



DESCRIPCIÓN DE LA ASIGNATURA

| | |
|---|---|
| Grado/Máster en: | Master Universitario en INGENIERÍA INFORMÁTICA por la Universidad de Málaga |
| Centro: | Escuela Técnica Superior de Ingeniería Informática |
| Asignatura: | ACELERACIÓN DE ANALYTICS (ACCELERATING ANALYTICS) |
| Código: | 211 |
| Tipo: | Optativa |
| Materia: | ESPECIALIDAD EN INGENIERÍA Y CIENCIA DE DATOS |
| Módulo: | COMPLEMENTOS EN TECNOLOGÍAS INFORMÁTICAS |
| Experimentalidad: | |
| Idioma en el que se imparte: | Español |
| Curso: | 2 |
| Semestre: | 1 |
| Nº Créditos | 4,5 |
| Nº Horas de dedicación del estudiante: | 112,5 |
| Nº Horas presenciales: | 33,8 |
| Tamaño del Grupo Grande: | |
| Tamaño del Grupo Reducido: | |
| Página web de la asignatura: | http://mop.cv.uma.es |

EQUIPO DOCENTE

Departamento: ARQUITECTURA DE COMPUTADORES
Área: ARQUITECTURA Y TECNOLOGÍA DE COMPUTADORES

| Nombre y Apellidos | Mail | Teléfono Laboral | Despacho | Horario Tutorías |
|---------------------------------------|---------------|------------------|--|------------------|
| Coordinador/a: RAFAEL ASENJO PLAZA | asenjo@uma.es | 952132791 | 2.2.31 - E.T.S. Ing. Telecomunicación | |
| LUIS FELIPE ROMERO GOMEZ | felipe@uma.es | 952134169 | 2.2.25 - E.T.S.I. INFORMÁTICA | |

RECOMENDACIONES Y ORIENTACIONES

It is recommended that the student had taken and passed the previous courses 'Data Engineering and Science', and 'High Performance for Analytics and Cybersecurity'. This course is complementary and dovetails well with the ones included in the Intensification of Data Engineering and Science, mainly with 'Big Data', 'On-line Data Mining' and 'Visual Data Analysis'.

CONTEXTO

This course aims at achieving three main goals: i) to provide a high-level survey of key analytics models and algorithms; ii) to analyze the usage patterns of these models; and iii) to discuss HW and SW opportunities for accelerating analytics workloads.

The particular technologies used for the hands-on sessions will be Python, and R to a lesser extent. Python and R have been covered in previous courses, like Data engineering and science I and II, but now the goal is to study particular implementations of common data analytics patterns and optimize these implementations taking full advantage of the underlying hardware. More precisely, analytics algorithms implemented in Python and R will be accelerated by exploiting modern multicore and heterogeneous architectures. To this end, interfaces from Python and R to C and C++ (as Cython and Rcpp) will be leveraged. Among the different analytics kernels, the course will focus on the more widely used, as regression, clustering, decision trees and time series.

COMPETENCIAS

1 Competencias generales y básicas.

Competencias básicas

- 1.1 CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación
- 1.2 CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
- 1.3 CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios
- 1.4 CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades
- 1.5 CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo

Competencias generales

- 1.1 CG1 - Capacidad para proyectar, calcular y diseñar productos, procesos e instalaciones en todos los ámbitos de la ingeniería informática.



1 Competencias generales y básicas.

Competencias generales

- 1.4 CG4 - Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la ingeniería en informática
- 1.8 CG8 - Capacidad para la aplicación de los conocimientos adquiridos y de resolver problemas en entornos nuevos o poco conocidos dentro de contextos más amplios y multidisciplinares, siendo capaces de integrar estos conocimientos.
- 1.10 CG10 - Capacidad para aplicar los principios de la economía y de la gestión de recursos humanos y proyectos, así como la legislación, regulación y normalización de la informática.

