

# MÓDULO 2: HERRAMIENTAS BIG DATA

## CONTENIDO

### 0. Introducción a Linux

#### 1. El BigData en el Contexto del Machine Learning a Gran Escala

#### 2. Limitaciones Teóricas de las Bases de Datos Relacionales

#### 3. Escalado Vertical vs. Escalado Horizontal

#### 4. Estrategias de computación distribuida y en paralelo

#### 5. Arquitectura de la Solución Machine Learning para BigData

5.1. Capa de Ingesta de Datos

5.2. Capa de Ingestión

5.3. Capa de Almacenamiento (Hadoop Storage)

5.4. Plataforma Hadoop – Capa de Procesamiento

5.5. Capa Analítica

5.6. La Capa de Consumo

5.7. La Capa de Seguridad y Monitorización

#### 6. Ecosistema Hadoop

6.1. Componentes del Ecosistema Hadoop

6.2. Distribuciones Hadoop y Proveedores

#### 7. DevOps – Desarrollo (Dev) y Operaciones (Ops)

7.1. La importancia de productivizar/industrializar modelos estadísticos

#### 8. Le Evolución de los Entornos BigData

#### 9. ¿Cómo se organizan los equipos de Analítica Avanzada?

#### 10. Soluciones y Librerías Machine Learning

#### 11. Herramientas Machine Learning una Visión General

11.1. Apache Mahout

11.2. Julia

11.3. Python

11.4. Apache Spark

11.5. H2O

#### 12. Paquetes Avanzados de R

#### 13. Introducción al RGPD (Reglamento General Protección de Datos)

13.1. Introducción

13.2. Ámbito de aplicación del RGPD

13.3. Conceptos Básicos

13.4. Principios

13.5. Derechos de las personas

13.6. Obligaciones del responsable del tratamiento

## PROFESORES RESPONSABLES

- D. Carlos Ortega Fernández
- Dña. María Rosario Heras Carrasco

(Para mayor información de los profesores, puede consultar el menú “Equipo Docente” de la Web)