

Introducción a la Ciencia de datos y Machine Learning

JUSTIFICACIÓN

Debido a la gran variedad de retos presentes en la industria y la academia surge la necesidad de dar nuevos enfoques a métodos estadísticos tradicionales y el desarrollo de nuevas metodologías que puedan satisfacer las necesidades en procesamiento y análisis de datos que requiere el mundo de hoy. Este curso ofrece una serie de herramientas para el procesamiento y transformación de datos estructurados y la aplicación de modelos de machine learning para resolver problemas de regresión y clasificación.

METODOLOGÍA

Aprendizaje basado en casos
Aprendizaje Colaborativo
Simulaciones, entre otros.

DIRIGIDO A

Directivos y coordinadores en áreas de analítica y toma de decisiones.
Profesionales en áreas de innovación y tecnología.

RESULTADO DE APRENDIZAJE

- Manejar paquetes en Python para análisis de datos y ETL.
- Conocer e implementar el método de KNN como herramienta de clasificación y Regresión.
- Conocer e implementar el Modelo de regresión lineal (Ridge, Lasso, polinómico) y logística como herramientas explicativas y predictivas.
- Conocer las métricas de evaluación de un modelo de clasificación y regresión.
- Conocer e implementar los diferentes modelos de clasificación supervisada (SVM, Decision Trees, Naive Bayes, random forest y GB Decision trees)
- Optimizar su modelo de aprendizaje de máquinas mediante técnicas de selección de parámetros.

CONTENIDO

Módulo I: Supervised Learning part 1

- KNN: Classification and Regression
- Linear Regression(Ridge, Lasso and Polynomial)
- SVM
- Decision Trees

Módulo II: Model Evaluation

- UnderFitting, Over fitting
- Confusion Matrix and evaluation Metrics
- Model Selection

Módulo III: Supervised Learning part 2

- Naives Bayes
- Random Forest
- Gradient Boosted Decision Tree.
- Data Leakage

DURACIÓN DEL PROGRAMA

48 HORAS

MODALIDAD

REMOTA

EXPERTO FACILITADOR

Jairo Andrés Díaz Rodríguez

Certified Data Scientist with solid fundamental background, experienced in the whole DS pipeline, from data acquisition and transformation, to visualization, model selection (Statistics and ML based), and deployment into fully functional production systems. I combine both Industry and Academia worlds with top-level research and teaching experience in Statistics, Applied Mathematics, Optimization and Machine Learning.

PhD in Mathematics (Statistics), University of Geneva (Switzerland). Master in Mathematics and Computer Science, University of Geneva (Switzerland). Bachelor in Electronic Engineering, Industrial University of Santander (Colombia) 2010.

Manuel Mendoza Becerra