

El camino de un Data Scientist

Disciplinas y competencias

Matemáticas ✓

- Algebra lineal
- Cálculo vectorial
- Cálculo matricial

- Optimización
- Métodos iterativos

Probabilidad ✓

- Distribución de probabilidad
- Inferencia bayesiana

- Variables aleatorias
- Máxima verosimilitud

Estadística ✓

- Descripción de datos
- Muestras aleatorias
- Estimación de parámetros
- Test de Hipótesis

- Regresión lineal
- Correlación
- Regresión múltiple
- Estadística no paramétrica
- Grafos

Programación ✓

- Python
- Bases de datos
- SQL, MongoDB

- R
- Linux
- Git

Machine Learning

- Scikit-learn
- Underfitting, Overfitting
- Random Forests

- Pipelines
- XGBoost
- Data Leakage

Deep Learning

- Keras
- Pytorch
- TensorFlow

- Redes neuronales
- Descenso por gradiente
- Clasificación binaria

Ingeniería de factores

- Baseline Model
- Codificación categórica

- Generación de factores
- Selección de factores

Natural Language Processing

- Clasificación de texto

- Word Vectors

Visualización de datos

- BI (Business Intelligence)
- Tableau, Power BI, Qlik

- Excel VBA

Deployment ✓

- Microsoft Azure
- Heroku
- Google Cloud Platform

- Flask
- Django

Otros

- Conocimiento de negocio
- Habilidades comunicacionales

- Proyectos
- Portafolios

Práctica