

Guía docente

Identificación de la asignatura

Asignatura / Grupo	11631 - Aprendizaje Estadístico y Toma de Decisiones / 1
Titulación	Máster Universitario en Análisis de Datos Masivos en Economía y Empresa
Créditos	6
Período de impartición	Primer semestre
Idioma de impartición	Castellano

Profesores

Horario de atención a los alumnos

Profesor/a	Hora de inicio	Hora de fin	Día	Fecha inicial	Fecha final	Despacho / Edificio
Jan Olof William Nilsson - william.nilsson@uib.es	12:15	13:15	Viernes	01/09/2018	31/07/2019	DB257
Heiko Jürgen Rachinger - heiko.rachinger@uib.es	15:00	16:00	Miércoles	01/10/2018	28/06/2019	DB 210/ Jovellanos (cita previa por email)

Contextualización

La asignatura de Aprendizaje Estadístico y Toma de Decisiones es una asignatura de carácter obligatorio. La asignatura tiene cuatro partes. La primera parte consta de un tema dedicado a los métodos de remuestreo, en concreto, validación y selección de modelos. La segunda parte consta de dos temas dedicados al aprendizaje no-supervisado (*unsupervised learning*), donde se discuten técnicas para reducción de la dimensionalidad y métodos de agrupación. La tercera parte consiste en aprendizaje supervisado (*supervised learning*), destinado a métodos de clasificación y regresión. La última parte trata efectos de tratamiento heterogéneos.

El curso es especialmente útil para introducir al alumno en algunas de herramientas más utilizadas para el tratamiento de datos masivos, tales como las técnicas de reducción de la dimensionalidad, métodos de agrupación y de clasificación y regresión. La cuarta parte es muy importante en aplicaciones cuando la toma de decisión implica implementar una política, una campaña o un tratamiento y el objetivo es encontrar quien debe recibir el tratamiento.

En el contexto de los Métodos Económicos Cuantitativos, el curso de Aprendizaje Estadístico y Toma de Decisiones se presenta como la continuación de la formación estadística y econométrica básica. Concretamente, la asignatura introduce al alumno a las herramientas econométricas necesarias para el análisis de los datos masivos.

Requisitos

Ningún requisito obligatorio.

Guía docente

Recomendables

Se recomienda que para seguir correctamente la asignatura Aprendizaje Estadístico y Toma de Decisiones se disponga de un conocimiento intermedio de probabilidad y técnicas estadísticas.

Competencias

Específicas

- * CE1. Aplicar las técnicas econométricas más relevantes para el tratamiento de datos masivos .
- * CE2. Contribuir a la buena gestión de la asignación de recursos (tanto en el ámbito privado como en el público) .
- * CE3. Aportar racionalidad al análisis y a la descripción de cualquier aspecto de la realidad económica .

Genéricas

- * CG3. Aplicar al análisis de los problemas criterios profesionales basados en el manejo de instrumentos .
- * CG5. Analizar los problemas con razonamiento crítico, sin prejuicios, con precisión y rigor .
- * CG7. Capacidad de síntesis .

Básicas

- * Se pueden consultar las competencias básicas que el estudiante tiene que haber adquirido al finalizar el máster en la siguiente dirección: http://estudis.uib.cat/es/master/comp_basiques/

Contenidos

Parte 1. Validación y Selección de Modelos (1 tema)

En esta primera parte el alumno aprenderá las técnicas básicas de remuestreo. En concreto se analizan los métodos de validación cruzada.

Parte 2: Aprendizaje No-Supervisado (*Unsupervised Learning*) (2 temas)

En la segunda parte se introduce al alumno en los aspectos relacionados con la reducción de la dimensionalidad y métodos de agrupación (*clustering*).

Parte 3: Aprendizaje Supervisado (*Supervised Learning*) (2 temas)

En la tercera parte se introduce al alumno en los aspectos relacionados con la regresión logística, análisis discriminante, métodos basados en árboles; *bagging*, bosques aleatorios, *boosting* y clasificación de múltiples clases.

Parte 4: Modelos con efectos de tratamiento heterogéneos (1 tema)

La cuarta parte incluye modelos de efectos de tratamiento heterogéneos; *causal forest*, *causal boosting* y *causal MARS*.

Contenidos temáticos

1. Validación y Selección de Modelos

- * 1.1 Sesgo, Número de Parámetros y Varianza.

