



TECNOLÓGICO
NACIONAL DE MÉXICO

INSTITUTO TECNOLÓGICO DE DURANGO

DIPLOMADO EN CIENCIA DE LOS DATOS E INTERNET DE LAS COSAS



Módulo V

Machine Learning.

Rubén Pizarro Gurrola

Objetivo

- Explorar, analizar, interpretar y difundir resultados mediante la adecuada implementación y aplicación de técnicas y modelos de ***machine learning*** haciendo uso del lenguaje de programación R y R Studio.
- Implementar técnicas y modelos de ***machine learning*** para una adecuada interpretación y difusión mediante lenguaje de programación R y R Studio.
- Desarrollar prácticas en R y R Studio analizando e interpretando datos mediante técnicas de ***machine learning***

Contenido

- Machine Learning
- Regresión
 - Regresión lineal simple
 - Regresión lineal múltiple
 - Regresión Polinomial
 - Regresión logística
- Clasificación
- Árboles de decisión
- Clustering o conglomerados
- Reducción de dimensionalidad
- Reconocimiento de lenguaje natural (Probable)
- Reconocimiento de imágenes (Probable)
- Herramientas
 - R
 - Big ML, Rapid Miner

Calendario

Tema	Fecha	Tiempo	Actividades	Evidencia
Machine Learning Flujo de trabajo Machine Learning	Viernes 28 Feb 2020	1 Hr	Teoría	Presentación
Regresión Lineal simple	Viernes 28 Feb 2020	3 Hrs	Teoría, práctica, herramienta.	Prácticas
Regresión línea múltiple	Sábado 29 Feb 2020	3 Hrs	Teoría, práctica, herramienta	Prácticas
Regresión polinomial, generalizada logística	Sábado 29 Feb 2020	3 Hrs	Teoría, práctica, herramienta	Prácticas

Calendario

Tema	Fecha	Tiempo	Actividades	Evidencia
Clasificación	Viernes 6 Mar 2020	2 Hr	Teoría, práctica, herramienta.	Prácticas
Arboles	Viernes 28 Feb 2020	2 Hr	Teoría, práctica, herramienta.	Prácticas
Clustering	Sábado 7 Mar 2020	2 Hr	Teoría, práctica, herramienta	Prácticas
Reducción de dimensionalidad	Sábado 7 Mar 2020	2 Hr	Teoría, práctica, herramienta	Prácticas
Reconocimiento de lenguaje natural y Reconocimiento de imágenes	Sábado 7 Mar 2020	2 Hr	Teoría, práctica, herramienta	Prácticas

Recordando. Introducción y Fundamento de R



```
library(readr)  
library(dplyr)  
library(sqldf)  
library(plyr)
```

RPubs *brought to you by RStudio*



Antes de iniciar.

- Recordatorio introducción R y R Studio
- Requisitos
 - schoology: Su cuenta y contraseña e identificar curso
 - git y github: Descargar e instalar
 - rpubs: Probar su espacio y <https://rpubs.com/rpizarro>
- Práctica: Para recordar y sensibilizar
- crear un proyecto enlazado github para explorar datos “alumnos inscritos ene-jun 2018.csv” y llevar el control de versiones.
 - Crear proyecto en github
 - Crear proyecto en R Studio
 - Cargar datos
 - Explorar datos
 - Visualizar diagrama de dispersión de datos

Práctica control de versiones y recordando R

- Objetivo

- Recordar R y R Studio y sensibilizar con control de versiones git y github

- Control de Versiones:

- Instalar git
- Instalar y crear cuenta github
- Crear un proyecto en github con archivo readme
- Configurar git en R Studio tolos/global Options/git
- Crear un nuevo proyecto git en R Studio
 - Crear un script para explorar datos women()
 - Salvar cambios localmente
 - Añadir a git, commit, mensaje, commit pull y verificar en github

- Identificar conjunto de datos “alumnos inscritos ene-jun 2018.csv”

- Crear nuevo script

- Cargar librerías
- Cargar datos
- Explorar datos
- Filtrar datos
- Analizar datos por carrera
- Histograma de promedios por carrera
- Agregar a proyecto y a control de versiones github,
- Verificar url

<https://github.com/rpizarrog/pruebaControlVersiones/blob/master/scripts/explorar%20datos%20de%20alumnos%20inscrits%202018.R>