

Titulación	Tipo	Curso
2503852 Estadística Aplicada	OB	3

Contacto

Nombre: Rosario Delgado De la Torre

Correo electrónico: rosario.delgado@uab.cat

Equipo docente

Rosario Delgado De la Torre

Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

Prerrequisitos

Se supone que el alumno que cursa esta asignatura ha adquirido las competencias propias de las asignaturas de

- Càlcul 1,
- Eines informàtiques per a l'Estadística i Introducció a la Programació,
- Introducció a la Probabilitat i Inferència Estadística 1, i
- Aprenentatge Automàtic 1.

Será necesario un buen nivel y práctica en la programación con R.

Objetivos y contextualización

Aprender qué son y cómo se utilizan las Redes Bayesianas: son un modelo probabilístico que se utiliza en el Aprendizaje Automático Supervisado y que describe las relaciones de tipo probabilístico entre las variables que afectan a determinado fenómeno de interés (que puede ser un sistema complejo) y se pueden utilizar como clasificadores.

Entender cómo las Redes Bayesianas se utilizan para evaluar y cuantificar riesgos, entre otras aplicaciones. Conocer diferentes metodologías que se tendrán que aplicar, o no, cuando se trabaje con estos modelos, dependiendo de sus características, en la fase de pre-proceso de la base de datos o bien en la fase de construcción del modelo predictivo.

Conocer diferentes métricas de comportamiento para realizar la validación del modelo y entender su utilidad y adecuación, en función de las características de la base de datos.

Aprender cómo construir scripts de R que permitan aprender estos modelos a partir de una base de datos y hacer su validación, utilizando las librerías pertinentes. Aplicarlo a datos reales.

Resultados de aprendizaje

1. CM09 (Competencia) Valorar la adecuación de los modelos con la utilización e interpretación correcta de indicadores y gráficos.
2. CM09 (Competencia) Valorar la adecuación de los modelos con la utilización e interpretación correcta de indicadores y gráficos.
3. CM10 (Competencia) Modificar el software existente si el modelo estadístico propuesto lo requiere, o crear nuevo software, si fuera necesario.
4. KM12 (Conocimiento) Proporcionar las hipótesis experimentales de la modelización, teniendo en cuenta las implicaciones técnicas y éticas relacionadas.
5. KM12 (Conocimiento) Proporcionar las hipótesis experimentales de la modelización, teniendo en cuenta las implicaciones técnicas y éticas relacionadas.
6. SM12 (Habilidad) Interpretar los resultados obtenidos para formular conclusiones respecto a las hipótesis experimentales
7. SM13 (Habilidad) Comparar el grado de ajuste entre diversos modelos estadísticos.
8. SM14 (Habilidad) Emplear gráficos de visualización del ajuste y de la adecuación del modelo.

Contenido

1. Introducción a las Redes Bayesianas (RBs).
Definición.
Inferencia con RBs.
Aprendizaje de RBs (tanto de estructura como de parámetros).
2. Las RBs como clasificadores.
La labor de clasificación dentro del Aprendizaje Automático Supervisado.
El criterio MAP. Tipo de RB (*Naive Bayes*, *Augmented Naive*, *TAN*).
Tipo de clasificación: binaria, multi-clase, multi-etiqueta.
3. La validación y las métricas de comportamiento.
Validación cruzada.
Métricas para el caso binario y multi-clase.
Métricas por el caso de clasificación ordinal.
4. Otros aspectos.
Clasificación multi-dimensional.
Ensembles de clasificadores.
El "*concept drive*" y las RBs dinámicas.
RBs gaussianas e híbridas.
Clasificación *multi-instance*.

Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Problemas	14	0,56	

Prácticas (entregas, controles)	12	0,48
Teoría	26	1,04
Tipo: Supervisadas		
Tutorías	10	0,4
Tipo: Autónomas		
Estudio y pensar problemas	40	1,6
Trabajo práctico con ordenadores	30	1,2

La asignatura se estructura a partir de clases teóricas, de problemas y prácticas. El seguimiento de la asignatura es presencial, pero será necesario ampliar las explicaciones del profesor con el estudio autónomo del alumno, con el apoyo de la bibliografía de referencia y del material facilitado por el profesor.

