



Escuela Corporativa de Análisis de Datos

DATO.Institute es un emprendimiento cuyo objetivo principal es la de fomentar la cultura analítica empoderando a las personas a través de nuestro servicio educativo de calidad. Compuesta de capacitadores con más de 15 años de experiencia en Educación y en el ámbito Profesional, nos enfocamos en la formación profesional, capacitación y desarrollo de potencial humano en tecnología de información bajo marcas como:



PROGRAMA PROFESIONAL EN DATA SCIENCE AND BUSINESS ANALYTICS

Este programa consta de 8 módulos que le proporcionan las últimas habilidades y técnicas listas para el trabajo que cubren una amplia gama de temas de Data Science y Business Analytics, incluyendo: herramientas líderes de mercado y de código abierto, metodologías, Python, R, visualización de datos, análisis de datos y aprendizaje automático

MÓDULO 1

INTRODUCCIÓN AL DATA SCIENCE AND BUSINESS ANALYTICS

El arte de descubrir los conocimientos y las tendencias en los datos ha existido desde la antigüedad. Los antiguos egipcios utilizaban los datos del censo para aumentar la eficiencia en la recaudación de impuestos y predijeron con precisión la inundación del río Nilo cada año. Desde entonces, las personas que trabajan en ciencia de datos han creado un campo único y distinto para el trabajo que realizan. Este campo es ciencia de datos.

En este curso, conoceremos a algunos profesionales de la ciencia de los datos y obtendremos una descripción general de lo que es la ciencia de datos hoy en día y como aplicarla al negocio.

Duración: 20hs

Contenido:

- Conceptos (Data Mining, Visual Analytics, Business Intelligence, Business Analytics, Machine Learning, Data Science)
- Herramientas (IBM SPSS, IBM Watson, Tableau, KNIME, Python, R, ETL Spoon)
- Metodología para el desarrollo de Modelos Analíticos CRISP-DM
- Técnicas de Machine Learning y Data Science
- Visual Analytics
- Business Application

MÓDULO 2

MACHINE LEARNING PARA ANÁLISIS PREDICTIVO

El análisis predictivo consiste en la tecnología que aprende de la experiencia para predecir el futuro comportamiento de individuos para tomar mejores decisiones – Eric Siegel

El análisis predictivo es un área de la minería de datos que consiste en la extracción de información existente en los datos y su utilización para predecir tendencias y patrones de comportamiento, pudiendo aplicarse sobre cualquier evento desconocido, ya sea en el pasado, presente o futuro.

Tanto IBM SPSS Modeler y KNime son herramientas que proporcionan una interfaz gráfica intuitiva para ayudar a visualizar cada paso del proceso de la Minería de Datos y poder generar Modelos Predictivos. Los analistas y los usuarios empresariales pueden agregar fácilmente conocimientos especializados y conocimientos comerciales a dichos procesos.

Duración: 20hs

Contenido:

- Introducción a la Minería de Datos y Metodología CRISP-DM.
- Tipos de modelos predictivos.
- Técnicas y algoritmos principales (Regresión logística, Árboles de clasificación, Redes Neuronales, SVM, Time Series, ...).
- Métodos de Evaluación de modelos (Matriz de Clasificación, Matriz de Contingencia, Curvas ROC, Curvas Lift).
- Caso práctico de producción, evaluación e implementación de un modelo predictivo.

MÓDULO 3

DATA MINING PARA SEGMENTACIÓN

Clasificar a los diferentes clientes y potenciales es una labor necesaria para toda organización si se busca mejorar la rentabilidad. Para poder realizar dicha clasificación es necesario conocer cuáles son los criterios que se van a utilizar para clasificarlos, así como de qué manera vamos a analizarlos y evaluar si esa clasificación es la adecuada.

La segmentación de clientes es la base de una buena estrategia de marketing. Los clientes son diferentes entre sí, tienen necesidades diferentes. En el límite, cada cliente individual tiene necesidades e intereses únicos, y en esto se basa el "one to one" marketing.

