

Data Scientist con R

Información General

Fecha de inicio: -

Frecuencia: Semanal

Carga horaria por clase: 90 min

Presentación

R es un lenguaje y entorno de programación de software libre para análisis estadístico, cálculo numérico y representaciones gráficas (modelos lineales y no lineales, tests estadísticos, análisis de series temporales, algoritmos de clasificación y otros). Por su parte Big Data es el término que se utiliza para describir los grandes volúmenes de datos que componen Internet y la posibilidad de obtener información sobre ellos. Actualmente compañías como Google, Amazon, Facebook, Netflix, Walmart, GAP, IBM y General Electric, entre otras, están liderando procesos de gestión masiva de información y reclutando "Data Scientists" en la nueva era que deviene de Big Data. El trabajo de un buen Científico de Datos es hacer descubrimientos "buceando" en un océano de datos cada vez más masivos y complejos. Identifican potenciales fuentes ricas en datos valiosos y acercan la brecha entre los datos, la información y finalmente las decisiones de negocio. Este entrenamiento tiene como objetivo familiarizar a los participantes con R y sus principales algoritmos más difundidos. Se espera que los participantes incorporen criterios para diseñar y utilizar repositorios de datos.

Objetivos Generales

Formar profesionales Científicos de Datos para los que se requiere que los participantes:
Ubiquen a la ciencia de datos en el contexto de disciplinas al servicio de los negocios
Entiendan la estructura de R
Conozcan las convenciones utilizadas de sintaxis de R
Sepan usar los algoritmos más difundidos
Puedan diseñar un repositorio de datos
Incorporen la necesidad de limpiar los datos antes de usarlos
Usen sin dificultad las herramientas de inferencia estadística
Complementen su formación y entrenamiento como Data Scientists

Objetivos Específicos

- Puedan comparar las ventajas y desventajas de los distintos algoritmos que se ofrecen para las diferentes técnicas descriptivas cubiertas.
- Identifiquen los principales algoritmos aplicados a la estadística descriptiva que vienen dentro de R
Entiendan el significado de cada parámetro de la invocación de los algoritmos descriptivos desde R
- Comparen adecuadamente las herramientas que existen en el mercado para soportar repositorios de datos
- Manejen la terminología relacionada con los repositorios de datos
- Palpen las consecuencias nefastas de no realizar adecuadamente la limpieza de los datos
- Adquieran criterio propio para diseñar repositorios de datos
- Incorporen el lugar que R va tomando como estándar de facto dentro de la ciencia de datos.

- Ubiquen los principales repositorios de los que bajar algoritmos
- Manejen la sintaxis de comandos de R
- Entiendan la estructura general de R
- Puedan comparar las ventajas y desventajas de los distintos algoritmos avanzados que se ofrecen para las diferentes técnicas predictivas cubiertas. Conozcan las futuras tendencias de la disciplina de la ciencia de datos.
- Dominen el concepto de ciencia de datos
- Identifiquen los principales algoritmos avanzados aplicados a la inferencia estadística que vienen dentro de R Entiendan el significado de cada parámetro de la invocación de los algoritmos predictivos avanzados desde R
- Puedan comparar las ventajas y desventajas de los distintos algoritmos que se ofrecen para las diferentes técnicas predictivas cubiertas.
- Identifiquen los algoritmos más simples aplicados a la inferencia estadística que vienen dentro de R Entiendan el significado de cada parámetro de la invocación de los algoritmos predictivos más simples desde R
- Manejen la sintaxis de R para generar los gráficos.
- Comprendan como distintas visualizaciones sirven o no para transmitir diferentes conceptos.
- Identifiquen las principales herramientas gráficas que vienen dentro de R

Destinatarios

- Junto con una formación previa en BI este curso complementa a los destinatarios en su formación como Data Scientists.
- Todos aquellos profesionales que quieran profundizar en el uso de R aplicado a la "Inteligencia de Negocios" (BI) ya sea para incorporarse en una posición de Analista de Business Intelligence como para mejorar sus posibilidades de crecimiento en este competitivo y creciente mercado.

Plan de Estudios

Módulo 1: Data Scientist con R

- Introducción a BI, R y la Ciencia de los Datos
- Diseño y construcción de repositorios de datos
- Algoritmos Descriptivos en R
- Herramientas de Graficación en R
- Algoritmos Predictivos básicos en R
- Algoritmos Predictivos avanzados en R
- Evaluación Final Integradora
- Recuperatorio de Evaluación Final Integradora

Bibliografía

- Pyle D. Business Modeling and Data Mining. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann Publisher. (2003)
- Hernández Orallo et al. Introducción a la Minería de Datos. Pearson Educación. (2004).
- Han, J., & Kamber, M. Data mining: Concepts and techniques. Morgan Kaufmann. (2001).
- Franco Jean Michel El Data Warehouse, El Data Mining. EDS Institut Prométhéus. Ed. F.

Devolt (1998).

- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. Advances in knowledge discovery and data mining. Menlo Park, CA: AAAI Press. . (1996).
- Christopher D. Manning, Prabhakar Raghavan y Hinrich Schütze,
- Cannataro M & Comito C. A Data Mining Ontology for Grid Programming. In Proceedings of the 1st International Workshop on Semantics Peer to Peer and Grid Computing. Budapest, 20-24 May 2003.
- Bernstein A., Provost F. and Hill S.
- Anand Rajaraman, Jure Leskovec y Jeffrey D. Ullman,
- Trevor Hastie, Robert Tibshirani y Jerome Friedman. Springer,

UTN.BA Campus Online — Universidad Tecnológica Nacional, Facultad Regional Buenos Aires

Mozart 2300, C1407 CABA, Argentina • +54 11 4867-7500 • campus@frba.utn.edu.ar

Documento generado el 23/07/2026 09:24 • UUID: 306dd134-31ad-4dd3-8057-9d8f95a75089