



| Guía Docente | | | | |
|-----------------------|--|--------------------|--|----------|
| Datos Identificativos | | | | 2022/23 |
| Asignatura (*) | Aprendizaxe Automática a Gran Escala | Código | 614G02032 | |
| Titulación | | | | |
| Descritores | | | | |
| Ciclo | Período | Curso | Tipo | Créditos |
| Grao | 1º cuatrimestre | Cuarto | Optativa | 6 |
| Idioma | Castelán | | | |
| Modalidade docente | Presencial | | | |
| Prerrequisitos | | | | |
| Departamento | Ciencias da Computación e Tecnoloxías da Información | | | |
| Coordinación | Hernandez Pereira, Elena Maria | Correo electrónico | elena.hernandez@udc.es | |
| Profesorado | Alvarez Estevez, Diego Blanco Mallo, Eva Cancela Barizo, Brais Fontenla Romero, Oscar Hernandez Pereira, Elena Maria Morán Fernández, Laura | Correo electrónico | diego.alvareze@udc.es eva.blanco@udc.es brais.cancela@udc.es oscar.fontenla@udc.es elena.hernandez@udc.es laura.moranf@udc.es | |
| Web | campusvirtual.udc.gal | | | |
| Descrición xeral | <p>Nesta materia abórdase a área da aprendizaxe automática en contornas onde existe unha gran cantidade de datos a analizar. Neste contexto xorden certas problemáticas que fan que moitos dos sistemas de aprendizaxe clásicos non sexan directamente aplicables por motivos de complexidade computacional. A aprendizaxe automática a escala trata dous ámbitos de escalabilidade diferentes. A primeira é o adestramento dun modelo con grandes conxuntos de datos, que necesita as funcionalidades da escalabilidade sobre unha agrupación industrial de computadores para realizar o adestramento. O segundo céntrase na posta en operación do modelo adestrado de maneira que se poida escalar para cumprir as necesidades das aplicacións que o consomen.</p> <p>Coa aprendizaxe automática a gran escala, o enfoque desprázase cara aos datos e a tarefa. O tempo dedicado á tarefa e os datos é significativo e, a miúdo, moito máis grande do previsto. O obxectivo desta materia é proporcionar unha visión clara sobre os conceptos do "Big Data" no ámbito da aprendizaxe automática.</p> | | | |

| Competencias do título | |
|------------------------|------------------------|
| Código | Competencias do título |
| | |

| Resultados da aprendizaxe | | | |
|---|------------------------|-----------------|----------|
| Resultados de aprendizaxe | Competencias do título | | |
| Identificar e saber afrontar os problemas máis frecuentes relacionados coa explosión de datos, coñecidos como as "Vs do Big Data" e como algúns deles afectan á aplicación de técnicas de Aprendizaxe Automática. | A25 | B2 B9 B10 | C1 C4 |
| Coñecer os métodos e técnicas máis representativos e actuais de preprocesado de datos para tratar grandes volumes de datos. | A24 A25 A26 | B9 | C1 |
| Coñecer os métodos e técnicas máis representativas e actuais de Aprendizaxe Automática en contornas afectadas por problemas como o volume, a velocidade ou a privacidade dos datos. | A24 A25 A26 | B3 B10 | C1 C4 |
| Saber manexar as ferramentas e contornas de traballo máis actuais no ámbito da aprendizaxe automática para tratar grandes volumes de datos. | A24 | B2 B4 B7 | C1 |



| | | | |
|---|------------|---|----|
| Coñecer técnicas para a representación visual de datos complexos e saber utilizar ferramentas de visualización de datos para poder comunicar eficazmente os resultados das análises realizadas. | A24 A26 | B2 B3 B4 B7 B8 B9 B10 | C1 |
| Coñecer técnicas analíticas e escalables baseadas en grafos. | A24 A26 | B2 B8 B9 B10 | C1 |

| Contidos | |
|--|---|
| Temas | Subtemas |
| Problemáticas da análise de datos en contornas "Big Data"; | |
| Tratamento e visualización de grandes volumes de datos | Técnicas de preprocesado de datos Técnicas de visualización |
| Aprendizaxe distribuída | Aprendizaxe por lotes en plataformas paralelas e distribuídas Aprendizaxe distribuída en vertical e horizontal |
| Aprendizaxe federada | Aprendizaxe na contorna Preservación da privacidade |
| Tratamento de datos en continuo | Aprendizaxe incremental Aprendizaxe en tempo real Problemas de cambio de concepto |
| Analítica de datos baseada en grafos | |

| Planificación | | | | |
|--------------------------|-----------------------------------|-------------------|---|--------------|
| Metodoloxías / probas | Competencias | Horas presenciais | Horas non presenciais / traballo autónomo | Horas totais |
| Solución de problemas | A24 A25 A26 B2 B4 B7 B10 C1 C4 | 0 | 40 | 40 |
| Proba obxectiva | A24 A25 A26 B2 B3 B8 B9 | 3 | 0 | 3 |
| Prácticas de laboratorio | A24 A25 A26 B2 B3 B7 B9 B10 C1 | 21 | 21 | 42 |
| Sesión maxistral | A25 A26 B2 B3 B8 B9 | 21 | 42 | 63 |
| 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 |
| Solución de problemas | Desenvolveranse exemplos e exercicios nos que os estudantes terán que aplicar os coñecementos teóricos da materia a casos concretos. Garantirase a interactividade, resolvendo dúbidas por parte dos estudantes e animándoos a contrastar as súas solucións e expor cuestións relevantes. |
| Proba obxectiva | Levarase a cabo unha avaliación da materia mediante unha proba que incluírá tanto preguntas sobre os contidos teóricos, como supostos prácticos e exercicios de aplicación relacionados cos distintos temas vistos na materia. |