Pero cuando necesitamos definir una estrategia de clientes la personalización total no sirve. Una estrategia para cada uno de los clientes conlleva demasiada complejidad. Por eso es mejor asumir que existen grupos de clientes similares entre sí, y distintos de los demás. Esto permite definir estrategias diferenciadas para cada grupo, en el plan de marketing. Segmentar es identificar estos grupos homogéneos y clasificar a los clientes en ellos.

En este curso aprenderemos en cómo llevar adelante una segmentación de clientes haciendo uso de un conjunto de técnicas de Machine Learning provistas por herramientas como SPSS Modeler y KNime.

Duración: 20hs

Contenido:

- Introducción a la Minería de Datos y a los modelos de Segmentación.
- Tipos de Segmentación y Condiciones Necesarias para la misma.
- Proceso de Segmentación.
- Definición de la población a segmentar.
- Generación de la Matriz de datos.
- Selección de Variables/Reducción de Dimensionalidad.
- Selección de técnicas y algoritmos.
- Formación de Segmentos.
- Caracterización/Perfilado.
- Automatización e Implantación.

MÓDULO 4

VISUAL ANALYTICS: NIVEL BÁSICO

Este curso está pensado para quienes trabajan con datos, independientemente de sus conocimientos técnicos o analíticos. Este curso está diseñado para ayudarlo a comprender y aplicar los conceptos y las técnicas importantes de visualización, para pasar de visualizaciones simples a complejas, y aprender cómo combinarlas en dashboards interactivos.

Duración: 20hs

Contenido:

- Comprender la terminología de Tableau.
- Preparar los datos para el análisis
- Crear vistas de datos
- Utilizar el eje tiempo
- Crear cálculos básicos: cálculos aritméticos básicos, agregaciones y relaciones personalizadas, cálculo de fechas y cálculos de tablas rápidos.
- Uso de varias medidas en una vista
- Filtros (en medidas y dimensiones, en varias hojas, condicionales, top / bottom N)
- Utilizar varios tipos de gráficos
 - Tabulaciones cruzadas
 - Gráficos circulares y de barras
 - Mapas geográficos
 - Gráficos de eje doble y gráficos combinados con diferentes tipos de marcas
 - Mapas de calor
 - Tablas de resaltado
 - Mapas de árbol
 - Diagramas de dispersión

MÓDULO 5

VISUAL ANALYTICS: NIVEL AVANZADO

Este curso está diseñado para adquirir los conocimientos que se necesita para convertirse en un usuario avanzado para la generación de análisis visual. A partir del módulo anterior, que se enfoca en conocer los conceptos fundamentales relativos a la creación de hojas de trabajo y dashboards, en este se persigue el saber trabajar con problemas más complejos.

Duración: 20hs

Contenido:

- Crear visualizaciones y tipos de gráficos avanzados.
- Crear cálculos complejos para manipular datos.
- Utilizar técnicas estadísticas para analizar datos.
- Implementar técnicas avanzadas de mapeo geográfico y utilizar geocodificación e imágenes personalizadas para crear visualizaciones espaciales de datos no geográficos.
- Preparar los datos para el análisis.
- Combinar diferentes fuentes de datos.
- Combinar datos de diferentes tablas en una misma fuente de datos mediante uniones de columnas.
- Crear mejores dashboards con técnicas para análisis guiados, diseño de dashboards interactivos y prácticas recomendadas de análisis visual.
- Implementar consejos y trucos para mejorar la eficiencia.

MÓDULO 6

BUSINESS ANALYTICS: INVESTIGACIÓN DE MERCADO AVANZADA

Este curso está diseñado para usuarios que se desempeñan en el área de Marketing e Investigación de Mercado o para quienes requieren analizar los resultados de sus encuestas, y/o diseñar o mejorar productos y/o servicios. Los ejemplos de aplicación de este curso se desarrollan en IBM SPSS Statistics y IBM SPSS Text Analytics for Survey, además de hacer uso de otras herramientas como KNime y Tableau.

