

Guía docente

270953 - MVA - Análisis Multivariante

Última modificación: 23/11/2023

Unidad responsable: Facultad de Informática de Barcelona
Unidad que imparte: 715 - EIO - Departamento de Estadística e Investigación Operativa.

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

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

PROFESORADO

Profesorado responsable: SERGI RAMIREZ MITJANS - CARINA GIBERT OLIVERAS

Otros: Primer quadrimestre:
DANTE CONTI - 11, 12
CARINA GIBERT OLIVERAS - 11, 12

CAPACIDADES PREVIAS

La asignatura supone haber efectuado previamente un curso básico de estadística, programación y matemáticas, en particular tener adquiridos los conceptos:

- Concepto de media, matriz de covarianza y correlaciones.
- Concepto de prueba de hipótesis.
- Operaciones de álgebra matricial, valores y vectores propios.
- Programación de algoritmos.
- Regresión lineal múltiple.

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

- CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa
- 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.
- CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos
- 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
- CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.
- CE9. Aplicar métodos adecuados para el análisis de otro tipo de formatos, tales como procesos y grafos, dentro del ámbito de ciencia de datos

Genéricas:

- CG2. Identificar y aplicar métodos de análisis, extracción de conocimiento y visualización de datos recogidos en formatos muy diversos.
- CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

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

El curso tiene como objetivo proporcionar las bases estadísticas para la minería de datos. El aprendizaje se realiza mediante una combinación de explicación teórica y su aplicación a un caso real. Las clases desarrollarán los conocimientos científicos necesarios, mientras que las clases de laboratorio serán su aplicación para resolver problemas de minería de datos. La implementación de prácticas favorece habilidades genéricas relacionadas con el trabajo en equipo y la presentación de resultados y sirve para integrar diferentes conocimientos de la materia. El software utilizado será principalmente R.

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

- 1.Descripción multivariante de los datos
- 2.Visualización de los datos
- 3.Inferencia multivariada
- 4.Clasificación de nuevos individuos

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas grupo grande	24,0	16.00
Horas actividades dirigidas	6,0	4.00
Horas grupo pequeño	24,0	16.00
Horas aprendizaje autónomo	96,0	64.00

Dedicación total: 150 h



CONTENIDOS

Introducción al Análisis de Datos Multivariado

Descripción:

Ventajas del tratamiento multivariante. Ejemplos de datos multivariados. Métodos probables y libres de distribución. Enfoque exploratorio versus modelado.

Análisis de Componentes Principales

Descripción:

Análisis de individuos. Análisis de variables. Representación visual de la información. Reducción de dimensionalidad. Información suplementaria. Descomposición en valores singulares.

Descomposición en valores singulares

Descripción:

Método para explorar y visualizar filas y columnas de una tabla mediante descomposición en valores singulares.

Análisis Factorial

Descripción:

Método de reducción de dimensiones.

Escalamiento Multidimensional

Descripción:

Este método trata con datos relativos a distancias entre los elementos. Normalmente, a partir de datos procedentes de distancias o similitudes. El método pone de manifiesto una estructura común de todos los elementos y la especificidad de cada uno de ellos, evidenciando que hace que sean cerca o distantes.

Clustering Jerárquico y de Particiones

Descripción:

Dos enfoques para agrupar métodos utilizados para clasificar las observaciones, dentro de un conjunto de datos, en múltiples grupos en función de su semejanza.

Métodos de profiling automático

Descripción:

Los métodos de profiling ayudan a entender las características comunes de los clusters

Distribución normal multivariante

Descripción:

Particularidades de la distribución normal en el caso general de planteamientos multivariantes, donde los puntos se distribuyen en varias dimensiones.



Análisis Discriminante

Descripción:

El análisis discriminante (DA) y Naïve Bayes (NB) son métodos de clasificación. DA clasifica las observaciones en grupos no superpuestos, basándose en las puntuaciones de una o más variables predictoras cuantitativas. NB es un algoritmo de aprendizaje sencillo que utiliza la regla de Bayes junto con una suposición firme que los atributos son independientes condicionalmente, dada la clase.

Árboles de clasificación y regresión

Descripción:

Este método puede predecir o clasificar. Explica cómo se pueden predecir o clasificar los valores de una variable de resultado en función de otros valores. Tiene una estructura gráfica muy útil.

Reglas de asociación

Descripción:

Encontrar patrones, asociaciones, correlaciones o estructuras causales frecuentes entre conjuntos de elementos u objetos en bases de datos de transacciones, bases de datos relacionales y otros repositorios de información.



ACTIVIDADES

Presentación de la asignatura + Análisis Multivariante de datos

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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.

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.

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

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.

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.

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.

Dedicación: 7h

Grupo grande/Teoría: 2h

Aprendizaje autónomo: 5h



Análisis de Componentes Principales

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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.

