



Teaching Guide

Teaching Guide				
Identifying Data				2021/22
Subject (*)	Machine Learning III		Code	614G02026
Study programme	Grao en Ciencia e Enxeñaría de Datos			
Descriptors				
Cycle	Period	Year	Type	Credits
Graduate	2nd four-month period	Third	Obligatory	6
Language	Spanish			
Teaching method	Face-to-face			
Prerequisites				
Department	Ciencias da Computación e Tecnoloxías da Información			
Coordinador	Bolón Canedo, Verónica	E-mail	veronica.bolon@udc.es	
Lecturers	Bolón Canedo, Verónica Cancela Barizo, Brais Eiras Franco, Carlos Fernández Blanco, Enrique	E-mail	veronica.bolon@udc.es brais.cancela@udc.es carlos.eiras.franco@udc.es enrique.fernandez@udc.es	
Web				
General description	Esta asignatura presenta unha visión avanzada e máis específica da aprendizaxe automática. No temario explícanse distintas técnicas e métodos, incluíndo técnicas de preprocesado, e métodos menos clásicos como a ordenación ou os problemas dunha clase. Na parte práctica realizarase a resolución de casos reais.			
Contingency plan	1. Modifications to the contents 2. Methodologies *Teaching methodologies that are maintained *Teaching methodologies that are modified 3. Mechanisms for personalized attention to students 4. Modifications in the evaluation *Evaluation observations: 5. Modifications to the bibliography or webgraphy			

Study programme competences

Code	Study programme competences
A24	CE24 - Comprensión e dominio das principais técnicas básicas e avanzadas de aprendizaxe automática, incluíndo as dedicadas ao tratamento de grandes volumes de datos.
A25	CE25 - Capacidade para identificar a adecuación de cada unha das técnicas de aprendizaxe automática á resolución dun problema, incluíndo os aspectos relacionados coa súa complexidade computacional ou a súa capacidade explicativa, de acordo aos requisitos establecidos.
A26	CE26 - Coñecemento das ferramentas informáticas actuais no campo da aprendizaxe automática, e capacidade para seleccionar a máis adecuada para a resolución dun problema.
B2	CB2 - Que os estudantes saiban aplicar os seus coñecementos ao seu traballo ou vocación dunha forma profesional e posúan as competencias que adoitan demostrarse por medio da elaboración e defensa de argumentos e a resolución de problemas dentro da súa área de estudo
B3	CB3 - Que os estudantes teñan a capacidade de reunir e interpretar datos relevantes (normalmente dentro da súa área de estudo) para emitir xuízos que inclúan unha reflexión sobre temas relevantes de índole social, científica ou ética



B4	CB4 - Que os estudantes poidan transmitir información, ideas, problemas e solucións a un público tanto especializado como non especializado
B7	CG2 - Elaborar adecuadamente e con certa orixinalidade composicións escritas ou argumentos motivados, redactar plans, proxectos de traballo, artigos científicos e formular hipóteses razoables.
B8	CG3 - Ser capaz de manter e estender formulacións teóricas fundadas para permitir a introdución e explotación de tecnoloxías novas e avanzadas no campo.
B9	CG4 - Capacidade para abordar con éxito todas as etapas dun proxecto de datos: exploración previa dos datos, preprocesado, análise, visualización e comunicación de resultados.
B10	CG5 - Ser capaz de traballar en equipo, especialmente de carácter multidisciplinar, e ser hábiles na xestión do tempo, persoas e toma de decisións.
C1	CT1 - Utilizar as ferramentas básicas das tecnoloxías da información e as comunicacións (TIC) necesarias para o exercicio da súa profesión e para a aprendizaxe ao longo da súa vida.
C4	CT4 - Valorar a importancia que ten a investigación, a innovación e o desenvolvemento tecnolóxico no avance socioeconómico e cultural da sociedade.

Learning outcomes			
Learning outcomes		Study programme competences	
Coñecer e saber aplicar técnicas avanzadas de preprocesado de datos, incluíndo as de redución da dimensión ou de tratamento de valores ausentes		A24	B3 B8 B9
Coñecer as técnicas máis representativas de aprendizaxe para os problemas clásicos de clasificación, regresión e agrupación, e outros menos clásicos como problemas de ordenación, problemas de unha clase ou multitarea		A24	B3 B8 B9
Coñecer as técnicas máis representativas e actuais de aprendizaxe non supervisado, semisupervisado e supervisado, con e sen reforzo		A24	B3 B8 B9
Identificar as técnicas apropiadas de análise de datos segundo o problema		A25	B2 B3 B4 B7 B8 B10
Manexar as ferramentas e contornas de traballo máis actuais no ámbito da aprendizaxe automática		A26	C1

Contents	
Topic	Sub-topic
1. Técnicas avanzadas de preprocesamento de datos	1.1. Tratamento de datos sesgados e ausentes 1.2. Métodos de redución da dimensión.
2. Modelos combinados (Ensemble)	2.1. Métodos de combinación de modelos: Voting, Bagging, Boosting... 2.2. Bosques Aleatorios
3. Aprendizaxe por reforzo	3.1. Baseada en modelos 3.2. Baseada en diferenzas temporais
4. Aprendizaxe semisupervisada	4.1. Modelos xenerativos 4.2. Modelos baseados en grafo
5. Métodos de clasificación dunha clase	5.1. Baseados en densidade 5.2. Baseados en reconstrución 5.3. Discriminativos



6. Outras aproximacións	6.1. Algoritmos de ranking 6.2. Cuantificación 6.3. Aprendizaxe multitarefa
-------------------------	---

Planning				
Methodologies / tests	Competencies	Ordinary class hours	Student?s personal work hours	Total hours
Guest lecture / keynote speech	A24 A25 B2 B3 B4 B8 C4 C1	21	42	63
Laboratory practice	A24 A25 A26 B2 B3 B4 B7 B9 B10 C1	21	42	63
Objective test	A24 A25 B2 B3 B4 B7 B8 B9 C1 C4	2	20	22
Personalized attention		2	0	2
(*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.				