| | |
|--------------------------|---|
| Prácticas de laboratorio | Desenvolveranse exemplos e exercicios nos que os estudantes terán que aplicar os coñecementos teóricos da materia a casos concretos. Garantirase a interactividade, resolvendo dúbidas por parte dos estudantes e animándoos a contrastar as súas solucións e expor cuestións relevantes. |
| Sesión maxistral | Clases maxistrais onde se expoñerán os conceptos teóricos da materia, sen perder nunca de vista exemplos de aplicación para motivar e contextualizar os contidos da materia. Fomentarase a interactividade en clase mediante a formulación de preguntas e utilizaranse distintos recursos como transparencias ou demostracións. |

Atención personalizada

| Metodoloxías | Descrición |
|---|--|
| Solución de problemas Prácticas de laboratorio | <p>O desenvolvemento, tanto das clases maxistrais como das de resolución de problemas e os laboratorios de prácticas, realizarase atendendo ao progreso dos estudantes nas capacidades de comprensión e asimilación dos contidos impartidos. O avance xeral da clase compaxinarase cunha atención específica a aqueles estudantes que presenten maiores dificultades na tarefa da aprendizaxe e cun apoio adicional a aqueles que presenten maior desenvolvemento e desexen ampliar coñecementos.</p> <p>A atención personalizada se realiza a través dos seguintes canles:</p> <ul style="list-style-type: none">- Correo-e: De uso pra facer consultas de resposta curta ou solicitar encontros para resolver dúbidas.- Teams: De uso nos horarios oficiais de titorías nos que se garante unha resposta inmediata. Non entanto, dado o elevado número de estudantes, recoméndase solicitude previa de titoría via correo-e.- Campus virtual: Poñeránse a disposición dos/as estudantes "foros temáticos" que resolvan dúbidas xerais ligadas a actividades específicas como as prácticas. |

Avaliación

| Metodoloxías | Competencias | Descrición | Cualificación |
|--------------------------|-----------------------------------|---|---------------|
| Solución de problemas | A24 A25 A26 B2 B4 B7 B10 C1 C4 | Valoraranse os resultados, forma e condicións de realización de diversos traballos puntuables que se detallarán durante o curso. | 40 |
| Proba obxectiva | A24 A25 A26 B2 B3 B8 B9 | Realización obrigatoria. Avaliarase o dominio dos coñecementos teóricos e operativos da materia. | 60 |
| Prácticas de laboratorio | A24 A25 A26 B2 B3 B7 B9 B10 C1 | Son obrigatorias para poder aprobar os traballos da solución de problemas e inflúen na cualificación final destes, pero non se puntúan á marxe da nota outorgada á solución de problemas. | 0 |

Observacións avaliación



Para poder aprobar a materia o estudante deberá cumprir os seguintes requisitos (puntuación entre 0 e 10 en todas as actividades):

- Lograr unha nota superior ou igual a 3,5 na proba mixta final realizada ao final do cuadrimestre.
- Lograr unha nota superior ou igual a 5 ao realizar a suma de todas as probas de avaliación.

Notas sobre os traballos prácticos e solución de problemas:

- Todas as actividades terán unha única oportunidade para a súa entrega durante o curso académico, salvo a proba obxectiva final que terá dúas oportunidades oficiais de exame.

Avaliación no caso da convocatoria adiantada:

No caso de que o alumno solicite e asista á convocatoria adiantada, o 50% da súa nota será a proba mixta (exame final) e o outro 50% o traballo tutelado. O traballo tutelado deberá entregarse como data límite o mesmo día do exame oficial da convocatoria adiantada. Para aprobar a materia, o alumno debe cumprir os requisitos mencionados anteriormente.

Sobre a responsabilidade

compartida dos traballos en grupo.

- Nas

actividades que se levan a cabo en grupos, 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.

Matrícula a tempo parcial

-

Os alumnos matriculados a tempo parcial terán que entregar as actividades avaliáveis nas condicións e prazos específicos que se establecerán. Será obriga do estudante comunicar a súa situación ao profesorado.

Non presentado

- Quen non concurra á

proba obxectiva no período oficial de avaliación terá a condición de ?Non presentado? (NP). Na primeira oportunidade, isto será extensible a quen non entregue todas as prácticas en tempo e forma.

Fontes de información

Bibliografía básica

- Jules S. Damji, Brooke Wenig, Tathagata das, and Denny Lee (2020). Learning Spark : Lightning-fast big data analytics. Sebastopol, CA : O'Reilly
- Bill Chambers and Matei Zaharia (2018). Spark : the definitive guide :big data processing made simple . Sebastopol, CA : O'Reilly Media, Inc.

Bibliografía complementaria

Recomendacións

Materias que se recomenda ter cursado previamente

Procesamento Paralelo/614G02023
 Aprendizaxe Automática III/614G02026
 Aprendizaxe Automática I/614G02019
 Aprendizaxe Automática II/614G02021

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