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.

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

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.

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h

Descomposición en valores singulares

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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.

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.

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

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.

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Métodos de profiling automático

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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.

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.

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

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.

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Análisis Factorial

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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.

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.

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

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.

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Análisis Factorial

Objetivos específicos:

2, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

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.

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

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Escalamiento Multidimensional

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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.

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.

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

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.

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Análisis Discriminante

Objetivos específicos:

3, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

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

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

CE9. Aplicar métodos adecuados para el análisis de otro tipo de formatos, tales como procesos y grafos, dentro del ámbito de ciencia de datos

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.

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

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h

Árboles de Clasificación y Regresión

Objetivos específicos:

2, 3, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

CE9. Aplicar métodos adecuados para el análisis de otro tipo de formatos, tales como procesos y grafos, dentro del ámbito de ciencia de datos

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.

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

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.

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Clustering Jerárquico y de Particiones

Objetivos específicos:

2, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

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.

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

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Distribución normal multivariante

Objetivos específicos:

2, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

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.

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

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.

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Reglas de asociación

Objetivos específicos:

4

Competencias relacionadas:

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

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.

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

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.

Dedicación: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h

Práctica Final

Dedicación: 14h 54m

Actividades dirigidas: 1h 54m

Aprendizaje autónomo: 13h

Examen de conceptos

Dedicación: 13h 06m

Aprendizaje autónomo: 13h 06m



Resumen y ejercicios prácticos. 1a parte

Objetivos específicos:

1, 2, 3, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

CE9. Aplicar métodos adecuados para el análisis de otro tipo de formatos, tales como procesos y grafos, dentro del ámbito de ciencia de datos

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.

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

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.

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.

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.

Dedicación: 7h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Resumen y ejercicios prácticos. 2a parte

Objetivos específicos:

1, 2, 3, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

CE9. Aplicar métodos adecuados para el análisis de otro tipo de formatos, tales como procesos y grafos, dentro del ámbito de ciencia de datos

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.

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

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.

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.

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.

Dedicación: 7h

Grupo pequeño/Laboratorio: 2h

Aprendizaje autónomo: 5h



Dudas de la pràctica

Objetivos específicos:

1, 2, 3, 4

Competencias relacionadas:

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

CG3. Definir, diseñar e implementar sistemas complejos que cubran todas las fases en proyectos de ciencia de datos

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

CE5. Modelar, diseñar e implementar sistemas complejos de datos, incluyendo la visualización de datos

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

CE8. Extraer información de datos estructurados y no estructurados, teniendo en cuenta la naturaleza multivariante de los mismos.

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

CE10. Identificar los métodos de aprendizaje automático y modelización estadística a utilizar para resolver un problema específico de ciencia de datos y aplicarlos de forma rigurosa

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.

CE9. Aplicar métodos adecuados para el análisis de otro tipo de formatos, tales como procesos y grafos, dentro del ámbito de ciencia de datos

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.

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

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.

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.

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.

Dedicación: 2h

Grupo grande/Teoría: 2h



SISTEMA DE CALIFICACIÓN

La evaluación del curso se basará en las notas obtenidas del trabajo práctico realizados durante el curso, una nota de teoría y la nota obtenida en la práctica final.

Cada práctica dará lugar a la redacción de la correspondiente redacción del informe y podrá realizarse de forma conjunta, hasta un máximo de cuatro alumnos por grupo .

Los ejercicios realizados a lo largo del curso tienen como objetivo consolidar el aprendizaje de técnicas multivariantes.

La práctica final es que los estudiantes demuestren su madurez para resolver un problema real utilizando técnicas de visualización multivariante, interpretación de agrupamiento y predicción. Los estudiantes elegirán entre diferentes alternativas para resolver el problema. Esta práctica será presentada y defendida públicamente, en la que el alumno deberá responder a cualquier duda sobre los modelos teóricos y métodos utilizados en la solución. Las prácticas se realizan utilizando el software R.

Las pruebas escritas evaluarán la asimilación de los conceptos básicos de la asignatura. Habrá tres pruebas a lo largo del curso, en horas de clase de teoría. Mientras que la presentación de la práctica final se realizará durante el período de exámenes.

Los ejercicios realizados durante el curso tienen una ponderación del 20%, la nota de teoría del 40% y la práctica final del 40%.

BIBLIOGRAFÍA

Básica:

- Larose, Daniel T. Discovering knowledge in data : an introduction to data mining [en línea]. 2nd ed. Hoboken, N.J: John Wiley & Sons,, 2014 [Consulta: 15/01/2024]. Disponible a : <https://onlinelibrary-wiley-com.recursos.biblioteca.upc.edu/doi/book/10.1002/9781118874059>. ISBN 9781118874059.

RECURSOS

Enlace web:

- <https://cran.r-project.org/>- <https://rstudio.com/>