



Guía Docente

Datos Identificativos				
Asignatura (*)	IA en Contornos Big Data	Código	2024/25 614544016	
Titulación				
Descriptorios				
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	Bolón Canedo, Verónica Cancela Barizo, Brais	Correo electrónico	veronica.bolon@udc.es brais.cancela@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
--------	-------------------------------------

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 integran 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 AM11 AM15	BM2 BM3 BM7 BM8	CM3 CM9

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
Infraestruturas para o almacenamento e procesamento de Big Data	Paralelismo e sistemas distribuídos Computación de Altas Prestacións e Computación Big Data Apache Hadoop e MapReduce
Procesamento de datos a gran escala: Apache Spark	Procesamento en lote e procesamento continuo Arquitectura Spark Core (RDDs) e Spark SQL, DataSets & DataFrames Spark DataFrames
Aprendizaxe automática con Apache Spark	O fluxo de traballo en aprendizaxe automática Aprendizaxe automática supervisada e non supervisada Axuste de hiperparámetros, avaliación e pipelines

Planificación				
Metodoloxías / probas	Competencias / Resultados	Horas lectivas (presenciais e virtuais)	Horas traballo autónomo	Horas totais
Prácticas a través de TIC	A11 A12 A13 B2 B6 B3 B4 B5 B7 B8 B9 C3 C7 C8 C9	14	44	58
Proba obxectiva	A11 A12 A13 B2 B6 B7 B8 B9 C4 C8 C9	2	20	22



Aprendizaxe colaborativa	B3 B4 B5 B6 B8 B9 C4 C7 C8 C9	7	19	26
Sesión maxistral	A11 A12 A13 A16 B2 B3 B4 B6 B8 B9 C4 C8 C9	21	21	42
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.
Proba obxectiva	Proba de avaliación escrita na que o alumno deberá demostrar os coñecementos adquiridos na asignatura
Aprendizaxe colaborativa	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.
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
Aprendizaxe colaborativa Prácticas a través de TIC	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.

Avaliación			
Metodoloxías	Competencias / Resultados	Descrición	Cualificación
Aprendizaxe colaborativa	B3 B4 B5 B6 B8 B9 C4 C7 C8 C9	Avaliarase a realización de proxectos de aprendizaxe colaborativa, onde o alumnado fará traballos (preferiblemente en parellas ou grupos) para desenvolver en detalle un artigo científico relacionado cos temas vistos en teoría, e o presentará para toda a clase, onde se poderán facer preguntas. Estes proxectos poderán facerse nas horas de ensino non presencial e o seu obxectivo é profundizar nos contidos da materia, así como adquirir competencias de análise crítica, resumo, e presentación oral. Valorarase o grao de cumprimento das especificacións, a metodoloxía e rigorosidade e a presentación de resultados	5
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.	45



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. As entregas fóra de prazo non serán avaliadas.

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

Bibliografía básica	- Apuntamentos proporcionados polo profesorado- A. Polak, &i&t;Scaling Machine Learning with Spark&i&t;, O'Reilly, 2023- I. Triguero, M. Galar, &i&t;Large-Scale Data Analytics with Python and Spark&i&t;, Cambridge University Press, 2023- Apuntamentos proporcionados polo profesorado- A. Polak, &i&t;Scaling Machine Learning with Spark&i&t;, O'Reilly, 2023- I. Triguero, M. Galar, &i&t;Large-Scale Data Analytics with Python and Spark&i&t;, Cambridge University Press, 2023
Bibliografía complementaria	- T. White, &i&t;Hadoop: The Definitive Guide&i&t;, 4th Edition, O'Reilly, 2015- J. Damji, B. Wenig, T. Das and D. Lee. &i&t;Learning Spark&i&t;, 2nd Edition, O'Reilly, 2020- T. White, &i&t;Hadoop: The Definitive Guide&i&t;, 4th Edition, O'Reilly, 2015- J. Damji, B. Wenig, T. Das and D. Lee. &i&t;Learning Spark&i&t;, 2nd Edition, O'Reilly, 2020

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