



Guía Docente				
Datos Identificativos				2023/24
Asignatura (*)	IA en Contornos Big Data		Código	614544016
Titulación	Máster Universitario en Intelixencia Artificial			
Descritores				
Ciclo	Período	Curso	Tipo	Créditos
Mestrado Oficial	1º cuatrimestre	Segundo	Optativa	6
Idioma	Inglés			
Modalidade docente	Presencial			
Prerrequisitos				
Departamento	Ciencias da Computación e Tecnoloxías da Información			
Coordinación	Bolón Canedo, Verónica		Correo electrónico	veronica.bolon@udc.es
Profesorado	Alonso Ríos, David		Correo electrónico	david.alonso@udc.es
	Bolón Canedo, Verónica			veronica.bolon@udc.es
Web				
Descrición xeral	<p>A cada vez maior cantidade de información accesible a través de Internet fai que o procesamiento eficiente de grandes cantidades de datos sexa cada vez de maior interese. Isto levou ao desenvolvemento de novas técnicas de almacenamento e procesamiento de inxentes cantidades de información, técnicas que se adaptan de forma natural aos sistemas distribuídos.</p> <p>O obxectivo principal desta materia é proporcionar aos estudantes os coñecementos e habilidades necesarios para comprender, desenvolver e aplicar técnicas de intelixencia artificial (IA) en contornas de Big Data.</p>			

Competencias / Resultados do título	
Código	Competencias / Resultados do título
A11	CE10 - capacidade para a construción, validación e aplicación dun modelo estocástico dun sistema real a partir dos datos observados e o análise crítico dos resultados obtidos para seleccionar as máis axeitadas para a resolución de problemas
A12	CE11 - Comprensión e dominio das principais técnicas e ferramentas de análise de datos, tanto dende o punto de vista estatístico como da aprendizaxe automática, incluíndo as dedicadas ao tratamento de grandes volúmenes de datos, e capacidade para seleccionar as máis axeitadas para a resolución de problemas
A13	CE12 - capacidade para plantexar, formular e resolver todas as etapas dun proxecto de datos, incluíndo a comprensión e dominio de fundamentos e técnicas básicas para a búsqueda e o filtrado de información en grandes coleccións de datos
A16	CE15 - coñecemento das ferramentas informáticas no campo da aprendizaxe automática, e capacidade para seleccionar a máis axeitada para a resolución dun problema
B2	CG02 - Abordar con éxito todas as etapas dun proxecto de Intelixencia Artificial
B3	CG03 - Buscar e seleccionar a información útil necesaria para resolver problemas complexos, manexando con soltura as fontes bibliográficas do campo
B4	CG04 - Elaborar axeitadamente e con certa orixinalidade composicións escritas ou argumentos motivados, redactar plans, proxectos de traballo, artigos científicos e formular hipóteses razoables no campo
B5	CG05 - Traballar en equipo, especialmente de carácter multidisciplinar, e ser hábiles na xestión do tempo, persoas e toma de decisións
B6	CB01 - Poseer e comprender coñecementos que aporten unha base ou oportunidade de ser orixinais no desenvolvemento e/ou aplicación de ideas, a miúdo nun contexto de investigación
B7	CB02 - Que os estudantes saiban aplicar os coñecementos adquiridos e posúan capacidade de resolución de problemas en entornos novos ou pouco coñecidos dentro de contextos máis amplos (ou multidisciplinares) relacionados coa súa área de estudo
B8	CB03 - Que os estudantes sexan capaces de integrar coñecementos e enfrentarse 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
B9	CB04 - Que os estudantes saiban comunicar as súas conclusións e os coñecementos e razóns últimas que as sustentan a públicos especializados e non especializados dun xeito claro e sen ambigüidades



C3	CT03 - 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	CT04 - Desenvolverse para o exercicio dunha cidadanía respetuosa coa cultura democrática, os dereitos humanos e a perspectiva de xénero
C7	CT07 - Desenvolver a capacidade de traballar en equipos interdisciplinares ou transdisciplinares, para ofrecer propostas que contribúan a un desenvolvemento sostible ambiental, económico, político e social
C8	CT08 - Valorar a importancia que ten a investigación, a innovación e o desenvolvemento tecnolóxico no avance socioeconómico e cultural da sociedade
C9	CT09 - Ter a capacidade de xestionar tempos e recursos: desenvolver plans, priorizar actividades, identificar as críticas, establecer prazos e cumprilos

