



UNIVERSIDAD DE LAS PALMAS
DE GRAN CANARIA

GUÍA DOCENTE

CURSO: 2020/21

51191 - CIENCIA Y ANALÍTICA DE DATOS

CENTRO: 415 - IU de Microelectrónica Aplicada

TITULACIÓN: 5048 - MU Electrónica y Telecomunicación Aplicadas

ASIGNATURA: 51191 - CIENCIA Y ANALÍTICA DE DATOS

CÓDIGO ULPGC: 51191 **CÓDIGO UNESCO:** 120304

MÓDULO: MÓDULO COMÚN

MATERIA:

TIPO: Obligatoria

CRÉDITOS ECTS: 6

CURSO: 1

SEMESTRE: 1º semestre

LENGUA DE IMPARTICIÓN (Especificar créditos de cada lengua)

ESPAÑOL: 6

INGLÉS:

SUMMARY

Data science combines computer science and statistics to solve data-intensive problems in industry and in many fields of science. The objective of the subject "Science and Data Analytics" is to obtain a solid understanding of the methods used in data science, analyzing how and why they work.

In this subject the student:

- learns the implications of handling massive data and the existing tools to deal with them
- applies data mining algorithms in the resolution of massive data problems
- learns the existing tools for the parallel and distributed processing of large volumes of data
- learns the existing types of machine learning as well as the methodologies for the design of the resulting models and their validation
- solves different problems through the treatment and adjustment of supervised and unsupervised data models
- applies machine learning algorithms in the resolution of classification and recommendation problems
- acquires skills in the use of tools to solve automatic learning problems

REQUISITOS PREVIOS

Plan de Enseñanza (Plan de trabajo del profesorado)

Contribución de la asignatura al perfil profesional:

La asignatura "Ciencia y Analítica de Datos" complementa y amplía la asignatura de "Métodos de Investigación Científica y Técnica para Electrónica y Telecomunicación Aplicadas" agregando otras técnicas de análisis y diseño y contribuyendo de esta forma en la abundancia de la capacidad de seleccionar la más adecuada para resolver un problema concreto en el ámbito del aprendizaje automático.

Competencias que tiene asignadas:

- CG1. Que los estudiantes sean capaces de plantear y llevar a cabo un trabajo de investigación en el ámbito de la ingeniería electrónica y de telecomunicación avanzada.
- 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.
- CT1 - Liderar equipos y organizaciones, promoviendo el libre intercambio de ideas y experiencias, la búsqueda de soluciones originales y el compromiso permanente con la excelencia.
- CT2 - Impulsar responsablemente todas las formas de conocimiento y de acción que puedan contribuir al enriquecimiento del capital económico, social y cultural de la sociedad en la que el estudiante desarrolla su práctica profesional y en la que ejerce sus derechos y deberes de ciudadanía.
- CE.C2. Conocer los componentes disponibles de arquitectura de software y utilizar los modelos de algoritmia, estructura de datos y programación adecuados para la resolución de problemas de aprendizaje automático (ML).
- CE.C3 - Evaluar la calidad de los modelos de minería de datos en términos de indicadores de error relevantes en cada tarea..

Objetivos:

Los objetivos de la asignatura son:

- OBJ-1. Comprender los principios computacionales y estadísticos generales que subyacen en los algoritmos modernos de aprendizaje automático y minería de datos.
- OBJ-2. Aplicar diversos métodos para analizar datos científicos y de negocios.
- OBJ-3. Evaluar la idoneidad de cada método para la recopilación y el uso de datos.
- OBJ-4. Implementar soluciones de aprendizaje automático de última generación de manera eficiente utilizando plataformas informáticas de alto rendimiento.
- OBJ-5. Generar informes con los resultados de manera clara y comprensible.
- OBJ-6. Analizar datos científicos e industriales para diseñar nuevas aplicaciones y respaldar la toma de decisiones

Contenidos:

Introducción a los datos masivos
Herramientas para Data Science (Python, Jupiter, Numpy, Pandas, Keras, ...)
Minería de datos (recolección, limpieza, almacenamiento y visualización de datos)
Tipos de aprendizaje automático, conceptos básicos
Metodología de diseño de modelos para Machine Learning
Validación de modelos (tasas de error, matriz de confusión, curvas ROC)
Clasificación y regresión con máquinas de vectores de soporte (SVM)
Clustering y clasificadores con árboles de decisión

