

Presentación de la Asignatura Tecnologías para el Análisi de Datos Masivos

Máster de Análisis de Datos Masivos UIB: Juan Gabriel Gomila
& Ricardo Alberich

17/10/2020

¿Quiénes somos?



Figure 1: Juan Gabriel Gomila



Figure 2: Ricardo Alberich

¿Quiénes somos? Juan Gabriel Gomila

- ▶ Departamento de Ciencias Matemáticas e Informática e la UIB
- ▶ Profesor asociado del área de Ciencia de la Computación e Inteligencia artificial
- ▶ Licenciado en Matemáticas por la UIB
- ▶ CEO Frogames
- ▶ Y más cosas. . .
- ▶ Email

¿Quiénes somos? Ricardo Alberich

- ▶ Departamento de Ciencias Matemáticas e Informática e la UIB
- ▶ Profesor Titular del área de Ciencia de la Computación e Inteligencia Artificial
- ▶ Licenciado en matemáticas por la Universidad de Valencia
- ▶ Doctor en informática por la UIB
- ▶ Email

Asignatura 11630 - Tecnologías para el Analisis de Datos Masivos

- ▶ Guía docente(català)
- ▶ Cronograma: Horarios de clase
- ▶ Espacio discord de la asignatura
- ▶ Espacio moodle de la UIB de la asignatura

Temas de la asignatura

Todo será de carácter práctico y aplicado. Ya se profundizará en otras asignaturas según los itinerarios que hayáis elegido.

Grandes temas (no necesariamente en este orden) son tecnologías para:

- ▶ Tema 1 Repaso de estadística descriptiva, inferencia y Introducción a R y RStudio Este será un tema transversal al curso cubierto con los materiales del curso online de Udemy los anexos.
- ▶ Tema 2. Manipulación de datos con Tidyverse. Este tema se tratará sólo en R pero se extensible al lenguaje python
- ▶ Tema 3. Machine Learning y Aprendizaje Estadístico
- ▶ Tema 4 Protección de datos y Legislación.

Contenidos Tema 3 (y 1)

- ▶ Parte 1 - Preprocesamiento de datos
- ▶ Parte 2 - Regresión: Regresión Lineal Simple, Regresión Lineal Múltiple, Regresión Polinomial, SVR, Regresión en Árboles de Decisión y Regresión con Bosques Aleatorios
- ▶ Parte 3 - Clasificación: Regresión Logística, K-NN, SVM, Kernel SVM, Naive Bayes, Clasificación con Árboles de Decisión y Clasificación con Bosques Aleatorios
- ▶ Parte 4 - Clustering: K-Means, Clustering Jerárquico
- ▶ Parte 5 - Aprendizaje por Reglas de Asociación: Apriori, Eclat

Contenidos Tema 3 (y 2)

Continuación

- ▶ Parte 6 - Reinforcement Learning: Límite de Confianza Superior, Muestreo Thompson
- ▶ Parte 7 - Procesamiento Natural del Lenguaje: Modelo de Bag-of-words y algoritmos de NLP
- ▶ Parte 8 - Deep Learning: Redes Neuronales Artificiales y Redes Neuronales Convolucionales
- ▶ Parte 9 - Reducción de la dimensión: ACP, LDA, Kernel ACP
- ▶ Parte 10 - Selección de Modelos & Boosting: k-fold Cross Validation, Ajuste de Parámetros, Grid Search, XGBoost

Metodología

Actividades de trabajo presencial (2,16 créditos, 54 horas)

1. Clases teóricas *Descripción: Sesiones expositivas, explicativas y / o demostrativas de contenido. Las presentaciones pueden ser realizadas por el profesor o por los estudiantes supervisado por el profesor.
 - ▶ Horas: 36
2. Clases prácticas
 - ▶ Descripción: Trabajos prácticos en el aula en el que se desarrollan actividad de aplicación de los conocimientos a situaciones concretas y la adquisición y habilidades relacionadas con la asignatura.
 - ▶ Horas: 18

Evaluación

1. Actividades de trabajo no presencial (6,84 créditos, 171 horas)

a. Presentación de un tema de TECNOLOGÍAS

- Descripción: Preparación de un tema de formación complementario referente a las TECNOLOGÍAS en grupo para entregar al profesor y con exposición ORAL en clase. Puede haber evaluación en forma de test si los profesores lo consideran adecuado.
- Horas: 31

2. Estudio y trabajo autónomo individual o en grupo

Descripción: Preparación de seminarios, lecturas, investigaciones, trabajos, memorias u obtención y análisis de datos para entregar al profesor o exponer en clase. Horas: 140 Criterio de evaluación: -

Actividad de evaluación

- ▶ Clases teóricas peso 30% recuperable si ? ; nota mínima 5
- ▶ Clases prácticas peso 20% Recuperable no
- ▶ Presentación de un tema de NT 15% Recuperable no
- ▶ Estudio y trabajo autónomo individual o en grupo 35%, Recuperable sí; nota mínima 5
- ▶ La nota mínima es la nota a partir de la cual se puede aplicar el peso que corresponda sobre la nota final.