2 Competencias específicas.

- 2.1 ET11 - Capacidad para modelar, diseñar, definir la arquitectura, implantar, gestionar, operar, administrar y mantener aplicaciones, redes, sistemas, servicios y contenidos informáticos.
- 2.8 ET18 - Capacidad de diseñar y desarrollar sistemas, aplicaciones y servicios informáticos en sistemas empotrados y ubicuos.

3 Competencias transversales.

- 3.2 CT2 - Capacidad para identificar estrategias, herramientas y métodos que responden a situaciones de éxito que pueden ser abordadas con los recursos disponibles.

CONTENIDOS DE LA ASIGNATURA

Contenido

- * Introduction. Data Analytic exemplars
 - Patterns
 - Accelerators: SIMD, GPU, FPGA

- * Accelerating Python with C++
 - Pybynd11
 - Cython
 - Intel Distribution for Python

- * Accelerating R with C++
 - Rcpp

- * Case studies
 - Regression
 - Clustering
 - Decision trees
 - Time series
 - Monte Carlo simulation

- * Project
 - Geospatial analytics
 - Visibility problems, traffic planning,...

ACTIVIDADES FORMATIVAS

Actividades presenciales

Actividades expositivas

Lección magistral

Actividades prácticas en instalaciones específicas

Prácticas en laboratorio



ACTIVIDADES DE EVALUACIÓN

Actividades de evaluación presenciales

Actividades de evaluación del estudiante

Realización de trabajos y/o proyectos

RESULTADOS DE APRENDIZAJE / CRITERIOS DE EVALUACIÓN

As learning outcomes, the student will be able to:

- 1.- know the software and hardware technologies that offer opportunities to accelerate and efficiently implement engineering applications that process massive volumes of data.
- 2.- hone the skills to design and develop data science and engineering systems that exploit large volumes of multimedia data using graphics processor-based architectures.
- 3.- know and apply techniques and tools for large data sequences that are generated continuously and that must be processed on-line, including the detection and treatment of trend change.

Assessment methods and criteria linked to learning outcomes:

- Outcome 1, 2 and 3 will be evaluated from practical exercises, lab practices and a final project.

PROCEDIMIENTO DE EVALUACIÓN

The final assessment will take into account the exercises, lab practices and the project (LAB), as well as other activities related to the student active participation in class and course forums (SAP):

Final Grade= $0,9 \cdot \text{LAB} + 0,1 \cdot \text{SAP}$

For the second and special calls, the student must pass a final exam consisting of two parts: theory (a maximum of 3 points) and laboratory (a maximum of 7 points).

BIBLIOGRAFÍA Y OTROS RECURSOS

Básica

- Kurt Smith. "Cython, A guide for Python programmers", O'Reilly Media, 2015. ISBN: 978-1-491-90155-7
- Norman Matloff. "Parallel Computing for Data Science", CRC Press, 2016, ISBN: 978-1-4665-8703-8
- Rajesh Bordawekar, Bob Blainey, Ruchir Puri, "Analyzing Analytics", Morgan & Claypool, 2016, ISBN: 978-1-6270-5835-3
- Wes McKinney, "Python for Data Analysis", 2nd edition. O'Reilly Media, 2017, ISBN: 978-1-491-95766-0

DISTRIBUCIÓN DEL TRABAJO DEL ESTUDIANTE

ACTIVIDAD FORMATIVA PRESENCIAL

| Descripción | Horas | Grupo grande | Grupos reducidos |
|--|--------------|-------------------------------------|--------------------------|
| Prácticas en laboratorio | 22,8 | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Lección magistral | 11 | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| TOTAL HORAS ACTIVIDAD FORMATIVA PRESENCIAL | 33,8 | | |
| TOTAL HORAS ACTIVIDAD FORMATIVA NO PRESENCIAL | 67,45 | | |
| TOTAL HORAS ACTIVIDAD EVALUACIÓN | 11,25 | | |
| TOTAL HORAS DE TRABAJO DEL ESTUDIANTE | 112,5 | | |