Construcción de modelos predictivos basados en redes neuronales
Sistemas de computación para datos masivos (herramientas en la nube)
Clasificadores de imágenes (sistemas supervisados y no supervisados)
Introducción al procesado de texto (lingüística y extracción básica de características)
Análisis de sentimiento y redes sociales
Introducción a sistemas de recomendación

Bloque 1. Aprendizaje automático. 40 horas (30 Teoría + 10 Laboratorio)

Tema 1. Introducción y herramientas de aprendizaje automático. 10 horas (10T)

Tema 2. Técnicas clásicas de aprendizaje automático. 10 horas (8T +2L)

Tema 3. Deep Learning. 10 horas (6T+4L)

Tema 4. Aprendizaje automático en la práctica. 10 hora (6T + 4L)

Bloque 2. Big Data. 20 horas (15 Teoría + 5 Laboratorio)

Tema 5. Ecosistema Hadoop. 6 horas (6T)

Tema 6. Herramientas de procesado y ejecución en Hadoop. 8 horas (5T + 3L)

Tema 7. Algoritmos de aprendizaje automático en entornos distribuidos. 6 horas (4T + 2L)

Metodología:

MD1. Método expositivo/Lección magistral. Enseñanza directa expositiva y demostrativa para aquellos contenidos esenciales y/o que requieran una explicación detallada por parte del profesor. En el caso de que se interrumpan las clases presenciales, las clases magistrales se impartirán mediante vídeos o videoconferencia.

MD2. Actividades prácticas. Actividades presenciales que requieren la transferencia de conocimientos conceptuales con los procedimentales, favoreciendo la autonomía y la capacidad de reflexión de los estudiantes, así como fomentando las habilidades personales, y las interpersonales mediante el trabajo en equipo.

MD3. Trabajos, proyectos y memorias. Realización y exposición individual o en grupo de trabajos monográficos sobre la asignatura

MD4. Actividades no presenciales: destinadas al fomento del estudio y al desarrollo por parte del alumno de las competencias de trabajo autónomo y de autoaprendizaje.

Evaluación:

Criterios de evaluación

La evaluación de la asignatura se realizará teniendo en cuenta tres aspectos principales: los contenidos teóricos, la realización de las actividades prácticas y la calidad del trabajo expuesto en clase.

Las fuentes y los criterios de evaluación que se emplearán serán los siguientes:

1) Evaluación teórica

FE1. Exámenes escritos.

Esta fuente de evaluación está relacionada con las actividades formativas AF1, AF2 y AF5.

Criterios relativos a los exámenes escritos:

- Aplicación de la técnica de diseño más adecuada para resolver un problema concreto.
- Análisis de los problemas en función de la eficiencia de sus soluciones.
- Estructuración de los algoritmos.

- Se valorarán los cuestionarios de estudiantes.

2) Trabajo práctico en grupos

FE2. Prácticas.

Esta fuente de evaluación está relacionada con las actividades formativas AF2 y AF5.

Criterios relativos a las prácticas:

-Los programas e información entregados se someterán a los análisis y pruebas oportunos para formular una valoración global que tenga en cuenta todos los aspectos implicados en su realización:

- Diseño (relaciones entre clases y adecuación de los métodos a la funcionalidad descrita en las especificaciones).
- Ejecución (funcionamiento, robustez y eficiencia).
- Estilo (formato, comentarios y elección de identificadores).
- Tipos de errores mostrados por las pruebas diseñadas.

3) Exposición del trabajo realizado en grupos

FE3. Contenido del trabajo.

FE4. Exposición oral de cada estudiante.

Ambas fuentes de evaluación están relacionadas con las actividades formativas AF1, AF2 y AF5.

Criterios relativos al contenido de los trabajos:

- Análisis de un problema de aprendizaje automático.
- Valoración adecuada de la documentación estudiada indicando las fuentes bibliográficas.

