

Curso Académico: (2023 / 2024)

Fecha de revisión: 19-12-2023

Departamento asignado a la asignatura: Departamento de Teoría de la Señal y Comunicaciones

Coordinador/a: MOLINA BULLA, HAROLD YESID

Tipo: Optativa Créditos ECTS : 6.0

Curso : 1 Cuatrimestre : 1

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Conocimientos básicos en Python

OBJETIVOS

El objetivo principal de esta asignatura es formar analistas en ciencias de datos, ya sea para investigación o el mercado laboral, usando las técnicas más novedosas que se aplican en aprendizaje máquina.

Nos orientaremos aprender como sacarle el mayor partido a los recursos computacionales que podemos tener a nuestra disposición, desde el ordenador propio, recursos en la nube para programación con GPUs y programación para Big Data; conocimientos con alta demanda en diversos entornos.

Para se explicarán los conceptos básicos de programación paralela:

- 1- usando los recursos de los ordenadores de propósito general,
- 2- programación híbrida: usando los recursos de un ordenador normal y hardware específico como las tarjetas de procesamiento gráfico (GPUs)
- 3- programación distribuida y en la nube, para grandes casos, como Big Data.

Competencias Básicas

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.

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.

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.

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.

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:

CG1 Capacidad para mantener una formación continuada una vez graduado, permitiéndole enfrentarse a tecnologías de nueva aparición.

CG2 Capacidad para aplicar los conocimientos de las habilidades y métodos de investigación relacionados con las Ingenierías.

CG3 Capacidad para aplicar los conocimientos de las habilidades y métodos de investigación relacionados con las Ciencias de la Vida.

CG4 Capacidad para contribuir a la ampliación de las fronteras del conocimiento a través de una investigación original, parte de la cual merezca la publicación referenciada a nivel internacional.

Competencias específicas:

CE4 Capacidad para utilizar técnicas para el tratamiento de cantidades masivas de datos e imágenes médicas.

CE5 Capacidad para poner en práctica los métodos de tratamiento de información e imagen médica.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

Programación Paralela:

- * Programación Multiprocesador/Multicore
- * Programación Paralela con Memoria Compartida (y los peligros para los datos)
- * Programación Paralela con Memoria Compartida con semáforos y bloqueo

Programación Híbrida

- * Uso de las GPUs y su diferencia con el procesador del ordenador común.
- * Como se programan las GPUs: cuando y como se pueden utilizar.
- * Uso de técnicas avanzadas de programación en GPUs, gestión de los recursos de GPU.
- * Como usar la GPU en Aprendizaje Máquina

Programación Distribuida

- * Que es la computación distribuida y como podemos aprovecharla.
- * Uso de plataformas para programación distribuida para BigData y Aprendizaje Máquina
- * Casos de uso de Apache Spark para aprendizaje máquina

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Es una asignatura orientada a la práctica. Se impartirán los conceptos básicos teóricos de computación intensiva, para realizar las prácticas guiadas. Se implementarán casos de uso de las tecnologías, orientadas a formar un investigador avanzado en ciencia de datos.

Actividades Formativas:

- AF3 Clases teórico prácticas
- AF4 Prácticas de laboratorio
- AF5 Tutorías
- AF6 Trabajo en grupo
- AF7. Trabajo individual del estudiante
- AF8. Exámenes parciales y finales

METODOLOGÍA

MD1: Exposiciones en clase del profesor con soporte de medios informáticos y audiovisuales, en las que se desarrollan los conceptos principales de la materia y se proporciona la bibliografía para complementar el aprendizaje de los alumnos.

MD2: Lectura crítica de textos recomendados por el profesor de la asignatura

MD3: Resolución de casos prácticos, problemas, etc.... planteados por el profesor de manera individual o en grupo

MD4: Exposición y discusión en clase, bajo la moderación del profesor de temas relacionados con el contenido de la materia, así como de casos prácticos

MD5: Elaboración de trabajos e informes de manera individual o en grupo

Se establecerán 2 horas a la semana de tutorías para los alumnos donde el profesor estará disponible en su despacho.

SISTEMA DE EVALUACIÓN

EVALUACIÓN CONTINUA.

- * Prácticas realizadas durante la clase: 20%
- * Practica Computacion Masiva en el Ordenador 20%
- * Práctica Computación Heterogénea (Ordenador + GPU): 20%
- * Practica Final Especial (Distribuida y Heterogénea): 40%

CONVOCATORIA EXTRAORDINARIA

- * Examen presencial con preguntas sobre las prácticas realizadas.

Peso porcentual del Examen Final:	0
Peso porcentual del resto de la evaluación:	100

BIBLIOGRAFÍA BÁSICA

- Benjamin Bengfort ; Jenny Kim Interactive Spark using PySpark, O'Reilly Media, 2016
- Holden Karau ; Rachel Warren High Performance Spark, O'Reilly Media, 2017
- Ian Gorton, Deborah K. Gracio Data-Intensive Computing: Architectures, Algorithms, and Applications,

