



FICHA TÉCNICA DE LA ASIGNATURA

| Datos de la asignatura | |
|------------------------|---|
| Nombre completo | Machine Learning I: Regresión y Clasificación / Machine Learning I: Regression and Classification |
| Código | FCEE-BA-315 |
| Título | Grado en Análisis de Negocios / Business Analytics por la Universidad Pontificia Comillas |
| Impartido en | Grado en Análisis de Negocios/Business Analytics y Grado en Derecho [Tercer Curso] Grado en Administración y Dirección de Empresas y Grado en Análisis de Negocios/Business Analytics [Tercer Curso] Grado en Ingeniería en Tecnologías de Telecom. y Grado en Análisis de Negocios/Business Analytics [Tercer Curso] |
| Nivel | Reglada Grado Europeo |
| Cuatrimestre | Semestral |
| Créditos | 6,0 ECTS |
| Carácter | Obligatoria (Grado) |
| Departamento / Área | Departamento de Métodos Cuantitativos |
| Responsable | Carlos Martínez de Ibarreta Zorita |
| Horario de tutorías | previa petición de cita por correo electrónico. El horario concreto estará disponible en cuanto los horarios sean definitivos |

| Datos del profesorado | |
|-----------------------|---------------------------------------|
| Profesor | |
| Nombre | Cristina Lozano Colomer |
| Departamento / Área | Departamento de Métodos Cuantitativos |
| Despacho | Alberto Aguilera 23 |
| Correo electrónico | clozano@icade.comillas.edu |
| Profesor | |
| Nombre | Carlos Martínez de Ibarreta Zorita |
| Departamento / Área | Departamento de Métodos Cuantitativos |
| Despacho | Alberto Aguilera 23 [OD-413] |
| Correo electrónico | charlie@icade.comillas.edu |
| Teléfono | 2261 |
| Profesor | |
| Nombre | Miguel Ángel Sanz Bobi |



| | |
|----------------------------|--|
| Departamento / Área | Departamento de Telemática y Computación |
| Despacho | Alberto Aguilera 25 [D-419] |
| Correo electrónico | Miguelangel.Sanz@iit.comillas.edu |
| Teléfono | 4240 |

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura

Aportación al perfil profesional de la titulación

Muchas decisiones y problemas de carácter empresarial se pueden mejorar mediante la aplicación de técnicas analíticas a la inmensa cantidad de datos de todo tipo disponibles.

La asignatura es una introducción a las técnicas de machine learning, entrando luego en detalle en las técnicas básicas de predicción y clasificación.

Para que sea útil, la asignatura se centrará en la aplicación de las técnicas a datos y problemas reales, de manera que el alumno tenga las herramientas para poder enfrentarse a problemas empresariales reales más adelante.

Prerequisitos

Fundamentos de análisis estadístico de datos

Fundamentos de programación en R

Fundamentos de Business Analytics

Competencias - Objetivos

Competencias

GENERALES

| | | | |
|-------------|---|------------|--|
| CG03 | Resolución de problemas y toma de decisiones en un entorno de datos masivos tanto cuantitativos como cualitativos | | |
| | <table border="1"><tr><td>RA1</td><td>Saber seleccionar para cada problema la técnica o técnicas de análisis de datos más adecuada para poder convertir los datos ¿en bruto¿ en información y ésta en conocimiento que ayude a la toma de decisiones y a mejorar la gestión.</td></tr></table> | RA1 | Saber seleccionar para cada problema la técnica o técnicas de análisis de datos más adecuada para poder convertir los datos ¿en bruto¿ en información y ésta en conocimiento que ayude a la toma de decisiones y a mejorar la gestión. |
| RA1 | Saber seleccionar para cada problema la técnica o técnicas de análisis de datos más adecuada para poder convertir los datos ¿en bruto¿ en información y ésta en conocimiento que ayude a la toma de decisiones y a mejorar la gestión. | | |
| CG04 | Capacidad para elaborar proyectos e informes de manera oral y escrita, difundiendo estas ideas a través de canales digitales | | |



| | | |
|--------------------|------------|---|
| | RA1 | Ser capaz de resumir, sintetizar y comunicar de una forma atractiva y eficaz los resultados de la aplicación de las técnicas de análisis de datos, incluso de las más sofisticadas, de manera que resulten comprensibles a destinatarios no técnicos y ayuden de forma eficiente a la toma de decisiones empresariales. |
| ESPECÍFICAS | | |
| CE19 | | Conocer los fundamentos de las principales técnicas tanto de la estadística clásica (descriptiva e inferencial) como del data mining |
| | RA3 | Conocer los fundamentos de las principales técnicas de data mining supervisado (predictivo). |
| CE20 | | Saber modelizar un problema empresarial real que precise análisis de datos y seleccionar críticamente la técnica o combinación de técnicas más adecuada |
| | RA2 | Saber modelizar un problema de predicción (regresión, clasificación) |
| | RA4 | Conocer las ventajas y limitaciones de los procedimientos de data mining y saber cómo enfocar un problema desde diferentes técnicas complementarias |
| CE21 | | Saber interpretar, evaluar y comunicar resultados derivados de las técnicas de análisis de datos así como usarlos para la ayuda en la gestión y la toma de decisiones empresariales |
| | RA2 | Saber interpretar, evaluar y comunicar los resultados derivados de un análisis que emplee técnicas de data mining predictivo. |
| CE22 | | Saber aplicar las técnicas de análisis de datos (tanto las de la estadística clásica como las técnicas de data mining) a un conjunto de datos reales, mediante el empleo de algún software apropiado para tal fin |
| | RA3 | Saber aplicar técnicas de data mining predictivo (aprendizaje supervisado) a un conjunto de datos reales usando software apropiado para tal fin |

