



Guía Docente

| Datos Identificativos | | | | | 2023/24 |
|-----------------------|---|--------------------|---|----------|---------|
| Asignatura (*) | Comprensión da Linguaxe Natural | Código | 614544008 | | |
| Titulación | Máster Universitario en Intelixencia Artificial | | | | |
| Descritores | | | | | |
| Ciclo | Período | Curso | Tipo | Créditos | |
| Mestrado Oficial | 1º cuatrimestre | Primeiro | Obrigatoria | 6 | |
| Idioma | Inglés | | | | |
| Modalidade docente | Presencial | | | | |
| Prerrequisitos | | | | | |
| Departamento | Ciencias da Computación e Tecnoloxías da Información | | | | |
| Coordinación | Gómez Rodríguez, Carlos | Correo electrónico | carlos.gomez@udc.es | | |
| Profesorado | Gómez Rodríguez, Carlos Vilares Calvo, David Vilares Ferro, Jesus | Correo electrónico | carlos.gomez@udc.es david.vilares@udc.es jesus.vilares@udc.es | | |
| Web | campusvirtual.udc.es | | | | |
| Descrición xeral | <p>A materia introduce os conceptos e técnicas básicas asociadas ao procesamento da linguaxe natural, punto de partida para o deseño de contornas de explotación de información e de diálogo baseadas na linguaxe humana, tanto a nivel léxico como sintáctico, semántico e pragmático.</p> <p>O obxectivo é introducir ao estudante na complexidade inherente á análise da linguaxe natural humana, fundamentalmente asociada á ambigüidade e dependencias contextuais que presenta, e no deseño de estruturas de datos e algoritmos que permitan o seu tratamento práctico.</p> | | | | |

Competencias / Resultados do título

| Código | Competencias / Resultados do título |
|--------|---|
| A2 | CE01 - Comprensión e dominio de técnicas para o procesamento léxico, sintáctico e semántico de textos en linguaxe natural |
| A3 | CE02 - Comprensión e dominio dos fundamentos e técnicas de procesamento de documentos enlazados, estruturados e non estruturados, e da representación do seu contido |
| A4 | CE03 - Comprensión e coñecemento das técnicas de representación e procesado de coñecemento mediante ontoloxías, grafos e RDF, así como das ferramentas asociadas ás mesmas |
| B1 | CG01 - Manter e estender os plantexamentos teóricos fundados para permitir a introducción e explotación de tecnoloxías novas e avanzadas no eido da 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 |
| 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 su área de estudo |
| B10 | CB05 - Que os estudantes posúan as habilidades de aprendizaxe que lles permitan continuar estudando dun xeito que terá que ser en gran medida autodirixido ou autónomo |
| C2 | CT02 - Dominar a expresión e comprensión, de xeito oral e escrito, dun idioma extranxeiro |
| 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 |
| C7 | CT07 - Desenvolver a capacidade de traballar en equipos interdisciplinares ou transdisciplinares, para ofrecer propostas que contribúan a un desenrolo sostible ambiental, económico, político e social |
| C8 | CT08 - Valorar a importancia que ten a investigación, a innovación e o desenrolo tecnolóxico no avance socioeconómico e cultural da sociedade |



Resultados da aprendizaxe

| Resultados de aprendizaxe | Competencias / Resultados do título | | |
|--|-------------------------------------|---|--------------------------|
| Coñecer, comprender e analizar a representación formal de diversos fenómenos léxicos, sintácticos e semánticos da linguaxe natural. | AM1 | BM1 BM3 BM4 BM6 BM10 | CM2 CM8 |
| Coñecer, comprender e saber usar as tecnoloxías, marcos e librarías para a construción de sistemas de procesamento da linguaxe natural. | AM1 AM2 | BM3 BM4 BM6 BM7 BM10 | CM2 CM3 CM7 |
| Deseñar, implementar e saber usar algoritmos e estruturas de datos para tratar e dar soporte aos diversos fenómenos característicos da linguaxe natural. | AM1 AM2 AM3 | BM1 BM3 BM4 BM6 BM7 BM10 | CM2 CM3 CM7 CM8 |
| Coñecer, comprender e analizar as técnicas de procesamento da linguaxe natural para o procesado e desambiguación a nivel léxico, sintáctico e semántico. | AM1 AM2 AM3 | BM1 BM3 BM4 BM6 BM7 BM10 | CM2 CM3 CM7 CM8 |
| Coñecer e comprender os problemas que presenta a ambigüidade e imprecisión nas fontes de datos en linguaxe natural e técnicas para resolvelos. | AM1 AM2 | BM1 BM3 BM4 BM6 BM7 BM10 | CM2 CM3 CM7 CM8 |

Contidos

| Temas | Subtemas |
|---------------------|---|
| Introdución. | Niveis de análise. Ambigüidade e dependencias contextuais. |
| Análise léxica. | Segmentación. Dicionarios e tesauros. Técnicas de etiquetaxe morfosintáctica. |
| Análise sintáctica. | Gramáticas alxebraicas. Gramáticas suavemente sensibles ao contexto. Gramáticas de dependencias. Gramáticas probabilísticas. |
| Análise semántica. | Semántica léxica. Dependencias semánticas. Grafos semánticos. |

