19-20

MÁSTER UNIVERSITARIO EN
INGENIERÍA Y CIENCIA DE DATOS

GUÍA DE ESTUDIO PÚBLICA

MODELOS BAYESIANOS JERÁRQUICOS

CÓDIGO 31110094

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC424007847FCC148

19-20

MODELOS BAYESIANOS JERÁRQUICOS

CÓDIGO 31110094

ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

EQUIPO DOCENTE

HORARIO DE ATENCIÓN AL ESTUDIANTE

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

RESULTADOS DE APRENDIZAJE

CONTENIDOS

METODOLOGÍA

SISTEMA DE EVALUACIÓN

BIBLIOGRAFÍA BÁSICA

BIBLIOGRAFÍA COMPLEMENTARIA

RECURSOS DE APOYO Y WEBGRAFÍA

ADENDA AL SISTEMA DE EVALUACIÓN CON MOTIVO DE LA PANDEMIA COVID 19

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/

4F3D161454DDF26FC42007847FCC148

Nombre de la asignatura	MODELOS BAYESIANOS JERÁRQUICOS
Código	31110094
Curso académico	2019/2020
Título en que se imparte	MÁSTER UNIVERSITARIO EN INGENIERÍA Y CIENCIA DE DATOS
Tipo	CONTENIDOS
Nº ETCS	4
Horas	100.0
Periodo	SEMESTRE 2
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

Bienvenidos a la guía de la asignatura de Modelos Bayesianos Jerárquicos.
La asignatura es una extensión natural de Modelado Estadístico de Datos (obligatoria, primer cuatrimestre) que añade a todas las técnicas estudiadas allí los modelos jerárquicos multinivel. Estos modelos (introducidos en el tema 2) permiten, entre otras cosas, inferir las probabilidades a priori a partir de los datos mediante hiperpriors y tratar de manera diferente a grupos distintos de observaciones o parámetros. Dadas las complejidades inherentes a los modelos multinivel y, en particular, la dificultad de obtener soluciones analíticas en espacios de alta dimensionalidad, se le presta una especial atención a las técnicas aproximadas de simulación de probabilidades a posteriori (tema 4). Finalmente, se dedica el último tema a la implementación práctica de este tipo de modelos, a su evaluación y crítica (introducida en el tema 3) y a su utilización en inferencia en el entorno proporcionado por al menos una librería. La librería elegida puede estar sujeta a evolución a medida que aparezcan en la bibliografía nuevas técnicas y entornos de prototipado más comprensivos.

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Esta asignatura tiene un contenido teórico exigente en materias de estadística multivariante e inferencia Bayesiana. En realidad, la formación previa necesaria se imparte en la asignatura de Modelado Estadístico de Datos. En cualquier caso, es recomendable que el estudiante se asegure de que posee las destrezas asociadas al razonamiento probabilístico y que conoce las propiedades de las diferentes distribuciones de probabilidad, maneja con soltura el Álgebra matricial y conoce los fundamentos del Cálculo Infinitesimal.

EQUIPO DOCENTE

Nombre y Apellidos	LUIS MANUEL SARRO BARO (Coordinador de asignatura)
Correo Electrónico	lsb@dia.uned.es
Teléfono	91398-8715
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	INTELIGENCIA ARTIFICIAL

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC42007847FCC148

Nombre y Apellidos	JOSE LUIS AZNARTE MELLADO
Correo Electrónico	jlaznarte@dia.uned.es
Teléfono	91398-9688
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	INTELIGENCIA ARTIFICIAL

