

Guía docente

270958 - MUD - Minería de Datos No Estructurados

Última modificación: 31/01/2024

Unidad responsable: Facultad de Informática de Barcelona
Unidad que imparte: 723 - CS - Departamento de Ciencias de la Computación.

Titulación: MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS (Plan 2021). (Asignatura obligatoria).

Curso: 2023 **Créditos ECTS:** 6.0 **Idiomas:** Inglés

PROFESORADO

Profesorado responsable: JORGE TURMO BORRÁS

Otros: Segon quadrimestre:
CARLOS ESCOLANO PEINADO - 12
SALVADOR MEDINA HERRERA - 11, 12
BARDIA RAFIEIAN - 11
JORGE TURMO BORRÁS - 11, 12

CAPACIDADES PREVIAS

Nivel alto de programación en python.
Conocimientos de estadística y matemáticas a nivel de grado universitari en ingeniería o similar.
Conceptos básicos de aprendizaje automático

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE12. Aplicar la ciencia de datos en proyectos multidisciplinares para resolver problemas en dominios nuevos o poco conocidos y que sean económicamente viables, socialmente aceptables, y de acuerdo con la legalidad vigente

CE13. Identificar las principales amenazas en el ámbito de la ética y la privacidad de datos en un proyecto de ciencia de datos (tanto en el aspecto de gestión como de análisis de datos) y desarrollar e implantar medidas adecuadas para mitigar dichas amenazas.

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

Genéricas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

Transversales:

CT1. ESPÍRITU EMPRENDEDOR E INNOVADOR: Conocer y entender la organización de una empresa y las ciencias que rigen su actividad; tener capacidad para entender las normas laborales y las relaciones entre la planificación, las estrategias industriales y comerciales, la calidad y el beneficio. Conocer y entender los mecanismos en que se basa la investigación científica, así como los mecanismos e instrumentos de transferencia de resultados entre los diferentes agentes socioeconómicos implicados en los procesos de I+D+i.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

Básicas:

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

METODOLOGÍAS DOCENTES

Clase expositiva participativa de contenidos teóricos y prácticos.

Clase práctica de resolución, con la participación de los estudiantes, de casos prácticos y/o ejercicios relacionados con los contenidos de la materia.

Prácticas de laboratorio - Trabajo en grupo

Tutorías.

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

1. Conocer y comprender las tareas básicas de procesamiento del lenguaje natural y su aplicación en tareas de análisis de texto.

2. Conocer, comprender y aplicar métodos de minería de textos, incluyendo reconocimiento de entidades, análisis de sentimiento y recuperación de documentos.

3. Conocer, comprender y aplicar los principios básicos del aprendizaje profundo en tareas sobre datos no estructurados, tales como el procesamiento del lenguaje o la visión por computador..

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas grupo grande	27,0	18.00
Horas grupo pequeño	27,0	18.00
Horas aprendizaje autónomo	96,0	64.00

Dedicación total: 150 h

CONTENIDOS

Procesamiento de lenguaje natural y su aplicación en análisis de textos.



Descripción de las etapas del procesamiento de lenguaje natural.

clasificación del texto y similitud de textos.

Extracción de información: Reconocimiento de entidades, extracción de relaciones

Técnicas de aprendizaje profundo para análisis de datos no estructurados.

Principales arquitecturas de aprendizaje profundo para datos no estructurados

ACTIVIDADES

Practica de laboratorio

Objetivos específicos:

3

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE13. Identificar las principales amenazas en el ámbito de la ética y la privacidad de datos en un proyecto de ciencia de datos (tanto en el aspecto de gestión como de análisis de datos) y desarrollar e implantar medidas adecuadas para mitigar dichas amenazas.

