A7. Métodos supervisados

MÁSTER UNIVERSITARIO EN INVESTIGACIÓN EN INTELIGENCIA ARTIFICIAL

UNIVERSIDAD INTERNACIONAL MENÉNDEZ PELAYO

Másteres Universitarios

Este documento puede utilizarse como documentación de referencia de esta asignatura para la solicitud de reconocimiento de créditos en otros estudios. Para su plena validez debe estar sellado por la Secretaría de Estudiantes UIMP.

DATOS GENERALES

Título asignatura

A7. Métodos supervisados

Código asignatura

102123

Curso académico

2016-17

Planes donde se imparte

MÁSTER UNIVERSITARIO EN INVESTIGACIÓN EN INTELIGENCIA ARTIFICIAL

Créditos ECTS

4,5

Carácter de la asignatura

OPTATIVA

Duración

Anual

Idioma

Castellano

CONTENIDOS

Contenidos

Desde sus inicios la Inteligencia Artificial ha prestado atención a cómo hacer que aprendan los ordenadores, más que a programar tareas simples. Este campo, denominado aprendizaje automático, se aplica a tareas que se desea que hagan los ordenadores y sobre las que tenemos más datos que conocimientos detallados de cómo han de llevarse a cabo.

Se presentan métodos de aprendizaje supervisado que son altamente utilizados en aplicaciones reales cuyo planteamiento es un problema de optimización y cuyo fin es resolver una tarea de clasificación, regresión y ranking. Se incluyen métodos tales como Regresión Logística, modelos Bayesianos, Redes de Neuronas Artificiales, Máquinas de Vectores Soporte (SVM) y métodos de Factorización de matrices. También se estudiarán en esta materia las técnicas de meta-aprendizaje:

- Aprendizaje como optimización.
- Regresión y Clasificación.
- Regresión Logística, Métodos Bayesianos, Redes neuronales, Métodos Kernel (SVM), Modelos lineales generalizados, Factorización.
- Meta-aprendizaje: boosting, bagging, random-forest.

Unidades

1. Módulo 1: Introducción a la clasificación supervisada y aplicaciones

2. Módulo 2: Principales algoritmos de aprendizaje supervisado

3. Módulo 3: Uso de tests estadísticos para la comparativa de modelos supervisados

4. Ejercicio final: flujo de análisis en clasificación supervisada

Elabora y estudia el material de manera secuencial, siguiendo el orden de unidades-temas propuesto. Realiza los ejercicios de cada tema-unidad tras estudiar la teoría correspondiente a éste. Anunciaremos los plazos de entrega de los ejercicios durante el curso.

RESULTADOS DE APRENDIZAJE Y DE FORMACIÓN

Generales

CG1 - Entender los conceptos, los métodos y las aplicaciones de la inteligencia artificial.

CG3 - Gestionar de manera inteligente los datos, la información y su representación.

Específicas

CE2 - Aplicar las técnicas de aprendizaje automático utilizando la metodología de validación y presentación de resultados más apropiada en cada caso.

CE5 - Analizar las fuentes documentales propias del ámbito de la investigación en Inteligencia Artificial para poder determinar cuáles de ellas son relevantes en la resolución de problemas concretos.

PLAN DE APRENDIZAJE

Actividades formativas

A1 - **Sesiones presenciales virtuales**: visionado inicial del material audiovisual (vídeos introductorios, presentaciones, animaciones) que se elabore en cada una de las materias y que servirán presentación de cada uno de los temas a los estudiantes (12 horas - 100% presencialidad).

A2 - **Trabajos individuales**: realización de ejercicios, resolución de problemas, realización de prácticas y/o trabajos/proyectos individuales (17 horas - 0% presencialidad).

A3 - **Trabajo autónomo**: estudio del material básico, lecturas complementarias y otros contenidos y estudio (72 horas - 0% presencialidad).

A4 - **Foros y chats**: lanzamiento de cuestiones y temas para la discusión general (5,5 horas - 0% presencialidad).

A5 - **Tutorías**: consultas y resolución de dudas, aclaraciones, etc (6 horas - 100% presencialidad).

SISTEMA DE EVALUACIÓN

Descripción del sistema de evalución

E1 - **Valoración de los cuestionarios de evaluación**: los estudiantes realizarán por cada unidad didáctica un cuestionario de evaluación que será objeto de puntuación en la nota final (ponderación mínima 20% y máxima 40%).