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Introducción a las técnicas de Machine Learning. Aprendizaje supervisado y no supervisado.

Trabajando con datos. introducción a R Studio. Preprocesamiento, enriquecimiento y análisis descriptivo preliminar

Predicción: Regresión lineal múltiple



Performance predictiva. Partición de los datos: conjunto de entrenamiento y de validación

Clasificación: Regresión logística

Performance en tareas de clasificación.

Otras técnicas: KNN , árboles de clasificación y random forest

Introducción a las redes neuronales y a los SVM

Herramientas de construcción, evaluación y mejora de un modelo. Ensemble de modelos.

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

La asignatura es de orientación eminentemente práctica y aplicada.

Los alumnos antes de clase deberán haber preparado los contenidos a ver.

En la clase presencial se reforzarán los contenidos teóricos y conceptos principales, para pasar enseguida a casos sencillos de aplicación.

Algunas semanas se encargarán para la semana siguiente casos y ejercicios para aplicar lo visto y profundizar más en los conceptos.

El trabajo de aplicación práctica tratará de aplicar las técnicas de machine learning a datos reales aplicados a un problema preferentemente de carácter empresarial o económico

Metodología Presencial: Actividades

Exposición de los principales conceptos teóricos

Realización de ejemplos de aplicación sencillos

Puesta en común y corrección de casos y problemas realizados por los alumnos

Actividades de evaluación

CG03, CG04,
CE19, CE20, CE21,
CE22

Metodología No presencial: Actividades

Estudio personal

Realización de casos y trabajos de aplicación práctica, empleando datos y programación

CG03, CG04,
CE19, CE20, CE21,
CE22

RESUMEN HORAS DE TRABAJO DEL ALUMNO



| HORAS PRESENCIALES | | |
|---|--|---|
| Lecciones de Carácter expositivo | Ejercicios y resolución de casos y de problemas | |
| 30.00 | 30.00 | |
| HORAS NO PRESENCIALES | | |
| Ejercicios y resolución de casos y de problemas | Estudios individual y/o en grupo, y lectura organizada | Trabajos monográficos y de investigación, individuales o colectivos |
| 30.00 | 30.00 | 30.00 |
| CRÉDITOS ECTS: 6,0 (150,00 horas) | | |

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

| Actividades de evaluación | Criterios de evaluación | Peso |
|---|---|------|
| Examen final de carácter teórico práctico | Calificación numérica de 0 a 10 | 50 |
| Pruebas periódicas y evaluación continua | Calificación numérica de 0 a 10 | 20 |
| Trabajo de aplicación práctica | Uso de datos reales y programación Calificación A,B,C,D Propuesta 30% y Trabajo final 70% | 20 |
| Casos prácticos periódicos | A realizar en casa Calificación A,B,C,D | 10 |

Calificaciones

Es necesario obtener un 5 en el examen final para aprobar la asignatura en cualquiera de las convocatorias.

El trabajo de aplicación práctica hay que realizarlo para poder aprobar la asignatura, tanto en la convocatoria ordinaria como extraordinaria. En una asignatura de carácter aplicado, es necesario demostrar que se tienen adquiridas las habilidades para aplicar los conceptos teóricos a datos reales mediante el empleo de las técnicas de análisis adecuadas y las habilidades de programación necesarias

En tercera y posteriores convocatorias, se recomienda realizar nuevamente el trabajo, si bien no es obligatorio. La calificación final en ese caso será la mejor de las dos siguientes: a) 70% examen y 30% trabajo y b) 100% examen

Alumnos de intercambio (IN): mismo regimen que los alumnos ordinarios



Alumnos de intercambio (OUT): mismo regimen que para terceras y posteriores convocatorias

PLAN DE TRABAJO Y CRONOGRAMA

| Actividades | Fecha de realización | Fecha de entrega |
|------------------------|----------------------|---------------------------------------|
| Propuesta de trabajo | | hacia la mitad de la asignatura |
| Trabajo práctico final | | día del examen final |
| Pruebas tipo test | | al final de cada bloque de contenidos |

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

Materiales de Moodle

Shmueli et al. (2017) *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R*. Ed. Wiley

Bibliografía Complementaria

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media, Inc.

James, Gareth, et al. *An introduction to statistical learning*. Vol. 112. New York: springer, 2013.

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

<https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>