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 52h

Aprendizaje autónomo: 52h

Examen final

Objetivos específicos:

1, 2

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE12. Aplicar la ciencia de datos en proyectos multidisciplinares para resolver problemas en dominios nuevos o poco conocidos y que sean económicamente viables, socialmente aceptables, y de acuerdo con la legalidad vigente

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT1. ESPÍRITU EMPRENDEDOR E INNOVADOR: Conocer y entender la organización de una empresa y las ciencias que rigen su actividad; tener capacidad para entender las normas laborales y las relaciones entre la planificación, las estrategias industriales y comerciales, la calidad y el beneficio. Conocer y entender los mecanismos en que se basa la investigación científica, así como los mecanismos e instrumentos de transferencia de resultados entre los diferentes agentes socioeconómicos implicados en los procesos de I+D+i.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 53h

Actividades dirigidas: 3h

Aprendizaje autónomo: 50h



NLP y sus aplicaciones

Objetivos específicos:

1

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT1. ESPÍRITU EMPRENDEDOR E INNOVADOR: Conocer y entender la organización de una empresa y las ciencias que rigen su actividad; tener capacidad para entender las normas laborales y las relaciones entre la planificación, las estrategias industriales y comerciales, la calidad y el beneficio. Conocer y entender los mecanismos en que se basa la investigación científica, así como los mecanismos e instrumentos de transferencia de resultados entre los diferentes agentes socioeconómicos implicados en los procesos de I+D+i.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 2h

Grupo grande/Teoría: 2h



etapas del procesamiento del lenguaje natural

Objetivos específicos:

1

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT1. ESPÍRITU EMPRENDEDOR E INNOVADOR: Conocer y entender la organización de una empresa y las ciencias que rigen su actividad; tener capacidad para entender las normas laborales y las relaciones entre la planificación, las estrategias industriales y comerciales, la calidad y el beneficio. Conocer y entender los mecanismos en que se basa la investigación científica, así como los mecanismos e instrumentos de transferencia de resultados entre los diferentes agentes socioeconómicos implicados en los procesos de I+D+i.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 11h 30m

Grupo grande/Teoría: 8h 30m

Grupo mediano/Prácticas: 3h



clasificación de textos. Similitud de textos.

Objetivos específicos:

2

Competencias relacionadas:

CE12. Aplicar la ciencia de datos en proyectos multidisciplinares para resolver problemas en dominios nuevos o poco conocidos y que sean económicamente viables, socialmente aceptables, y de acuerdo con la legalidad vigente

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 2h

Grupo grande/Teoría: 1h 30m

Grupo mediano/Prácticas: 0h 30m

Extracción de información: Reconocimiento de entidades. Extracción de relaciones.

Objetivos específicos:

1, 2

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE12. Aplicar la ciencia de datos en proyectos multidisciplinares para resolver problemas en dominios nuevos o poco conocidos y que sean económicamente viables, socialmente aceptables, y de acuerdo con la legalidad vigente

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT1. ESPÍRITU EMPRENDEDOR E INNOVADOR: Conocer y entender la organización de una empresa y las ciencias que rigen su actividad; tener capacidad para entender las normas laborales y las relaciones entre la planificación, las estrategias industriales y comerciales, la calidad y el beneficio. Conocer y entender los mecanismos en que se basa la investigación científica, así como los mecanismos e instrumentos de transferencia de resultados entre los diferentes agentes socioeconómicos implicados en los procesos de I+D+i.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 2h

Grupo grande/Teoría: 1h 30m

Grupo mediano/Prácticas: 0h 30m



Técnicas de deep learning para el análisis de datos no estructurados.

Objetivos específicos:

3

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE13. Identificar las principales amenazas en el ámbito de la ética y la privacidad de datos en un proyecto de ciencia de datos (tanto en el aspecto de gestión como de análisis de datos) y desarrollar e implantar medidas adecuadas para mitigar dichas amenazas.

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 7h

Grupo grande/Teoría: 5h

Grupo mediano/Prácticas: 2h



Principales arquitecturas de deep learning para datos no estructurados

Objetivos específicos:

3

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE13. Identificar las principales amenazas en el ámbito de la ética y la privacidad de datos en un proyecto de ciencia de datos (tanto en el aspecto de gestión como de análisis de datos) y desarrollar e implantar medidas adecuadas para mitigar dichas amenazas.

