



## Teaching Guide

Identifying Data				2022/23
Subject (*)	Statistical Learning	Code	614493125	
Study programme	Mestrado Universitario en Técnicas Estadísticas (Plan 2019)			
Descriptors				
Cycle	Period	Year	Type	Credits
Official Master's Degree	1st four-month period	Second	Optional	5
Language	Spanish			
Teaching method	Face-to-face			
Prerequisites				
Department	Matemáticas			
Coordinador	Fernández Casal, Rubén	E-mail	ruben.fcasal@udc.es	
Lecturers	Fernández Casal, Rubén	E-mail	ruben.fcasal@udc.es	
Web	rubenfcasal.github.io/aprendizaje_estadistico			
General description	Nesta materia preténdese achegar ao alumno ao modelado e resolución de problemas de aprendizaxe estatística procedentes de aplicacións reais			

## Study programme competences / results

Code	Study programme competences / results
A16	CE1 - Coñecer, identificar, modelar, estudar e resolver problemas complexos de estatística e investigación operativa, nun contexto científico, tecnolóxico ou profesional, xurdidos en aplicacións reais.
A17	CE2 ? Desenvolver autonomía para a resolución práctica de problemas complexos surtidos en aplicación reais e para a interpretación dos resultados cara á axuda na toma de decisións.
A18	CE3 - Adquirir coñecementos avanzados dos fundamentos teóricos subxacentes ás distintas metodoloxías da estatística e a investigación operativa, que permitan o seu desenvolvemento profesional especializado.
A19	CE4 - Adquirir as destrezas necesarias no manexo teórico-práctico da teoría de probabilidade e as variables aleatorias que permitan o seu desenvolvemento profesional no eido científico/académico, tecnolóxico ou profesional especializado e multidisciplinar.
A20	CE5 - Profundizar no coñecemento dos fundamentos teórico-prácticos especializados de modelado e estudo de distintos tipos de relacións de dependencia entre variables estatísticas.
A21	CE6 - Adquirir coñecementos teórico-prácticos avanzados de distintas técnicas matemáticas, orientadas especificamente á axuda na toma de decisións, e desenvolver a capacidade de reflexión para avaliar e decidir entre distintas perspectivas en contextos complexos.
A22	CE7 - Adquirir coñecementos teórico-prácticos avanzados de distintas técnicas de optimización matemática, tanto en contextos unipersoais como multipersoais, e saber aplicalos con autonomía suficiente nun contexto científico, tecnolóxico ou profesional.
A23	CE8 - Adquirir coñecementos teórico-prácticos avanzados das técnicas destinadas á realización de inferencias e contrastes relativos a variables e parámetros dun modelo estatístico, e saber aplicalos con autonomía suficiente nun contexto científico, tecnolóxico ou profesional.
A24	CE9 - Coñecer e saber aplicar con autonomía en contextos científicos, tecnolóxicos ou profesionais, técnicas de aprendizaxe automático e técnicas de análise de datos de alta dimensión (big data).
A25	CE10 - Adquirir coñecementos avanzados sobre metodoloxías para a obtención e o tratamento de datos derivados de distintas fontes, como enquisas, internet, ou entornos ?na nube&quot;.
B1	CB6 - Posuír e comprender coñecementos que acheguen unha base ou oportunidade de ser orixinais no desenvolvemento e/ou aplicación de ideas, a miúdo nun contexto de investigación
B2	CB7 - Que os estudantes saiban aplicar os coñecementos adquiridos e a súa capacidade de resolución de problemas en ámbitos novos ou pouco coñecidos dentro de contextos máis amplos (ou multidisciplinares) relacionados coa súa área de estudo
B3	CB8 - Que os estudantes sexan capaces de integrar coñecementos e enfrontarse á complexidade de formular xuízos a partir dunha información que, sendo incompleta ou limitada, inclúa reflexións sobre as responsabilidades sociais e éticas vinculadas á aplicación dos seus coñecementos e xuízos
B4	CB9 - Que os estudantes saiban comunicar as súas conclusións e os coñecementos e razóns últimas que as sustentan a públicos especializados e non especializados dun modo claro e sen ambigüidades



