

Curso Académico: (2021 / 2022)

Fecha de revisión: 10/06/2021 17:47:14

Departamento asignado a la asignatura: Departamento de Informática

Coordinador/a: ALER MUR, RICARDO

Tipo: Optativa Créditos ECTS : 6.0

Curso : 4 Cuatrimestre :

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Programación II

OBJETIVOS**1.) DE CONOCIMIENTO:**

- Dominar los conceptos básicos sobre la extracción de conocimiento a partir de datos
- Conocer las diferentes tareas que se pueden resolver con aprendizaje automático
- Conocer las técnicas de aprendizaje automático y su tipología
- Conocer la metodología de extracción de conocimiento y las fases que conlleva
- Conocer herramientas disponibles para la extracción de conocimiento

2.) DE COMPRENSIÓN:

- Comprender los conceptos básicos de la extracción del conocimiento
- Comprender los fundamentos y las motivaciones de la minería de datos
- Comprender la metodología de trabajo y las distintas fases de extracción de conocimiento
- Comprender la utilidad de las diferentes técnicas de extracción de conocimiento
- Comprender las diferencias de diferentes representaciones: proposicionales y relacionales
- Comprender la relación entre complejidad del modelo, cantidad de datos, características del problema y sobreaprendizaje

3.) DE APLICACIÓN:

- Analizar los dominios y diseñar procesos de extracción de conocimiento acordes al problema.
- Evaluar las prestaciones y eficiencia de los distintos métodos de extracción de conocimiento
- Trabajar sobre dominios específicos y contrastar distintas técnicas para comprobar su rendimiento en la extracción de conocimiento

4.) CRÍTICA O VALORACIÓN

- Selección de algoritmos, selección de modelos y ajuste de parámetros.
- Considerar la relación entre coste computacional y mejora marginal de diferentes soluciones
- Valoración de si los resultados obtenidos son adecuados, comparados con el azar o algoritmos básicos

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1. Introducción al aprendizaje automático
2. Métodos básicos para clasificación y regresión:
 - 2.1 Vecino más cercano (KNN)
 - 2.2 Árboles y reglas
3. Metodología en aprendizaje automático
 - 3.1. Entrenamiento de modelos
 - 3.2. Ajuste de hiper-parámetros
 - 3.3. Evaluación de modelos
 - 3.4. Preproceso de datos y selección de atributos
4. Aprendizaje automático con R y con la librería MLR (mlr3).
5. Métodos avanzados para clasificación y regresión:
 - 5.1. Conjuntos de modelos (ensembles): bagging, Random Forests, boosting
 - 5.2. Máquinas de vectores de soporte
6. Breve introducción al aprendizaje automático en Python (librería scikit-learn)

1. Introducción al aprendizaje automático
2. Métodos básicos para clasificación y regresión:
 - 2.1 Vecino más cercano (KNN)
 - 2.2 Árboles y reglas
3. Metodología en aprendizaje automático
 - 3.1. Entrenamiento de modelos
 - 3.2. Ajuste de hiper-parámetros
 - 3.3. Evaluación de modelos
 - 3.4. Preproceso de datos y selección de atributos
4. Aprendizaje automático con R y con la librería MLR (mlr3).
5. Métodos avanzados para clasificación y regresión:
 - 5.1. Conjuntos de modelos (ensembles): bagging, Random Forests, boosting
 - 5.2. Máquinas de vectores de soporte
6. Breve introducción al aprendizaje automático en Python (librería scikit-learn)

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

La asignatura se impartirá mediante lecciones magistrales las clases de teoría y mediante clases tutoradas las clases prácticas.

Las clases magistrales estarán enfocadas a enseñar todos los conceptos relativos a aprendizaje automático.

Las clases prácticas (grupos reducidos) se desarrollarán para que, de un modo tutorado, el alumno aprenda a resolver problemas reales. Las prácticas se realizarán en grupos de 2 personas. Existirán varias prácticas relacionadas con los distintos temas de la asignatura.

Se realizarán tutorías para ayudar en una comprensión mas personalizada de los temas teóricos y prácticos

SISTEMA DE EVALUACIÓN

| | |
|--|----|
| Peso porcentual del Examen/Prueba Final: | 20 |
| Peso porcentual del resto de la evaluación: | 80 |

A) Prueba final (examen): 20% (2 puntos)

B) Prácticas a desarrollar durante el cuatrimestre: 80% (8 puntos)

La calificación final de la asignatura se obtendrá sumando la calificación obtenida en los dos apartados. El examen no tiene nota mínima. La asignatura se aprueba si $A+B \geq 5$.

La principal herramienta para las prácticas será el lenguaje R. Se explicará tanto usando librerías de R, como usando la librería MLR (mlr3), que integra en una librería todas las tareas de aprendizaje automático. También se hará una breve introducción al aprendizaje automático usando Python y la librería scikit-learn.

BIBLIOGRAFÍA BÁSICA

- Aurélien Géron Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly Media, 2019
- Brett Lantz Machine Learning with R, Packt Publishing, 2019
- Hefin I. Rhy Machine Learning with R, the tidyverse, and mlr, Manning Publications, 2019
- Max Kuhn Applied Predictive Modeling, Springer, 2013

BIBLIOGRAFÍA COMPLEMENTARIA

- Hadley Wickham, Garrett Grolemund, R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, O'Reilly Media, 2016

- Hadley Wickham, Garrett Golemund, R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, O'Reilly Media, 2016

RECURSOS ELECTRÓNICOS BÁSICOS

- MLR . Machine Learning in R: <https://mlr3.mlr-org.com/>
- Scikit-learn team . Scikit-learn: <https://scikit-learn.org/>