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 5h 30m

Grupo grande/Teoría: 4h

Grupo mediano/Prácticas: 1h 30m



practica de laboratorio en reconocimiento de entidades

Objetivos específicos:

1, 2

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE12. Aplicar la ciencia de datos en proyectos multidisciplinares para resolver problemas en dominios nuevos o poco conocidos y que sean económicamente viables, socialmente aceptables, y de acuerdo con la legalidad vigente

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT1. ESPÍRITU EMPRENDEDOR E INNOVADOR: Conocer y entender la organización de una empresa y las ciencias que rigen su actividad; tener capacidad para entender las normas laborales y las relaciones entre la planificación, las estrategias industriales y comerciales, la calidad y el beneficio. Conocer y entender los mecanismos en que se basa la investigación científica, así como los mecanismos e instrumentos de transferencia de resultados entre los diferentes agentes socioeconómicos implicados en los procesos de I+D+i.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 5h

Grupo pequeño/Laboratorio: 5h

practica de laboratorio en extracción de relaciones

Objetivos específicos:

1, 2

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE12. Aplicar la ciencia de datos en proyectos multidisciplinares para resolver problemas en dominios nuevos o poco conocidos y que sean económicamente viables, socialmente aceptables, y de acuerdo con la legalidad vigente

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT1. ESPÍRITU EMPRENDEDOR E INNOVADOR: Conocer y entender la organización de una empresa y las ciencias que rigen su actividad; tener capacidad para entender las normas laborales y las relaciones entre la planificación, las estrategias industriales y comerciales, la calidad y el beneficio. Conocer y entender los mecanismos en que se basa la investigación científica, así como los mecanismos e instrumentos de transferencia de resultados entre los diferentes agentes socioeconómicos implicados en los procesos de I+D+i.

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 5h

Grupo pequeño/Laboratorio: 5h

practica de laboratorio sobre redes neuronales

Objetivos específicos:

3

Competencias relacionadas:

CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.

CE13. Identificar las principales amenazas en el ámbito de la ética y la privacidad de datos en un proyecto de ciencia de datos (tanto en el aspecto de gestión como de análisis de datos) y desarrollar e implantar medidas adecuadas para mitigar dichas amenazas.

CE11. Analizar y extraer conocimiento de información no estructurada mediante técnicas de procesamiento de lenguaje natural, minería de textos e imágenes

CE6. Diseñar el proceso de Ciencia de Datos y aplicar metodologías científicas para obtener conclusiones sobre poblaciones y tomar decisiones en consecuencia, a partir de datos estructurados o no estructurados y potencialmente almacenados en formatos heterogéneos.

CE7. Identificar las limitaciones impuestas por la calidad de datos en un problema de ciencia de datos y aplicar técnicas para disminuir su impacto

CT4. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.

CT5. TERCERA LENGUA: Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

CB8. Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CB9. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CB10. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CB6. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CB7. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

Dedicación: 5h

Grupo pequeño/Laboratorio: 5h

SISTEMA DE CALIFICACIÓN

práctica de laboratorio 50% + examen final 50%



BIBLIOGRAFÍA

Básica:

- Jurafsky, Dan; Martin, James H. *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition*. 2nd ed. Upper Saddle River: Prentice Hall, 2008. ISBN 9789332518414.
- Indurkha, Nitin ; Damerau, Fred J. *Handbook of natural language processing*. 2nd ed. Boca Raton: Taylor–Francis, 2010. ISBN 9780429149207.
- Manning, Christopher D; Schütze, Hinrich. *Foundations of statistical natural language processing*. Cambridge, Mass: MIT Press, 1999. ISBN 0262133601.
- Mitkov, Ruslan. *The Oxford handbook of computational linguistics*. Oxford: Oxford University Press, 2003. ISBN 0198238827.
- Indurkha, Nitin; Damerau, Frederick J. *Handbook of natural language processing*. Boca Raton: Chapman and Hall/CRC, 2010. ISBN 9781420085938.
- Clark, Alexander; Fox, Chris; Lappin, Shalom. *The Handbook of computational linguistics and natural language processing*. Oxford: Wiley-Blackwell, 2010. ISBN 9781444324044.
- Manning, C.; See, A. *Natural language processing with deep learning*. Stanford University, 2022.
- Collins, M. *Natural language processing*. Columbia University,