B5	CB10 - Que os estudantes posúan as habilidades de aprendizaxe que lles permitan continuar estudando dun modo que haberá de ser en gran medida autodirixido ou autónomo.
B17	CG1 - Coñecer, comprender e saber aplicar os principios, metodoloxías e novas tecnoloxías na estatística e a investigación operativa en contextos científico/académicos, tecnolóxicos ou profesionais especializados e multidisciplinares, así como adquirir as destrezas e competencias descritas nos obxectivos xerais do título.
B18	CG2 - Desenvolver autonomía para identificar, modelar e resolver problemas complexos da estatística e da investigación operativa en contextos científico/académicos, tecnolóxicos ou profesionais especializados e multidisciplinares.
B19	CG3 - Desenvolver a capacidade para realizar estudos e tarefas de investigación e transmitir os resultados a públicos especializados, académicos e xeneralistas.
B20	CG4 - Integrar coñecementos avanzados e enfrontarse á toma de decisións a partir de información científica e técnica.
B21	CG5 - Desenvolver a capacidade de aplicación de algoritmos e técnicas de resolución de problemas complexos no eido da estatística e a investigación operativa, manexando o software especializado axeitado.
C11	CT1 - Desenvolver firmes capacidades de razoamento, análise crítica e autocrítica, así como de argumentación e de síntese, contextos especializados e multidisciplinares.
C12	CT2 - Desenvolver destrezas avanzadas no manexo de Tecnoloxías da Información e a Comunicación (TIC), tanto para a obtención de información como para a difusión do coñecemento, nun ámbito científico/académico, tecnolóxico ou profesional especializado e multidisciplinar.
C13	CT3 - Ser capaz de resolver problemas complexos en novos escenarios mediante a aplicación integrada dos coñecementos.
C14	CT4 - Desenvolver unha sólida capacidade de organización e planificación do estudo, asumindo a responsabilidade do seu propio desenvolvemento profesional, para a realización de traballos en equipo e de xeito autónomo.
C15	CT5 - Desenvolver capacidades para o aprendizaxe e a integración no traballo en equipos multidisciplinares, nos ámbitos científico/académico, tecnolóxico e profesional.

## Learning outcomes

Learning outcomes	Study programme competences / results		
Ser capaz de identificar e modelar problemas complexos de aprendizaxe estatística xurdidos en aplicacións reais	AC16	BJ1	CJ11
	AC17	BJ2	CJ12
	AC18	BJ3	CJ13
	AC19	BJ4	CJ14
	AC20	BJ5	CJ15
	AC21	BJ17	
	AC22	BJ18	
	AC23	BJ19	
	AC24	BJ20	
	AC25	BJ21	
Adquirir coñecementos avanzados sobre aprendizaxe estatística tanto supervisado como non supervisado	AC16	BJ1	CJ11
	AC17	BJ2	CJ12
	AC18	BJ3	CJ13
	AC19	BJ4	CJ14
	AC20	BJ5	CJ15
	AC21	BJ17	
	AC22	BJ18	
	AC23	BJ19	
	AC24	BJ20	
	AC25	BJ21	



Coñecer os principios xerais das novas metodoloxías de aprendizaxe supervisada para clasificación e regresión	AC16 AC17 AC18 AC19 AC20 AC21 AC22 AC23 AC24 AC25	BJ1 BJ2 BJ3 BJ4 BJ5 BJ17 BJ18 BJ19 BJ20 BJ21	CJ11 CJ12 CJ13 CJ14 CJ15
Identificar e adquirir destreza no uso das principais ferramentas de R de aprendizaxe estatística	AC16 AC20 AC24 AC25	BJ2 BJ3 BJ4 BJ5 BJ20 BJ21	CJ12 CJ13 CJ14 CJ15
Saber analizar datos utilizando técnicas de Aprendizaxe Estatística en contextos multidisciplinares	AC16 AC17 AC18 AC20 AC21 AC24 AC25	BJ1 BJ2 BJ3 BJ4 BJ5 BJ17 BJ18 BJ19 BJ20 BJ21	CJ11 CJ12 CJ13 CJ14 CJ15
Ser capaz de aplicar con autonomía os resultados adquiridos na análise de datos masivos ou de alta dimensión	AC16 AC17 AC18 AC20 AC21 AC24 AC25	BJ1 BJ2 BJ3 BJ5 BJ17 BJ18 BJ19 BJ20 BJ21	CJ11 CJ12 CJ13 CJ14 CJ15
Saber comunicar os resultados propios da Aprendizaxe Estatística a un público especializado ou non especializado		BJ4 BJ19	CJ11 CJ12 CJ14 CJ15

Contents	
Topic	Sub-topic
Introdución á aprendizaxe estatística	Aprendizaxe estatística e aprendizaxe automática. Métodos de aprendizaxe estatística: supervisados e non supervisados. Construción e avaliación dos modelos. A maldición da dimensionalidade. Análise e interpretación dos modelos. Paquetes de R.
Métodos de aprendizaxe supervisada para clasificación	Métodos baseados en árbores. Bagging, bosques aleatorios e boosting. Máquinas de soporte vectorial (support vector machines). Outros métodos de clasificación.