Resultados da aprendizaxe			
Resultados de aprendizaxe	Competencias / Resultados do título		
Coñecer as técnicas que permiten o deseño de técnicas de IA escalables a nivel software e de recursos hardware	AM10 AM11 AM12 AM15	BM2 BM7	CM3 CM4
Adquirir as competencias que permitan integrar gran volume e variedade de datos en proxectos de Big Data en IA	AM10 AM11 AM12 AM15	BM3 BM4 BM5 BM6 BM7 BM8 BM9	CM3 CM4 CM7 CM8 CM9
Coñecer as paradigmas de escalabilidade en algoritmos de aprendizaxe automática	AM10 AM11 AM12 AM15	BM2 BM3 BM4 BM5 BM6 BM7 BM8 BM9	CM3 CM4 CM7 CM8 CM9
Comprender, analizar e deseñar as infraestruturas necesarias para proxectos de IA en BigData: contorna local/nube e equipamento físico/virtual con sistemas de almacenamento de baixa latencia e sistemas de ficheiros distribuídos.	AM12 AM15	BM2 BM6 BM7 BM8	CM3 CM4 CM7 CM9
Coñecer as linguaxes, frameworks e compoñentes que nos permiten incrementar o rendemento nas infraestruturas hardware con CPU e GPU	AM11 AM15	BM3 BM7 BM8	CM3 CM4 CM7 CM9
Coñecer as técnicas que permiten, con baixa latencia, a visualización de datos en contornas con gran volume de información.	AM11 AM12 AM15	BM2 BM3 BM5 BM6 BM7 BM8 BM9	CM3 CM4 CM7 CM8 CM9



Usar e poder aplicar os KPI correctos en cada contorna.	AM10	BM2	CM3
	AM11	BM3	CM9
	AM15	BM7	
		BM8	

Contidos	
Temas	Subtemas
Introdución ao Big Data	Que é Big Data Aplicacións Big Data Analítica Big Data Problemática da análise de datos en contornas Big Data
Preparación e visualización de datos	Técnicas de preprocesado de datos Técnicas de visualización
Aprendizaxe federada	Aprendizaxe no borde Preservación da privacidade
Infraestructuras para o almacenamento e procesamento de Big Data: Apache Hadoop e Apache Spark	Procesamento distribuído e infraestructuras Aprendizaxe por lotes en plataformas paralelas e distribuídas Aprendizaxe distribuída en vertical e horizontal
Tratamento de datos en continuo	Aprendizaxe incremental Aprendizaxe en tempo real Problemas de cambio de concepto

Planificación				
Metodoloxías / probas	Competencias / Resultados	Horas lectivas (presenciais e virtuais)	Horas traballo autónomo	Horas totais
Prácticas a través de TIC	A11 A12 A13 A16 B2 B3 B4 B5 B6 B7 B8 B9 C3 C7 C8 C9	14	44	58
Traballos tutelados	A11 A12 B3 B4 B5 B6 B9 C4 C7 C8	7	20	27
Proba obxectiva	A11 A12 A13 B2 B6 B7 B8 B9 C4 C8 C9	2	20	22
Sesión maxistral	A11 A12 A13 A16 B2 B3 B4 B6 B8 B9 C4 C8 C9	21	20	41
Atención personalizada		2	0	2

\*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado

Metodoloxías	
Metodoloxías	Descrición
Prácticas a través de TIC	Clases prácticas na aula de informática, que permiten ao alumno familiarizarse desde un punto de vista práctico coas cuestións expostas nas clases teóricas.
Traballos tutelados	Aprendizaxe baseada en problemas, seminarios, estudo de casos ou proxectos, que permiten que o alumnado adquira determinadas competencias en base á resolución de exercicios, estudo de casos e realización de proxectos.
Proba obxectiva	Proba de avaliación escrita na que o alumno deberá demostrar os coñecementos adquiridos na asignatura
Sesión maxistral	Clases de teoría, nas que se expón o contido de cada tema. O alumno disporá de copias das transparencias con anterioridade e o profesor promoverá unha actitude activa, realizando preguntas que permitan aclarar aspectos concretos e deixando cuestións abertas para a reflexión do alumno.