La clase de problemas se dedicará a la resolución orientada de algunos problemas propuestos. En las clases prácticas se trabajará con R y sus librerías. Se valorará especialmente la participación de los estudiantes en las clases de problemas y prácticas.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Evaluación

Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Examen	60%	3	0,12	CM09, SM12, SM13
PAC1	20%	6	0,24	CM09, CM10, KM12, SM12, SM13, SM14
PAC2	20%	9	0,36	CM09, SM13

La calificación final de esta asignatura se obtiene como la media ponderada de las notas de:

- PAC1 (20%)
- PAC2 (20%)
- Examen (60%)

Las pruebas de evaluación continuada PAC1 y PAC2 consisten en una entrega de problemas/ejercicios de prácticas/trabajos con R, que se concretarán a lo largo del curso, y en su desarrollo en las clases presenciales a lo largo del semestre.

Sólo se tendrán en cuenta en el cálculo de la media ponderada aquéllas notas que sean como mínimo 3.5 sobre 10 (las que no lo cumplan ponderarán 0).

Para aprobar la asignatura es necesario que esta media sea de al menos 5.0 sobre 10.

En caso de no aprobar la asignatura en primera convocatoria, el alumno puede presentarse a recuperación. El examen de recuperación representa el 100% de la nota final para aquellos alumnos que se presenten a la recuperación, que sólo pueden ser alumnos que no han superado la asignatura en primera convocatoria (el examen de recuperación no sirve para mejorar nota a alumnos que ya han aprobado).

Se considerará evaluable el alumno que haya presentado las entregas PAC1 o PAC2, o se haya presentado al examen o al examen de recuperación. En caso contrario constará en el acta como No Evaluable.

Para la eventual asignación de Matrículas de Honor no se tendrá en cuenta las notas de la segunda convocatoria.

Bibliografía

- Norman Fenton and Martin Neil, "Risk Assessment and Decision Analysis with Bayesian Networks", CRC Press. A Chapman & Hall Book, 2013. (Disponible en línea)
- Radhakrishnan Nagarajan, Marco Scutari and Sophie Lèbre, "Bayesian Networks in R with applications in Systems Biology", Springer, 2013. (Disponible en línea)
- Oliver Pourret, Patrick Naïm and Bruce Marcot, "Bayesian Networks. A practical guide to Applications", Series: Statistics in Practice. Wiley, 2008. (Disponible en línea)
- Richard E. Neapolitan, "Learning Bayesian Networks", Prentice Hall Series in Artificial Intelligence, 2004.
- Adnan Darwiche, "Modeling and reasoning with Bayesian networks", Cambridge, 2009.
- Kevin B. Korb and Ann E. Nicholson, "Bayesian Artificial Intelligence" (2nd edition), Series: Computer Science and Data Analysis. CRC Press. A Chapman & Hall book, 2011. (Disponible en línea)
- Daphne Koller and Nir Friedman, "Probabilistic Graphical Models", The MIT Press Cambridge, Massachusetts London, England, 2009. <http://mcb111.org/w06/KollerFriedman.pdf>
- Radhakrishnan Nagarajan, Marco Scutari and Sophie Lèbre, "Bayesian Networks in R with applications in systems biology", Series: Use R! Springer, 2013. (Disponible en línea)
- Marco Scutari and Jean-Baptiste Denis, "Bayesian networks with examples in R", Series: Texts in Statistical Science. CRC Press. A Chapman & Hall Book, 2015.

Software

Se utilizará el software R con algunas librerías que se indicarán oportunamente a lo largo del curso. Preferiblemente en el entorno RStudio.

Lista de idiomas

Nombre	Grupo	Idioma	Semestre	Turno
(PLAB) Prácticas de laboratorio	1	Catalán	segundo cuatrimestre	tarde
(TE) Teoría	1	Catalán	segundo cuatrimestre	tarde