Guía docente

- * 1.2 Validación Cruzada (Cross-Validation) de Modelos.
- 2. Aprendizaje No-Supervisado (Unsupervised Learning)
 - 2.1 Técnicas para la Reducción de la Dimensionalidad:
 - * 2.1.1 Componentes Principales.
 - * 2.1.2 Otras Técnicas.
 - 2.2 Métodos de Agrupación (*Clustering*)
 - * 2.2.1 K-Medias *Clustering*.
 - * 2.2.2 Agrupación Jerárquica (*Hierarchical Clustering*)
- 3. Aprendizaje Supervisado (Supervised Learning)
 - 3.1 Métodos de Clasificación
 - * 3.1.1 Regresión Logística.
 - * 3.1.2 Análisis Discriminante
 - 3.2 Métodos Basados en Árboles (*Tree-Based Methods*)
 - * 3.2.1 Árboles de Regresión y Clasificación.
 - * 3.2.2 *Bagging*.
 - * 3.2.3 Bosques Aleatorios (*Random Forests*).
 - * 3.2.4 *Boosting*.
 - * 3.2.5 Clasificación de múltiples clases.
- 4. Modelos con efectos de tratamiento heterogéneos
 - 4.1 *Causal forest*
 - 4.2 *Causal boosting*
 - 4.3 *Causal MARS*

Metodología docente

Volumen

Es recomendable leer el material correspondiente antes de asistir a las clases magistrales, para facilitar el aprendizaje del contenido. También es importante revisar el temario después de cada clase para asegurarse de que todas las dudas se han solucionado. Al finalizar un tema teórico el estudiante realizará ejercicios y prácticas para asimilar y aplicar la teoría expuesta en clase. Una parte de este trabajo se hace en clase, pero es importante que cada estudiante realice también estos estudios fuera de clase.

En el volumen de trabajo se contempla la posibilidad de cesión de horas de docencia a actividades comunes del máster: seminarios, charlas, etc. que sean propuestos por la dirección del máster.

Actividades de trabajo presencial (2,4 créditos, 60 horas)



Guía docente

Modalidad	Nombre	Tip. agr.	Descripción	Horas
Clases teóricas	Clases teóricas	Grupo grande (G)	Las lecciones magistrales proporcionan una exposición detallada de lo más importante de cada tema, incluyendo conceptos nuevos, ejemplos y ejercicios ilustrativos. Las clases teóricas desarrollan los fundamentos estadísticos y econométricos de los modelos y herramientas analizadas. Una función importante de estas lecciones es la de facilitar que los estudiantes consideren las técnicas y modelos en el contexto del análisis económico aplicado. Por ello, además de fundamentar estadísticamente los modelos mediante su estudio teórico, las clases magistrales inciden en cómo evaluar y analizar los resultados econométricos, derivando a partir de los mismos las principales conclusiones económicas.	30
Clases prácticas	Clases prácticas	Grupo grande (G)	Para complementar la exposición de los temas teóricos, el estudiante realizará ejercicios y prácticas para asimilar y aplicar la teoría analizada en clase. Las sesiones prácticas también incluyen el uso del programa econométrico específico llamado R. Especialmente al finalizar cada uno de los temas, el alumno realizará aplicaciones prácticas con datos que ilustren el empleo de las técnicas.	28.5
Evaluación	Examen Parcial	Grupo grande (G)	Se realizará solamente un examen parcial durante el curso. Se valorará el conocimiento de los conceptos, técnicas y modelos expuestos en los tres primeros temas. La prueba será de carácter objetivo, valorándose tanto los conocimientos de tipo teórico, como la capacidad de interpretar y extraer resultados del análisis empírico.	1.5
Evaluación	Trabajo Final	Grupo grande (G)	Habrà un trabajo final en la convocatoria oficial. Esta evaluación permitirá valorar el conocimiento tanto teórico como aplicado de la econometría para datos masivos aplicada a datos que abarcan el programa (temas 1 a 4). Se valorará de manera objetiva tanto el nivel de conocimiento teórico, como la capacidad del alumno para interpretar y extraer conclusiones de los resultados obtenidos con los modelos y técnicas utilizadas.	0
Evaluación	Participación y ejercicios en clases	Grupo grande (G)	Puntalmente durante el curso el estudiante realizará ejercicios y prácticas para asimilar y aplicar la teoría desarrollada en clase. Se valorará la asistencia y la participación en las clases. También se valorará la realización de ejercicios en clase que serán entregados al profesor (con o sin preaviso).	0

Al inicio del semestre estará a disposición de los estudiantes el cronograma de la asignatura a través de la plataforma UIBdigital. Este cronograma incluirá al menos las fechas en las que se realizarán las pruebas de evaluación continua y las fechas de entrega de los trabajos. Asimismo, el profesor o la profesora informará a los estudiantes si el plan de trabajo de la asignatura se realizará a través del cronograma o mediante otra vía, incluida la plataforma Aula Digital.

Actividades de trabajo no presencial (3,6 créditos, 90 horas)

Guía docente

Modalidad	Nombre	Descripción	Horas
Estudio y trabajo autónomo individual	Preparación de unidades didácticas	Es recomendable leer el material correspondiente antes que asistir las clases magistrales para facilitar el aprendizaje del contenido. También es importante revisar el temario después de cada clase para asegurarse que todas las dudas se han solucionado. Estudiar la literatura y los recursos ofrecidos por los profesores es importante para profundizar en el aprendizaje y ver el contexto de cada apartado del temario.	90

Riesgos específicos y medidas de protección

Las actividades de aprendizaje de esta asignatura no conllevan riesgos específicos para la seguridad y salud de los alumnos y, por tanto, no es necesario adoptar medidas de protección especiales.