## Atención personalizada

Metodoloxías	Descrición
Prácticas a través de TIC Traballos tutelados	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 co profesor ou en sesións públicas dentro da clase.

## Avaliación

Metodoloxías	Competencias / Resultados	Descrición	Cualificación
Prácticas a través de TIC	A11 A12 A13 A16 B2 B3 B4 B5 B6 B7 B8 B9 C3 C7 C8 C9	Avaliación de traballos prácticos: 50% dá nota avaliaranse as solucións propostas polo alumnado ás prácticas expostas. A avaliación de prácticas pode levar a cabo mediante unha corrección por parte do profesor, unha defensa da solución achegada por parte do alumno ante o profesor ou unha presentación oral da solución desenvolvida. Todos os traballos deberán ser entregados antes das datas que se especificarán e deberán cumprir uns requisitos mínimos de calidade para ser tidos en consideración. Valorarase o grao de cumprimento das especificacións, a metodoloxía e rigorosidade e a presentación de resultados	50
Proba obxectiva	A11 A12 A13 B2 B6 B7 B8 B9 C4 C8 C9	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

## Observacións avaliación

Para superar a materia, debe conseguirse unha puntuación total de 5 ou superior. É imprescindible para aprobar entregar todas as prácticas indicadas como obrigatorias.

Condición para cualificación de Non Presentado: non presentar ningunha práctica e non asistir ao exame final.

Os alumnos que non sexan de nova matrícula non conservan notas de cursos anteriores.

Oportunidade de recuperación (xullo) e extraordinaria:

A valoración será igual que na oportunidade ordinaria. Os alumnos que non entregaron os traballos propostos ao longo do cuadrimestre deberanos entregar antes da data establecida.

Condición para cualificación de Non Presentado: non presentar ningunha práctica e non asistir ao exame final.

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.

## Fontes de información

<b>Bibliografía básica</b>	- Apuntes proporcionados por el profesor- T. White, Hadoop: The Definitive Guide, 4th Edition, O'Reilly, 2015- B. Chambers, M. Zaharia, Spark: The Definitive Guide, O'Reilly, 2018- Apuntes proporcionados por el profesor- T. White, Hadoop: The Definitive Guide, 4th Edition, O'Reilly, 2015- B. Chambers, M. Zaharia, Spark: The Definitive Guide, O'Reilly, 2018
----------------------------	--



<b>Bibliografía complementaria</b>	- Karim, Md. Rezaul, Sridhar Alla. Scala and Spark for Big Data Analytics: Tame Big Data with Scala and Apache Spark! 1st edition. Birmingham: Packt, 2017.- Pentreath, Nick. Machine Learning with Spark Create Scalable Machine Learning Applications to Power a Modern Data-Driven Business Using Spark Packt Publishing Ltd., 2015.- Bowles, Michael. Machine Learning with Spark and Python: Essential Techniques for Predictive Analytics 2nd ed. Wiley, 2019- Karim, Md. Rezaul, Sridhar Alla. Scala and Spark for Big Data Analytics: Tame Big Data with Scala and Apache Spark! 1st edition. Birmingham: Packt, 2017.- Pentreath, Nick. Machine Learning with Spark Create Scalable Machine Learning Applications to Power a Modern Data-Driven Business Using Spark Packt Publishing Ltd., 2015.- Bowles, Michael. Machine Learning with Spark and Python: Essential Techniques for Predictive Analytics 2nd ed. Wiley, 2019
------------------------------------	--

## Recomendacións

### Materias que se recomenda ter cursado previamente

Fundamentos de IA/614544001  
Aprendizaxe Automática I/614544012  
Aprendizaxe Automática II/614544014

### Materias que se recomenda cursar simultaneamente

### Materias que continúan o temario

### Observacións

(\*)A Guía docente é o documento onde se visualiza a proposta académica da UDC. Este documento é público e non se pode modificar, salvo casos excepcionais baixo a revisión do órgano competente dacordo coa normativa vixente que establece o proceso de elaboración de guías