HORARIO DE ATENCIÓN AL ESTUDIANTE

El equipo docente atenderá preferentemente a los/las estudiantes de manera virtual a través de los foros del curso virtual. De esta manera, las respuestas a un/una estudiante particular serán de utilidad al resto de compañeros. Por supuesto, también atenderá dudas a través del correo electrónico lsb@dia.uned.es siempre que sea requerido por un/una estudiante.
El horario de atención personal en el despacho será los lunes de 10 a 14 horas, pero recomendamos a los estudiantes que notifiquen con antelación su visita porque puede ocurrir que los profesores tengan compromisos académicos que de manera ocasional ocurran dentro de este horario y cuya agenda escape a su control. Dentro de este horario también se puede optar por la llamada telefónica (Extensión 8715).
La atención durante el horario de mencionado en el párrafo anterior está garantizada (salvo imponderables u compromisos académicos). Ello no significa que ése horario sea exclusivo: lxs estudiantes podrán contactar con el equipo docente en cualquier momento por las vías antes citadas (foros, correo electrónico y teléfono).
Los estudiantes pueden proponer videoconferencias en grupo para aclarar aspectos particulares del temario, pero dichas videoconferencias deberán prepararse con anterioridad a través de los foros. La interacción habitual consiste en que unx o varixs estudiantes proponen una videoconferencia sobre un tema particular, entre todos definimos el problema y las cuestiones sobre las que queremos debatir o que pretendemos aclarar y finalmente se fija un dia y una hora que sea lo más compatible con las restricciones temporales de equipo docente y estudiantes.

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

COMPETENCIAS BÁSICAS

CB6 - 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.
CB7 - 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 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.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC42400784 7FCC148

CB9 - 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.
CB10 - 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.

COMPETENCIAS GENERALES

CG1 - Identificar los métodos apropiados para la solución de problemas asociados a la ciencia de datos y la analítica de información
CG2 - Ser capaz de aplicar diferentes técnicas de aprendizaje máquina, seleccionando el algoritmo óptimo que genere modelos precisos y permita el desarrollo de soluciones predictivas en diferentes ámbitos de uso

COMPETENCIAS TRANSVERSALES

CT1 - Ser capaz de abordar y desarrollar proyectos innovadores en entornos científicos, tecnológicos y multidisciplinares.
CT2 - Ser capaz de tomar decisiones y formular juicios basados en criterios objetivos (datos experimentales, científicos o de simulación disponibles).

COMPETENCIAS ESPECÍFICAS

CE1 - Conocer los fundamentos de la inferencia estadística y el análisis probabilístico y desarrollar diferentes tipos de modelos probabilísticos
CE6 - Diseñar mecanismos de evaluación de modelos de aprendizaje y comprender las métricas usadas para dicha evaluación

RESULTADOS DE APRENDIZAJE

Los resultados más relevantes que se pretenden alcanzar con el estudio de esta asignatura son los siguientes:

1. Comprender la importancia y utilidad de un modelado multi-nivel para datos con dependencias probabilísticas complejas y diferentes grupos de parámetros u observaciones.
1. Discriminar los diferentes escenarios en los que el modelado jerárquico ofrece ventajas respecto a los modelos planos.
1. Construir descripciones en forma de grafo de las relaciones de dependencia condicional entre observaciones, parámetros e hiperparámetros de un modelo.
1. Traducir dichos grafos a formulaciones matemáticas analíticas en términos de verosimilitudes y distribuciones a priori.
2. Seleccionar las técnicas de muestreo más apropiadas para un modelo dado.
3. Evaluar la bondad de un modelo dado (incluyendo su capacidad de predicción y las propiedades de convergencia de su muestreo de la probabilidad a posteriori en caso de ser necesario) y ser capaz de comparar y seleccionar modelos.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC42007847FCC148

CONTENIDOS

Introducción a la Inferencia Bayesiana

En este tema se repasarán los conceptos básicos de Inferencia Bayesiana vistos con anterioridad en la asignatura de Modelado Estadístico de Datos y se profundizará en algunos de ellos con ejemplos tomados del primer tema del texto base.