Planificación



| Metodoloxías / probas | Competencias / Resultados | Horas lectivas (presenciais e virtuais) | Horas traballo autónomo | Horas totais |
|--------------------------|---|---|-------------------------|--------------|
| Sesión maxistral | A2 A3 A4 B1 B3 B6 B7 B10 C2 C8 | 21 | 21 | 42 |
| Prácticas de laboratorio | A2 A3 A4 B3 B4 B6 B7 B10 C2 C3 C7 C8 | 14 | 48 | 62 |
| Solución de problemas | A2 A3 A4 B3 B4 B6 B7 B10 C2 | 7 | 25 | 32 |
| Proba obxectiva | A2 A3 A4 B1 B6 B7 C2 | 3 | 9 | 12 |
| 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 |
| Sesión maxistral | Clases teóricas, 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. |
| Prácticas de laboratorio | Clases prácticas con uso de computador, que permiten ao alumno familiarizarse desde un punto de vista práctico coas cuestións expostas nas clases teóricas. |
| Solución de problemas | Aprendizaxe baseada en problemas, seminarios, estudo de casos e proxectos. |
| Proba obxectiva | Avaliarase o dominio dos coñecementos teóricos e operativos da materia. |

| Atención personalizada | |
|--|--|
| Metodoloxías | Descrición |
| Sesión maxistral Prácticas de laboratorio Solución de problemas Proba obxectiva | O desenvolvemento, tanto das clases maxistras como das de resolución de problemas e os laboratorios de prácticas, realizarase atendendo ao progreso dos alumnos nas capacidades de comprensión e asimilación dos contidos impartidos. O avance xeral da clase compaxinarase cunha atención específica a aqueles alumnos que presenten maiores dificultades na tarefa da aprendizaxe e cun apoio adicional a aqueles que presenten maior desenvoltura e desexen ampliar coñecementos. No que respecta ás tutorías individuais, dado o seu carácter personalizado, non deben dedicarse a estender os contidos con novos conceptos, senón a aclarar os conceptos xa expostos. O profesor utilizaráas como unha interacción que lle permita extraer conclusións respecto ao grao de asimilación da materia por parte dos alumnos. |

| Avaliación | | | |
|--------------------------|---|---|---------------|
| Metodoloxías | Competencias / Resultados | Descrición | Cualificación |
| Prácticas de laboratorio | A2 A3 A4 B3 B4 B6 B7 B10 C2 C3 C7 C8 | As entregas das prácticas deben realizarse dentro do prazo establecido no campus virtual e deben seguir as especificacións indicadas no enunciado tanto para a súa presentación como o seu defensa. | 50 |
| Proba obxectiva | A2 A3 A4 B1 B6 B7 C2 | Realización obrigatoria. Avaliarase o dominio dos coñecementos teóricos e operativos da materia. | 50 |

| Observacións avaliación |
|-------------------------|
| |



Os alumnos deberán alcanzar polo menos un 40% da máxima nota cada parte (teoría, práctica) e en calquera caso a suma de ambas as partes debe superar un 5 para superar a materia. Se non se cumpre algún dos requisitos anteriores, a nota da convocatoria establecerase de acordo á menor nota obtida. En caso de non alcanzar o mínimo nalgunha das partes, o alumno terá unha segunda oportunidade na que soamente se lle esixirá a entrega de dita parte. As entregas das prácticas deben realizarse dentro do prazo establecido no campus virtual e deben seguir as especificacións indicadas no enunciado tanto para a súa presentación como a súa defensa. Terá a condición de "Presentado" quen entregue todas as prácticas obrigatorias ou concorra á proba obxectiva no período oficial de avaliación. No caso de realización fraudulenta de exercicios ou probas, aplicarase a Normativa de avaliación do rendemento académico dos estudantes e de revisión de cualificacións. En aplicación da normativa correspondente sobre plaxio, a copia total ou parcial dalgún exercicio de prácticas ou de teoría suporá o suspenso nas dúas oportunidades do curso, coa cualificación de 0,0 en ambos os casos.

Fontes de información

| | |
|------------------------------------|---|
| Bibliografía básica | <ul style="list-style-type: none"> - Manning, C., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT Press - Goldberg, Y. (2017). Neural network methods for natural language processing. Synthesis lectures on human language technologies. Morgan Claypool - Jacob Eisenstein (2019). Introduction to Natural Language Processing. MIT Press - Jurafsky, D. & Martin, J. H. (2022). Speech and Language Processing (3rd ed. draft). Disponible en: https://web.stanford.edu/~jurafsky/slp3/ - Manning, C., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT Press |
| Bibliografía complementaria | <ul style="list-style-type: none"> - Chollet, F. (2018). Keras: The python deep learning library. Astrophysics Source Code Library - Stuart Russell, Peter Norvig (2020). Artificial Intelligence: A Modern Approach, 4th Edition. Pearson - Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze (2008). Introduction to Information Retrieval. Cambridge University Press, Cambridge - Kübler, S., McDonald, R., & Nivre, J. (2009). Dependency Parsing. Synthesis lectures on human language technologies. Morgan Claypool <p>Adicionalmente, manexaranse textos científicos dispoñibles nas bibliotecas dixitais da área, como o ACL Anthology ou ACM.</p> |

Recomendacións

Materias que se recomenda ter cursado previamente

Materias que se recomenda cursar simultaneamente

Aprendizaxe Automática I/614544012

Materias que continúan o temario

Minería de Textos/614544011

Modelado da Linguaxe/614544009

Intelixencia Web e Tecnoloxías Semánticas/614544010

Observacións

Segundo se recolle nas distintas normativas de aplicación para a docencia universitaria, esta materia incorpora a perspectiva de xénero. Propiciarase a intervención en clase de alumnos e alumnas, e traballarase para identificar e modificar prexuízos e actitudes sexistas e influirase na contorna para modificalos e fomentar valores de respecto e igualdade. Deberanse detectar situacións de discriminación por razón de xénero e proponerse accións e medidas para corrixilas.

(*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