E2 - **Valoración de la participación en foros y chats**: se valorará el nivel de participación/debate de los estudiantes que contará para la nota final (ponderación mínima 10% y máxima 20%).

E3 - **Valoración de los trabajos individuales**: se valorarán los problemas, proyectos, trabajos realizados y entregados a través de la plataforma, y apoyado en los casos que sea necesario (sobre todo cuando se trate de desarrollo de código) por plataformas de gestión de código como GitHub. También se incluirá el video que el alumno deberá enviar al profesor para cada asignatura (ponderación mínima 40% y máxima 70%).

PROFESORADO

Profesor responsable

Larrañaga Múgica, Pedro

Licenciado en Matemáticas. Doctor en Informática.
Catedrático de Universidad (ciencias de la Computación e Inteligencia Artificial).
Universidad Politécnica de Madrid.

Profesorado

Inza Cano, Iñaki

Doctor en Informática.
Profesor Pleno -- equivalente laboral a Catedrático de Universidad.
Área: Ciencias de la Computación e Inteligencia Artificial.
UPV-EHU.

Mihaljevic , Bojan

Doctor en Inteligencia Artificial.
Profesor Permanente Laboral.
Universidad Politécnica de Madrid.

Bielza Lozoya, Concha

Doctora en Informática.
Catedrática de Universidad (área de Estadística e Investigación Operativa).
Universidad Politécnica de Madrid / Escuela Técnica Superior de Ingenieros Informáticos.

HORARIO

Horario

Las sesiones se desarrollarán en diciembre de 2016.

BIBLIOGRAFÍA Y ENLACES RELACIONADOS

Bibliografía

Bibliografía

M. Fernández-Delgado, E. Cernadas, S. Barro, D. Amorim (2014). Do we need hundreds of classifiers to solve real world classification problems? Journal of Machine Learning Research 15, 3133-3181. [link]

G. Brown, A. Pocock, M-J. Zhao, M. Luján (2012). Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. Journal of Machine Learning Research 13, 27-66. [link]

C. Bielza, P. Larrañaga (2014). Discrete Bayesian network classifiers: a survey. ACM Computing Surveys 47(1), 1-43. [link]

Y. Saeys, I. Inza, P. Larrañaga (2007). A review of feature selection techniques in bioinformatics. Bioinformatics 23(19), 2507-2517. [link]

M. Kuhn, K. Johnson (2013). Applied Predictive Modeling. Springer. [link]

Alpaydin. Introduction to Machine Learning. MIT Press.

R. Duda, P.E. Hart, D.G. Stork. Pattern Classification. Wiley. 2001.

J. Hernández-Orallo, M.J. Ramírez, C. Ferri. Introducción a la Minería de Datos. Pearson Educación. 2004.

L. Kuncheva. Combining Pattern Classifiers. Wiley. 2004.

S. Sharma. Applied Multivariate Techniques. Wiley. 1996.

A. Webb. Statistical Pattern Recognition. Wiley. 2002.

I. Witten, E. Frank. Data Mining. Morgan Kaufmann. 2ª ed. 2005.

Castillo, E., Gutiérrez, J.M., Hadi, A.S. (1997) Expert Systems and Probabilistic Network Models. Springer, New York. Versión en español, disponible en la red: Sistemas Expertos y Modelos de Redes Probabilísticas, Academia de Ingeniería, Madrid.

Koller, D, Friedman, N. (2009) Probabilistic Graphical Models Principles and Techniques. The MIT Press.

Koski, T., Noble, J. (2009) Bayesian Networks: An Introduction, Wiley.

Murphy, K.P. (2012) Machine Learning: A Probabilistic Perspective. The MIT Press.

Neapolitan, R., (2004) Learning Bayesian Networks, Prentice Hall.

Webs-enlaces

Software

- paquete caret (R): "classification and regression training with R" [link] [link CRAN]
- bnClassify (R): "learning discrete Bayesian network classifiers from data" [link en GitHub]
- a list of classification software for data mining and analytics [link]
- Bayesian networks and Bayesian classifier software [link]
- MEKA: a multilabel extension to WEKA [link]
- MULAN: a Java library for multilabel learning [link]

Lecturas - Reading

- 60+ free books on big data, data science, data mining, machine learning, python, R and more. [link]

Colección con aplicaciones de la minería de datos [link]