Criterios relativos a la calidad de la exposición oral realizada en clase:

- Exposición clara y fluída y adaptación al tiempo establecido.
- Réplica adecuada a las preguntas formuladas.

4) Examen global

FE5. Examen global.

Esta fuente de evaluación está relacionada con las actividades formativas AF1, AF2 y AF5.

Los criterios para la evaluación del examen escrito global son los mismos que los establecidos para la evaluación teórica. Los criterios para la evaluación del examen práctico global son los mismos que los establecidos para la evaluación de las prácticas.

Sistemas de evaluación

SE1. Exámenes escritos u orales orientados a evaluar las competencias adquiridas por los alumnos. En el caso de que se interrumpan las clases presenciales, los exámenes se harán por vía telemática.

SE2. Trabajos, proyectos y memorias escritas realizadas por el estudiante de manera individual o en grupo.

SE3. Exposición de trabajos, proyectos y memorias realizados por el estudiante de manera individual o en grupo. En el caso de que se interrumpan las clases presenciales, los estudiantes enviarán una presentación en vídeo al profesor.

SE4. Realización de las actividades pautadas en las prácticas de laboratorio y prácticas con ordenadores..

Criterios de calificación

SE1. Competencias evaluadas: CG1, CB6, CB7, CB8, CB9, CB10, CT1, CT2, CE.CE2, CE.CE.C3 Ponderación 30%

SE2. Competencias evaluadas: CG1, CB6, CB7, CB8, CB9, CB10, CT1, CT2, CE.CE2, CE.CE.C3 Ponderación 30%

SE3. Competencias evaluadas: CG1, CB6, CB7, CB8, CB9, CB10, CT1, CT2, CE.CE2, CE.CE.C3 Ponderación 10%

SE4. Competencias evaluadas: CG1, CB6, CB7, CB8, CB9, CB10, CT1, CT2, CE.CE2, CE.CE.C3

Ponderación 30%

Plan de Aprendizaje (Plan de trabajo de cada estudiante)

Tareas y actividades que realizará según distintos contextos profesionales (científico, profesional, institucional, social)

Las tareas que realizará el estudiante, todas ellas pertenecientes a los contextos científico y profesional, son las siguientes:

Ta1. Búsqueda, consulta y estudio de la documentación específica y técnica de la materia.

Ta2. Preparación de la exposición del trabajo.

Ta3. Trabajo práctico consistente en la realización de las prácticas de laboratorio que requieren solución individual o en grupo.

Ta4. Exámenes escritos.

Temporalización semanal de tareas y actividades (distribución de tiempos en distintas actividades y en presencialidad - no presencialidad)

Los estudiantes dedicarán 4 horas semanales presenciales a las actividades presenciales de clases de teoría, resolución de problemas en aula y prácticas de laboratorio. Estas actividades incluirán la realización de ejercicios de aplicación de las técnicas introducidas en la teoría y la solución de problemas (60 horas en 15 semanas). Los exámenes que se realizarán a lo largo del semestre consumen 2 horas presenciales en total. Asimismo, cada grupo dedicará 6 horas semanales (no presenciales) a la elaboración y exposición de los trabajos solicitados.

AF1. Clases de teoría. Clase presencial en la que se explican fundamentos teóricos y metodológicos mediante el uso de presentaciones con proyector o pantalla y pizarra. 1,72 ECTS

AF2 . Prácticas de laboratorio, prácticas de aula, seminarios, talleres y trabajos dirigidos. Ejercicios y actividades presenciales de carácter eminentemente práctico, para contextualizar el aprendizaje teórico a través de su aplicación en casos prácticos del ámbito de la electrónica y la telecomunicación aplicadas. 0,60 ECTS

AF4 Actividades de evaluación. Destinadas a valorar el grado de adquisición de las competencias

por parte del estudiante haciendo uso de los sistemas previstos en la presente memoria. 0,08 ECTS

AF5. Trabajo autónomo. Actividad no presencial que requiere en primer lugar el estudio por parte del alumno, y en segundo lugar que el alumno, de manera individual o en grupo, se aplique a la comprensión de los contenidos, la búsqueda de información, la realización de tareas, la redacción, presentación y exposición de trabajos, y la preparación de exámenes, cuestionarios y pruebas. 3,60

Recursos que tendrá que utilizar adecuadamente en cada uno de los contextos profesionales.