Methodologies	
Methodologies	Description
Guest lecture / keynote speech	Impartición teórica da materia da asignatura
Laboratory practice	Resolver problemas prácticos mediante o uso das distintas técnicas que se explicarán nas clases de teoría
Objective test	Proba de avaliación escrita na que o alumno deberá demostrar os coñecementos adquiridos na asignatura

Personalized attention	
Methodologies	Description
Laboratory practice	Realización do traballo práctico co asesoramento do profesor. Redacción de documentos de compendio dos resultados na forma de memorias ou artigos, así como a presentación dos resultados có profesor ou en sesións públicas dentro da clase.

Assessment			
Methodologies	Competencies	Description	Qualification
Laboratory practice	A24 A25 A26 B2 B3 B4 B7 B9 B10 C1	Resolución de problemas do mundo real utilizando a metodoloxía, para o cal se utilizarán varias técnicas explicadas en teoría, e estimularase ao alumno a xerar novas ideas para a resolución destes problemas.	50
Objective test	A24 A25 B2 B3 B4 B7 B8 B9 C1 C4	Preguntas sobre os contidos da asignatura (que poden ser de tipo test ou problemas para resolver), baseada nas distintas técnicas avanzadas de aprendizaxe automática e as súas aplicacións.	50

Assessment comments



Para superar a materia, o alumno deberá obter unha calificación mínima de 5 sobre 10 no resultado de combinar as calificacións da proba obxectiva e as prácticas de laboratorio. Ademais, o alumno deberá obter unha nota mínima de 2 sobre 5 puntos na proba obxectiva. Se non obtén esta nota mínima, a nota da materia será a correspondente á nota da proba obxectiva.

O traballo entregado deberá ser orixinal do alumno. De acordo ao artigo 14, apartado 4, da normativa, a entrega de traballos non orixinais ou con partes duplicadas (sexa por copias entre compañeiros ou por obtención doutras fontes...) levará unha nota global de SUSPENSO NA CONVOCATORIA ANUAL, tanto para o/a estudante que presente material copiado como a quen o facilitase.

Sobre a responsabilidade compartida dos traballos en grupo.

Nas actividades que se levan a cabo en grupos, tales como as prácticas, todos os membros do grupo serán responsables solidarios do traballo realizado e entregado, así como das consecuencias que se deriven do incumprimento das normas de autoría do mesmo.

Na segunda oportunidade, mantense a nota obtida nas prácticas de laboratorio. Opcionalmente, no seu lugar habilitarase unha entrega adicional para un traballo. A calificación deste traballo substituirá a nota de prácticas da primeira oportunidade. A entrega dun novo traballo implica a perda da calificación anterior independentemente de que esta fora superior. No que respecta ás prácticas que se realizan dentro das horas de laboratorio, estas non se poden recuperar xa que son froito da avaliación continua do traballo durante os créditos de práctica da materia. O alumno pode volver a facer o exame da proba obxectiva, sendo os criterios para obter a nota os indicados ao principio deste apartado.

Aqueles alumnos con matrícula a tempo parcial deberán entregar os traballos en data ao igual que os alumnos a tempo completo. É recomendable a súa asistencia ás clases de prácticas.

A maiores, as titorías considéranse unha parte importante dentro do desenvolvemento da asignatura. Están orientadas de tal maneira que os/as estudantes teñan e/ou poidan consultar distintas cuestións como:

1. Posibilidades de desenvolvemento profesional
2. Problemas no desenvolvemento das prácticas
3. Maneiras de enfocar/organizar as prácticas
4. Resolución de dúbidas sobre as cuestións teóricas

Debido a configuración baseada na non presencialidade das mesmas por parte dos centros, pedirase aos/ás estudantes que soliciten cita aos profesores responsables para realizar videochamadas por Teams dentro dos horarios de titorías do profesorado establecidos en espazos.udc.es.

Sources of information

Basic	- Kuncheva L.I. (2014). Combining pattern classifiers: methods and algorithms. John Wiley & Sons - Guyon, I., Gunn, S., Nikravesh, M., & Zadeh, L. A. (Eds.) (2008). Feature extraction: foundations and applications. Springer - Chapelle, O., Scholkopf, B., & Zien, A. (Eds.) (2006). Semi-Supervised learning. MIT Press - Tax, D. (2001). One-class classification: Concept-learning in the absence of counter-examples (Ph.D. thesis). The Netherlands: University of Delft (http://homepage.tudelft.nl/n9d04/thesis.pdf) - Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction.. MIT Press
Complementary	

Recommendations

Subjects that it is recommended to have taken before

Machine Learning I/614G02019
 Design and Analysis of Algorithms/614G02011
 Statistical Modeling of High Dimensional Data/614G02013
 Fundamentals of Programming II/614G02009
 Fundamentals of Programming I/614G02004
 Statistical Inference/614G02007

Subjects that are recommended to be taken simultaneously



Statistical Analysis of Dependent Data/614G02022
Subjects that continue the syllabus
Large Scale Machine Learning/614G02032 Numerical Methods for Data Science/614G02033 Image, Video and Audio Processing/614G02028
Other comments

(*)The teaching guide is the document in which the URV publishes the information about all its courses. It is a public document and cannot be modified. Only in exceptional cases can it be revised by the competent agent or duly revised so that it is in line with current legislation.