO 7

Redes Neuronales Simples:

Características.

Perceptrón.

XOR.

Práctica.

MODULO 8

Redes Neuronales:

Características.

Multiperceptrón.

Red CPN.

Práctica.

PROGRAMA