Evaluación del aprendizaje del estudiante

La evaluación del aprendizaje consiste en un trabajo final y una evaluación continua. El estudiante tendrá una calificación numérica entre 0 y 10 para cada una de las actividades evaluadas. La calificación global se calcula teniendo en cuenta los pesos asignados a las diferentes formas de evaluación.

El trabajo final es un trabajo escrito que se entrega en enero. A lo largo del semestre se realiza una sesión en la que cada estudiante, de forma individual, soluciona ejercicios con entrega al final de la sesión. La fecha de la sesión serán concretada al principio del semestre. Los alumnos que no asistan a la prueba parcial realizada durante el curso tendrán una calificación igual a 0 en dicha prueba. En el caso excepcional y debidamente documentado de que el alumno no pueda asistir a la prueba parcial por citación judicial a su persona, muerte de un familiar de primer grado u hospitalización del propio alumno, el porcentaje de la nota final previsto para dicha prueba parcial será trasladado al trabajo final.

A lo largo del semestre se pedirán a los estudiantes la resolución de ejercicios como parte de la nota de la evaluación continua. En las actividades evaluables correspondientes a competencias de aprendizaje y conocimientos para comprender ámbitos de aplicación (CE1-15), se considerará la asistencia y participación en actividades propuestas por la dirección del máster.

Alguien que no consiga aprobar el curso en enero tiene la posibilidad de recuperar el trabajo final. Hay que tener en cuenta que únicamente el trabajo final es recuperable, y su peso en la nota final es del 50%. Si el alumno ha obtenido una nota final del curso en el periodo de evaluación ordinario igual o superior a 5, no existirá la posibilidad de mejorar su evaluación en el periodo extraordinario.

Fraude en elementos de evaluación

De acuerdo con el artículo 33 del Reglamento académico, "con independencia del procedimiento disciplinario que se pueda seguir contra el estudiante infractor, la realización demostradamente fraudulenta de alguno de los elementos de evaluación incluidos en guías docentes de las asignaturas comportará, a criterio del profesor, una minusvaloración en su calificación que puede suponer la calificación de «suspense 0» en la evaluación anual de la asignatura".

Guía docente

Examen Parcial

Modalidad	Evaluación
Técnica	Pruebas objetivas (no recuperable)
Descripción	Se realizará solamente un examen parcial durante el curso. Se valorará el conocimiento de los conceptos, técnicas y modelos expuestos en los tres primeros temas. La prueba será de carácter objetivo, valorándose tanto los conocimientos de tipo teórico, como la capacidad de interpretar y extraer resultados del análisis empírico.
Criterios de evaluación	Resolución de ejercicios sobre los conceptos, técnicas y modelos analizados en la asignatura.
Porcentaje de la calificación final:	25%

Trabajo Final

Modalidad	Evaluación
Técnica	Trabajos y proyectos (recuperable)
Descripción	Habrà un trabajo final en la convocatoria oficial. Esta evaluación permitirá valorar el conocimiento tanto teórico como aplicado de la econometría para datos masivos aplicada a datos que abarcan el programa (temas 1 a 4). Se valorará de manera objetiva tanto el nivel de conocimiento teórico, como la capacidad del alumno para interpretar y extraer conclusiones de los resultados obtenidos con los modelos y técnicas utilizadas.
Criterios de evaluación	Trabajo aplicado a datos que abarcan el programa (temas 1 a 4).
Porcentaje de la calificación final:	60%

Participación y ejercicios en clases

Modalidad	Evaluación
Técnica	Otros procedimientos (no recuperable)
Descripción	Puntalmente durante el curso el estudiante realizará ejercicios y prácticas para asimilar y aplicar la teoría desarrollada en clase. Se valorará la asistencia y la participación en las clases. También se valorará la realización de ejercicios en clase que serán entregados al profesor (con o sin preaviso).
Criterios de evaluación	Se valorará la asistencia y la participación en las clases. También se valorará la realización de ejercicios en clase que serán entregados al profesor (con o sin preaviso).
Porcentaje de la calificación final:	15%

Recursos, bibliografía y documentación complementaria

Además de los libros especificados en la bibliografía, durante todo el curso se pondrá a disposición de los alumnos material específico. Se podrá descargar material adicional de la página Web de la asignatura, donde se encontrarán colgados las prácticas, los ejercicios complementarios, explicaciones más detalladas de algunos temas concretos y también las notas de los alumnos en los trabajos que se vayan realizando a lo largo del curso.

Bibliografía básica

Hastie, T., Tibshirani, R., & Friedman, J. (2009): The elements of statistical learning (2nd edition). Springer, Berlin: Springer series in statistics.
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 6). New York: Springer.



Guía docente

Bibliografía complementaria

Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. New York: Springer.