Métodos de aprendizaxe supervisada para regresión	Modelos lineais xeneralizados e modelos aditivos. Métodos non paramétricos. Selección de variables e métodos de regularización: regresión contraída (ridge) e lasso. Métodos de redución da dimensión: compoñentes principais e mínimos cadrados parciais. Regresión por projection pursuit. Redes neuronais. Métodos colaborativos.
---	--

Planning				
Methodologies / tests	Competencies / Results	Teaching hours (in-person & virtual)	Student?s personal work hours	Total hours
Guest lecture / keynote speech	A16 A17 A18 A19 A20 A21 A22 A23 A24 A25 B1 B2 B3 B4 B5 B17 B18 B19 B20 B21 C11 C12 C13 C14 C15	15	25	40
ICT practicals	A16 A17 A18 A19 A20 A21 A22 A23 A24 A25 B1 B2 B3 B4 B5 B17 B18 B19 B20 B21 C11 C12 C13 C14 C15	20	51	71
Mixed objective/subjective test	A16 A17 A24 B1 B2 B3 B4 B5 B17 B18 B19 B20 B21 C11 C12 C13 C14	4	0	4
Personalized attention		10	0	10

(\*)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	Presentación dos aspectos teóricos relevantes da materia, de modo que o alumnado poida abordar a realización das prácticas a través de TIC. Fomentárase en todo momento a participación e o debate
ICT practicals	Aplicación práctica das metodoloxías explicadas nas sesións maxistras. Farase uso de ferramentas de software libre, principalmente da contorna estatística R, e tamén de recursos web
Mixed objective/subjective test	O alumno deberá demostrar o seu dominio dos aspectos teóricos da materia e a súa capacidade para a resolución de problemas

Personalized attention	
Methodologies	Description
ICT practicals Guest lecture / keynote speech	Tanto nas sesións maxistras como nas sesións prácticas será importante atender persoalmente aos alumnos ante as posibles dúbidas que poidan xurdir. Esta atención servirá ao profesor para detectar posibles problemas na metodoloxía empregada para impartir a materia e, ademais, aos alumnos para consolidar os coñecementos e para expresar as súas inquietudes acerca da materia

Assessment			
Methodologies	Competencies / Results	Description	Qualification



ICT practicals	A16 A17 A18 A19 A20 A21 A22 A23 A24 A25 B1 B2 B3 B4 B5 B17 B18 B19 B20 B21 C11 C12 C13 C14 C15	Ao longo do curso, o alumno demostrará o seu interese e dominio da materia, e a súa capacidade de traballo en equipo, mediante a realización de traballos, individuais ou en grupo. A cualificación obtida conservarase entre as dúas oportunidades da convocatoria de cada curso	50
Mixed objective/subjective test	A16 A17 A24 B1 B2 B3 B4 B5 B17 B18 B19 B20 B21 C11 C12 C13 C14	O exame final consistirá nunha proba escrita teórico-práctica	50

### Assessment comments

Na data que estableza o Máster de Técnicas Estatísticas na súa programación anual, o alumno realizará, por escrito, o exame final da materia (proba mixta), no que terá que responder a preguntas teóricas, resolver cuestións teórico-prácticas, e calcular a solución de diversos problemas. Para esta proba o alumno só poderá levar consigo o material que se autorice de forma expresa. Para aprobar a materia será necesario superar as dúas partes (prácticas a través de TIC e proba mixta). A oportunidade de xullo (segunda oportunidade) estará sometida aos mesmos criterios que a de xaneiro (primeira oportunidade). Os alumnos que se presenten á segunda oportunidade poderán optar entre manter a nota de prácticas a través de TIC ou volver examinarse tamén desta parte.

### Sources of information

<b>Basic</b>	<ul style="list-style-type: none"> <li>- Fernández-Casal, R. y Costa, J. (2020). Aprendizaje Estadístico. Libro online: <a href="https://rubenfcasal.github.io/aprendizaje_estadistico">https://rubenfcasal.github.io/aprendizaje_estadistico</a></li> <li>- James, G., Witten, D., Hastie, T. y Tibshirani, R. (2017). An Introduction to Statistical Learning: with Applications in R. Springer</li> <li>- Kuhn, M. y Johnson, K. (2013). Applied predictive modeling. Springer</li> <li>- Williams, G. (2011). Data Mining with Rattle and R. Springer</li> </ul>
<b>Complementary</b>	<ul style="list-style-type: none"> <li>- Burger, S. V. (2018). Introduction to Machine Learning with R: Rigorous Mathematical Analysis. O'Reilly</li> <li>- Efron, B. y Hastie, T. (2016). Computer Age Statistical Inference. Cambridge University Press</li> <li>- Hastie, T., Tibshirani, R. y Friedman, J. (2017). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer</li> <li>- Hastie, T., Tibshirani, R. y Wainwright, M. (2015). Statistical Learning with Sparsity: the Lasso and Generalizations. CRC press</li> <li>- Torgo, L. (2011). Data Mining with R: Learning with Case Studies. Chapman &amp; Hall/CRC Press</li> </ul>

### Recommendations

#### Subjects that it is recommended to have taken before

Generalized Regression and Mixed Models/614493110

Regression Models/614493105

#### Subjects that are recommended to be taken simultaneously

#### Subjects that continue the syllabus

### Other comments

É recomendable que o alumno posúa coñecementos básicos de regresión e de estatística multivariante. Dispoñer de coñecementos xerais da contorna estatística R facilitará o aproveitamento das clases e a realización das prácticas. Aconséllase participar activamente no proceso de aprendizaxe da materia: asistencia e participación ás clases teóricas e prácticas, utilización de horas de tutorías e a realización dun esforzo responsable de traballo e asimilación persoal dos métodos estudados.



(\*)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.