•Probabilidad e Inferencia
•Notación general de la inferencia estadística
•Inferencia Bayesiana
•La probabilidad como una medida de la incertidumbre
•Resultados útiles de Teoría de la Probabilidad

Modelos Jerárquicos

Ésta será nuestra primera toma de contacto con los modelos jerárquicos Bayesianos. En este tema veremos cómo gracias a los modelos jerárquicos podremos, entre otras cosas, disminuir la dependencia de los resultados de nuestra inferencia de la elección de los priors haciendo a estos últimos sujetos de la propia inferencia.

•La construcción de una distribución a priori paramétrica
•La intercambiabilidad y los modelos jerárquicos
•El análisis Bayesiano de los modelos jerárquicos conjugados
•El ejemplo de las 8 escuelas.
•Modelado jerárquico aplicado al meta-análisis
•Priors débilmente informativos para las matrices de covarianza

Evaluación y comparación de modelos

En este tema abordaremos la cuestión fundamental de cómo comprobar que la inferencia que hemos realizado es confiable y que el modelo sobre el que hemos realizado la inferencia describe correctamente los datos. Por otra parte, estudiaremos como comparar modelos alternativos para un mismo conjunto de observaciones.

•Medidas de la exactitud de las predicciones
•Criterios basados en la Teoría de la Información y validación cruzada
•Comparación de modelos basada en el rendimiento predictivo
•Comparación de modelos mediante el factor de Bayes.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC42007847FCC148

Aspectos computacionales de la Inferencia Bayesiana

Los modelos multiparamétricos y especialmente los jerárquicos presentan desafíos y dificultades que habitualmente obligan a la estimación aproximada (no analítica) de las distribuciones a posteriori de los parámetros involucrados. Para ello se han desarrollado en los últimos años técnicas muy sofisticadas que quedan fuera del alcance de esta asignatura, pero cuyos fundamentos se hayan en una serie de técnicas más sencillas que explorremos aquí.

•Principios básicos de las cadenas de Markov
•El muestreador de Gibbs
•Los algoritmos de Metropolis y Metropolis-Hastings
•Los muestreadores de Gibbs y Metropolis como bloques básicos de la inferencia
•Cómo evaluar la convergencia
•El número efectivo de muestras
•Hamiltonian Monte Carlo
•Inferencia Variacional

Es importante enfatizar que el texto base contiene material más avanzado que el que constituye este tema y que, por muy interesante que sea, debemos dejar fuera por razones de carga de trabajo.

Librerías de modelado probabilístico

En este último tema se describirá Edward como entorno de especificación de modelos para inferencia probabilística.

METODOLOGÍA

El equipo docente de la asignatura ha desarrollado sus contenidos teniendo en mente un estudiante promedio de esta Universidad que compatibiliza sus estudios con otro tipo de actividades de tipo laboral o profesional y que tiene restricciones variables en cuanto a su disponibilidad de tiempo. Por ello, asume que el/la estudiante seguirá los contenidos teóricos a través del texto base y realizará una serie de ejercicios prácticos propuestos por dicho equipo docente ligados a los contenidos de cada tema. Tanto para el estudio de los contenidos como para la realización de los ejercicios, el estudiante contará con el apoyo del profesorado a través de la plataforma virtual o mediante otro tipo de canales de comunicación (correo electrónico, teléfono o videoconferencias).
Finalmente, el estudiante realizará una práctica cuya memoria debe demostrar que ha adquirido los conocimientos y las destrezas descritas en los contenidos de la asignatura.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC42007847FCC148

Dicha práctica incluirá el desarrollo de software en el entorno de programación elegido para la asignatura.

Las actividades formativas de la asignatura son las que se listan a continuación:

•Estudios de contenidos (55 horas)

•Actividades en la plataforma virtual (5 horas)

•Prácticas Informáticas (40 horas)

•Total: 100 horas.