Duración: 20hs

Al finalizar el curso usted podrá:

- Generar un modelo del proceso de toma de decisiones que realiza el cliente para luego diseñar productos con las características que este considera más importantes.
- Visualizar y explorar las relaciones entre las categorías de una variable y los casos de estudio.
- Identificar las características que más relacionan los consumidores con su producto o marca y saber cómo perciben sus productos en relación con otros.
- Convertir las variables cualitativas en cuantitativas mediante operaciones estadísticas.
- Conocer la utilidad de la creación de mapas perceptuales y el escalamiento multidimensional entre otras técnicas estadísticas.

MÓDULO 7

LENGUAJE R PARA DATA SCIENCE

R es un lenguaje de programación especialmente orientado al análisis estadístico y a la representación gráfica de los resultados obtenidos. Es un proyecto GNU. Por lo tanto, los usuarios son libres de modificarlo y extenderlo. Se trata de un lenguaje basado en comandos.

Este curso te proporcionará las bases del lenguaje de programación estadística R, el cual te permitirá escribir programas que lean, manipulen y analicen datos cuantitativos. Verás una introducción a los sistemas base de gráficos y al paquete para graficar ggplot2, para visualizar estos datos. Además también abordarás la utilización de uno de los IDEs más populares entre la comunidad de usuarios de R, llamado RStudio.

Duración: 20hs

Contenido:

- Introducción. Fundamentos teóricos de R.
- Instalación del ambiente de programación. Introducción a Listas y matrices. Operaciones generales.
- Introducción y operaciones con Dataframes. Exploración general de Datos.
- Conexión a fuentes de Datos. Preparación. Funciones de Filtro. Técnicas de manejo de "Missing Values".
- Familia de funciones "Apply".
- Introducción a distintas librerías de Analytics, Machine Learning y Visualizaciones (plyr, data.table, forecast, caret, ggplot2, etc.).

MÓDULO 8

LENGUAJE PYTHON PARA DATA SCIENCE

Este curso introduce un conjunto de conceptos básicos del entorno de programación de Python, incluidas las técnicas fundamentales de programación de Python, como lambdas, lectura y manipulación de archivos csv, y un conjunto de bibliotecas útiles para Data Science.

El curso introducirá técnicas de manipulación y limpieza de datos utilizando la popular biblioteca Pandas e introducirá la abstracción de la Series y el DataFrame como las estructuras de datos centrales para el análisis de datos, junto con tutoriales sobre cómo usar funciones como groupby, merge y tablas dinámicas con eficacia. Al final de este curso, ud. podrá tomar datos tabulares, limpiarlos, manipularlos y realizar análisis estadísticos inferenciales básicos.

Duración: 20hs

Contenido:

- Introducción. Fundamentos teóricos de Python.
- Instalación del ambiente de programación. Introducción a Listas, Tuplas, Diccionarios y Funciones. Operaciones generales.
- Controles de Flujo. Paquetes. Excepciones.
- Uso de distintas fuentes de Datos. Exploración y Preparación. Archivos, Bases de Datos.
- Consumo de datos a través de APIs.
- Introducción a distintas librerías de Analytics, Machine Learning y Visualizaciones (Numpy, Pandas, Scikit, matplotlib, geoplolib, etc).

NUESTROS DOCENTES



ING. WILFRIDO INCHAUSTTI
DATO / Head of Data Mining

+20 años experiencia
Profesor UNA, UCA, +15 tutorias de tesis
IBM SPSS Certified



ING. ROBERTO RUBIN
DATO / CEO

+15 años experiencia
Tableau Desktop and Server Qualified Associate
Tableau Dashboard Design Best Practices Deep Dive Certified



LIC. RODNEY ACEVEDO
COIN / CEO
DATO / Director

+20 años experiencia
Strategist and Marketing Specialist



ING. EMILIANO AMENDOLA
DATO / Data Scientist

+5 años experiencia
Tableau Desktop Qualified Associate
Data Science Researcher