uc3m Universidad Carlos III de Madrid

Análisis de Datos

Curso Académico: (2021 / 2022) Fecha de revisión: 29/06/2021 10:53:42

Departamento asignado a la asignatura: Coordinador/a: GARCIA HERRERO, JESUS Tipo: Optativa Créditos ECTS: 6.0

Curso: 4 Cuatrimestre:

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Programación (1º, 1C) Inteligencia Artificial (2º, 2C)

OBJETIVOS

Conocimientos:

- 1. Evaluación basada en varias tareas de análisis de datos
- 2. Conocimiento sobre las técnicas clásicas y de aprendizaje automático para desarrollar modelos que expliquen los datos
- 3. Conocimiento de técnicas que tratan con los errores habituales en los datos para poder utilizarlos

Procedimentales/Instrumentales:

- 1. Los estudiantes usarán diferentes técnicas de análisis, comparándolas mediante experimentos y analizando los resultados
- 2. Los estudiantes aplicarán la técnica de análisis apropiada con los parámetros correctos para resolver cada tarea (objetivo)

Actitud:

- 1. Los estudiantes harán los trabajos en equipo
- 2. Los estudiantes usarán herramientas de análisis para dar soluciones a problemas del mundo real desde la informática
- 3. Capacidad de aplicar los conocimientos en la práctica
- 4. Los estudiantes presentarán una memoria por trabajo y el trabajo final se presentará oralmente, además de hacer el examen final escrito
- 5. Los estudiantes deben saber usar el estado del arte de las herramientas de análisis de datos para resolver los trabajos propuestos

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

- 1. Introducción al análisis de datos y minería de datos
- 2. Aprendizaje con técnicas numéricas
- 2.1. Modelos estadísticos de los datos y relaciones causales
- 2.2. Clasificadores bayesianos. Atributos numéricos y simbólicos
- 3. Árboles para predicción numérica
- 3.1. Regresión lineal y árboles de regresión
- 3.2. Agrupamiento con técnicas numéricas: K-medias y EM
- 4. Evaluación de prestaciones
- 4.1. Matrices de confusión
- 4.2. Comparación de técnicas, contrastes de significatividad.
- 5. Análisis de atributos
- 5.1. Selección no supervisada
- 5.2. Transformación de atributos
- 5.3. Selección supervisada
- 6. Metodologías en análisis de datos

7. Otras técnicas avanzadas (combinación, SVM, aprendizaje de sistemas difusos, técnicas genéticas,...)

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Clases magistrales: 2 ECTS. Para los objetivos de conocimiento y análisis y en relación con todos los temas se presentarán las ideas fundamentales mediante clase magistral que se valorará mediante exámenes de conocimiento. Resolución de supuestos prácticos: 2,5 ECTS. Para los objetivos de comprensión, análisis, síntesis y crítica y en relación con todos los temas se desarrollarán casos concretos donde el alumno deberá demostrar su capacidad para plantear soluciones coherentes y analizar soluciones propuestas por otros alumnos, la valoración se hará en función de las propuestas de solución de cada alumnos a los problemas propuestos.

Desarrollo de un trabajo práctico: 1,5 ECTS. Para el objetivo de desarrollo y en relación con uno de los temas de la asignatura el alumno deberá desarrollar un trabajo práctico

SISTEMA DE EVALUACIÓN

Peso porcentual del Examen/Prueba Final: 30
Peso porcentual del resto de la evaluación: 70

Las clases magistrales se valorarán mediante un examen, o varios exámenes realizados durante el curso y que supondrán el 30% de la nota final.

Los supuestos prácticos se valorarán mediante la presentación de propuestas de solución a los problemas planteados y representarán un 20% de la nota final.

La evaluación vendrá en gran medida dada por un proyecto práctico desarrollado en el cuatrimestre que recoja los conceptos teóricos presentados y su aplicación a un supuesto práctico. Los equipos de trabajo estarán formados por grupos de dos personas. Se evaluará la entrega en plazo del trabajo, su desarrollo en el cuatrimestre y la presentación de resultados y conclusiones. El desarrollo del trabajo práctico representará un 50% de la nota final.

BIBLIOGRAFÍA BÁSICA

- I. Witten y E. Frank Data Mining: Practical Machine Learning Tools and Techniques (Third Edition) , Morgan Kaufmann, 2011
- Jesús García, Antonio Berlanga, José M. Molina, Miguel A. Patricio Ciencia de datos: Técnicas analíticas y aprendizaje estadístico en un enfoque práctico, Altaria, 2018

BIBLIOGRAFÍA COMPLEMENTARIA

- David Hand, Heikki Mannila Principles of data mining, MIT Press, 2002
- Pérez López, César Estadística aplicada a través de Excel, Prentice Hall, 2002