



2019 Data Scientist

CURSO BÁSICO



+54 9 11 49738574

| 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 otorgar al participante conocimientos básicos de datos y la utilidad que brinda a la toma de decisiones el conocer su potencial. Para ello, se brindará una introducción teórica a las bases de datos, el modo en que se relacionan las tablas y cómo se llega a explotar esos datos de manera accesible a personas que no son especialistas en Business Intelligence ni Bases de Datos.

Se utilizarán herramientas de software libre, de manera tutorada, que permitirán generar modelos predictivos sobre un conjunto de datos sencillos. Pero, sobre todo, donde podremos trabajar de manera conjunta en el análisis exploratorio de datos, previo a la implementación de los algoritmos.

Dado que se trata de un ámbito que utiliza especialmente la estadística, y que este curso no requiere conocimientos previos en este ámbito, se brindarán también herramientas para poder comprender por qué ciertas variables deben transformarse y cómo es que generan mejores resultados.

DETALLES

Modalidad: In-Company

Duración óptima: 2 días, 16 hs.

Costos: (Incluye material didáctico del curso)

Capacidad del curso: máximo 20 participantes

Requisito: Tener instalado Anaconda, Python y Jupyter notebook

PERFIL DEL PARTICIPANTE

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

Directivos, conducción intermedia y 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 qué 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 BÁSICO



MODULO 1

- Introducción a las bases de datos
 - Qué es análisis de datos y su utilidad
 - Proceso de descubrimiento del conocimiento:
Distintas fases
 - Tipos de tareas de Data Mining
 - Ejemplo de Aplicaciones
- Carga de un set de datos en Python y visualizar los tipos de datos con que se cuenta*

MODULO 2

- Preprocesamiento de datos
 - Limpieza de Datos
 - Integración y transformación de datos
 - Posibilidades de reducción de la dimensionalidad
 - Discretización de variables
- Implementación de conocimientos adquiridos al set de datos ya cargado*

MODULO 3

- Introducción estadística
- Análisis exploratorio y gráfico de los datos
- Características de la distribución de las variables
- Análisis de las relaciones entre las variables implicadas en el análisis
- Transformación y análisis de las mejoras

Exploración de resultados y aplicaciones prácticas

MODULO 4

- Aplicación de un modelo de clasificación (Árbol de clasificación) y análisis de resultados
- Mejoras de performance del modelo
- Análisis de parámetros del modelo
- Overfitting
- Comparativo por medio de curva ROC

Implementación y análisis de resultados para los sets de datos previamente trabajados

MODULO 5

- Otros tipos de modelos, finalidad y características básicas
- Cluster
- Regresión
- Redes

Práctica global

PROGRAMA