SISTEMA DE EVALUACIÓN

TIPO DE PRUEBA PRESENCIAL

Tipo de examen	Examen tipo test
Preguntas test	
Duración del examen	120 (minutos)
Material permitido en el examen	
Todo tipo de material	
Criterios de evaluación	
Corrección en las respuestas	
% del examen sobre la nota final	70
Nota del examen para aprobar sin PEC	
Nota máxima que aporta el examen a la calificación final sin PEC	
Nota mínima en el examen para sumar la PEC	0
Comentarios y observaciones	

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad	Si
Descripción	
Prueba presencial descrita en el apartado anterior.	
Criterios de evaluación	
Ponderación de la prueba presencial y/o los trabajos en la nota final	70%
Fecha aproximada de entrega	
Comentarios y observaciones	

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC?	Si,PEC no presencial
Descripción	
Criterios de evaluación	
Ponderación de la PEC en la nota final	30%
Fecha aproximada de entrega	Hacia el final del cuatrimestre.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC424007847FCC148

Comentarios y observaciones

La PEC consistirá en una implementación de un modelo de datos y en la inferencia de sus parámetros mediante la librería Edward.

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? No

Descripción

Criterios de evaluación

Ponderación en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

¿CÓMO SE OBTIENE LA NOTA FINAL?

Tanto la prueba presencial (PP) como la PEC serán evaluadas de 0 a 10. En caso de que ambas evaluaciones se encuentren por encima de 4, la nota final (NF) será la media ponderada de ambas de acuerdo a la siguiente fórmula:
NF = 0,70 * PP + 0,30 * PEC.

BIBLIOGRAFÍA BÁSICA

ISBN(13):
Título:BAYESIAN DATA ANALYSIS (Third Edition)
Autor/es:Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. Y Rubin, D. B. ;
Editorial:Chapman & Hall/CRC Texts in Statistical Science.

Para el seguimiento de la asignatura mediante el texto base, conviene seguir estas indicaciones:

•El tema 1 del contenido de esta asignatura se corresponde con el capítulo 1 del texto base.
•El tema 2 se corresponde con el capítulo 5 del texto base.
•El tema 3 se corresponde con el capítulo 7 del texto base.
•El tema 4 se corresponde con el capítulo 11 del texto base más las secciones 12.4, 12.5 y 13.7.

BIBLIOGRAFÍA COMPLEMENTARIA

ISBN(13):9780123814852
Título:DOING BAYESIAN DATA ANALYSIS: A TUTORIAL WITH R AND BUGS
Autor/es:John K. Kruschke ;
Editorial:ACADEMIC PRESS

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC42007847FCC148

RECURSOS DE APOYO Y WEBGRAFÍA

Los/as estudiantes dispondrán de los siguientes recursos de apoyo al estudio:

- ***Guía de la asignatura**. Incluye el plan de trabajo y orientaciones para su desarrollo. Esta guía será accesible desde el curso virtual.*
- ***Curso virtual**. A través de esta plataforma los/as estudiantes tienen la posibilidad de consultar información de la asignatura, realizar consultas al Equipo Docente a través de los foros correspondientes, consultar e intercambiar información con el resto de los compañeros/as.*
- ***Biblioteca**. El estudiante tendrá acceso tanto a las bibliotecas de los Centros Asociados como a la biblioteca de la Sede Central, en ellas podrá encontrar un entorno adecuado para el estudio, así como de distinta bibliografía que podrá serle de utilidad durante el proceso de aprendizaje. Además, desde la biblioteca digital de la UNED, el estudiante tendrá acceso a Safari Books Online, una biblioteca digital con más de 30.000 libros técnicos en constante actualización.*

Los recursos de apoyo, especialmente aquellos asociados al rápidamente cambiante mundo de los entornos de desarrollo y los lenguajes de programación, se harán disponibles a través del curso virtual para poder actualizarlos de manera ágil y frecuente.

ADENDA AL SISTEMA DE EVALUACIÓN CON MOTIVO DE LA PANDEMIA COVID 19

https://app.uned.es/evacaldos/asignatura/adendasig/31110094

IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4F3D161454DDF26FC42007847FCC148