Re1. Bibliografía, documentación específica sobre el contenido de la asignatura y documentación técnica. Contextos: científico y profesional.

Re2. Navegador web (búsquedas, consultas de documentación electrónica, interacción con compañeros de grupo). Contextos: científico, profesional y social.

Re3. Entorno Integrado de Desarrollo para Python. Contextos: científico y profesional.

Re4. Herramientas proporcionadas por el Campus Virtual. Contextos: científico, profesional y social.

Resultados de aprendizaje que tendrá que alcanzar al finalizar las distintas tareas.

El estudiante aprende las implicaciones de manejar datos masivos y las herramientas existentes para tratar con ellos

El estudiante aplica algoritmos de minería de datos en la resolución de problemas de datos masivos

El estudiante aprende las herramientas existentes para el procesado paralelo y distribuido de grandes volúmenes de datos

El estudiante aprende los tipos de aprendizaje automático existentes así como las metodologías para el diseño de los modelos resultantes y su validación

El estudiante resuelve diferentes problemas mediante el tratamiento y ajuste de modelos de datos de carácter supervisado y no supervisado

El estudiante aplica los algoritmos de aprendizaje automático en la resolución de problemas de clasificación y recomendación

El estudiante adquiere destreza en el manejo de herramientas para la resolución de problemas de aprendizaje automático

Plan Tutorial

Atención presencial individualizada (incluir las acciones dirigidas a estudiantes en 5ª, 6ª y 7ª convocatoria)

Se usará para aclarar dudas, asesorar al estudiante en las tareas individuales y optimizar su rendimiento. El profesor atenderá consultas del alumnado en su horario de tutorías. Las citas se pueden concertar a través del correo electrónico institucional o utilizando las herramientas proporcionadas por el Campus Virtual. Los estudiantes en 5ª, 6ª y 7ª convocatoria tendrán prioridad frente al resto de compañeros en el horario de tutoría del profesor.

Atención presencial a grupos de trabajo

Se usará para aclarar dudas, orientar las actividades colectivas dirigidas, mejorar el rendimiento académico, hacer un seguimiento del trabajo realizado y modelar las habilidades sociales. El profesor atenderá consultas del alumnado en su horario de tutorías. Las citas se pueden concertar a través del correo electrónico institucional o utilizando las herramientas proporcionadas por el Campus Virtual.

Atención telefónica

El profesor de la asignatura atenderá consultas telefónicas del alumnado en su horario de tutorías.

Atención virtual (on-line)

El profesor de la asignatura estará disponible para atender a los estudiantes virtualmente a través del correo electrónico institucional o utilizando las herramientas proporcionadas por el Campus Virtual.

Datos identificativos del profesorado que la imparte.

Datos identificativos del profesorado que la imparte

Dr./Dra. José María Quinteiro González (COORDINADOR)
Departamento: 238 - INGENIERÍA TELEMÁTICA
Ámbito: 560 - Ingeniería Telemática
Área: 560 - Ingeniería Telemática
Despacho: INGENIERÍA TELEMÁTICA
Teléfono: 928451236 **Correo Electrónico:** josemaria.quinteiro@ulpgc.es

Dr./Dra. Luis Miguel Hernández Acosta
Departamento: 238 - INGENIERÍA TELEMÁTICA
Ámbito: 560 - Ingeniería Telemática
Área: 560 - Ingeniería Telemática
Despacho: INGENIERÍA TELEMÁTICA
Teléfono: 928451383 **Correo Electrónico:** luismiguel.hernandez@ulpgc.es

Bibliografía

[1 Básico] Deep learning /

Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
The MIT Press,, Cambridge, MA : (2016)
978-0-262-03561-3

[2 Básico] Python machine learning [

Sebastian Raschka.
Packt Publishing,, Birmingham : (2015)
9781783555130

[3 Recomendado] Deep learning with Keras :implement deep learning models and neural networks with the power of Phyton /

Antonio Gulli, Sujit Pal.
Packt Publishing,, Birmingham ; : (2017) - (1ª ed.)
978-1-78712-842-2