

**ANX-PR/CL/001-02**  
**GUÍA DE APRENDIZAJE**

**ASIGNATURA**

Analisis inteligente de datos

**CURSO ACADÉMICO - SEMESTRE**

2015-16 - Primer semestre

## Datos Descriptivos

---

<b>Nombre de la Asignatura</b>	Análisis inteligente de datos
<b>Titulación</b>	10AN - Master Universitario en Ingeniería Informática
<b>Centro responsable de la titulación</b>	E.T.S. de Ingenieros Informaticos
<b>Semestre/s de impartición</b>	Primer semestre
<b>Carácter</b>	Obligatoria
<b>Código UPM</b>	103000607
<b>Nombre en inglés</b>	Intelligent Data Analysis

## Datos Generales

---

<b>Créditos</b>	4.5	<b>Curso</b>	1
<b>Curso Académico</b>	2015-16	<b>Período de impartición</b>	Septiembre-Enero
<b>Idioma de impartición</b>	Castellano	<b>Otros idiomas de impartición</b>	

## Requisitos Previos Obligatorios

---

### Asignaturas Superadas

El plan de estudios Master Universitario en Ingeniería Informática no tiene definidas asignaturas previas superadas para esta asignatura.

### Otros Requisitos

El plan de estudios Master Universitario en Ingeniería Informática no tiene definidos otros requisitos para esta asignatura.

## Conocimientos Previos

---

### Asignaturas Previas Recomendadas

El coordinador de la asignatura no ha definido asignaturas previas recomendadas.

### Otros Conocimientos Previos Recomendados

Basic knowledge of statistics will be helpful.

## Competencias

---

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.

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

CG19 - Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería en Informática

## Resultados de Aprendizaje

---

RA42 - Conocer y aplicar las principales técnicas de análisis de datos multivariantes.

RA44 - Formular, analizar y validar modelos de regresión, análisis discriminante, clasificación y "clustering".

RA45 - Aplicar la metodología apropiada para el ajuste de series temporales.

RA46 - Ser capaz de estructurar problemas de toma de decisiones bajo el paradigma bayesiano.

RA43 - Conocer y aplicar técnicas de reducción de la dimensionalidad en un conjunto de datos multivariantes.

RA47 - Ser capaz de modelizar problemas reales en los que la incertidumbre sea un componente esencial, mediante redes bayesianas.

## Profesorado

---

### Profesorado

Nombre	Despacho	e-mail	Tutorías
Moreno Diaz, Arminda ( <b>Coordinador/a</b> )	2112	arminda.moreno@upm.es	
Rodriguez Galiano, Maria Isabel	2112	mariaisabel.rodriguez@upm.es	
Gonzalez Pachon, Jacinto	2105	jacinto.gonzalez.pachon@upm.es	
Mateos Caballero, Alfonso	2110	alfonso.mateos@upm.es	
Fdez Del Pozo De Salamanca, Juan Antonio	2105	juan.fdezpozo.salamanca@upm.es	
Jimenez Martin, Antonio	2110	antonio.jimenez@upm.es	
Garcia Castro, Raul	2110	r.garcia@upm.es	

**Nota.-** Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

## Descripción de la Asignatura

---

The course is intended to be a non-exhaustive survey of techniques to convert multivariate data into useful information so that good decisions can be made. The perspective is twofold, theoretical and applied, covering topics such as: exploratory data analysis, statistical summaries and graphical representations, dimensionality reduction, regression techniques, time series analysis, decision theory and probabilistic graphical models. There will be an emphasis on hands-on application of the theory and methods throughout, with extensive use of R.

## Temario

---

1. Descriptive statistics and statistical modelling.
  - 1.1. Aspects of multivariate data. Descriptive statistics. Introduction to R.
  - 1.2. Getting Data from the Semantic Web with R.
  - 1.3. Dimensionality reduction: Principal Component Analysis and biplots.
  - 1.4. Regression models.
  - 1.5. Discrimination analysis and clustering.
2. Time Series.
  - 2.1. Definitions, Applications and Techniques.
  - 2.2. Stationarity and Seasonality.
  - 2.3. Common approaches.
  - 2.4. Box-Jenkins model identification, estimation and validation.
  - 2.5. Forecasting.
3. Introduction to Decision Analysis.
  - 3.1. Structure and representation of a decision problem.
  - 3.2. Decision making under certainty and uncertainty.
  - 3.3. Preferences and beliefs modelling.
  - 3.4. Collective decision making.
4. Graphical Models for Decision Making.
  - 4.1. Decision Trees and Influence Diagrams for optimal decisions.
  - 4.2. Bayesian networks for diagnosis and prognosis.
  - 4.3. Sensitivity Analysis for explanation of reasoning.

## Cronograma

**Horas totales:** 117 horas

**Horas presenciales:** 48 horas (41%)

**Peso total de actividades de evaluación continua:**  
100%

**Peso total de actividades de evaluación sólo prueba final:**  
100%

Semana	Actividad Presencial en Aula	Actividad Presencial en Laboratorio	Otra Actividad Presencial	Actividades Evaluación
Semana 1	<b>Module 1: Lecture</b> Duración: 03:00 LM: Actividad del tipo Lección Magistral			<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 2	<b>Module 1: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Computer Lab. Module 1.</b> Duración: 02:00 OT: Otras actividades formativas		<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 3	<b>Module 1: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Computer Lab. Module 1.</b> Duración: 02:00 OT: Otras actividades formativas		<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 4	<b>Module 1: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Computer Lab. Module 1.</b> Duración: 02:00 OT: Otras actividades formativas		<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Work on Assignment 1</b> Duración: 04:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 5	<b>Module 1: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Computer Lab. Module 1.</b> Duración: 02:00 OT: Otras actividades formativas		<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Work on Assignment 1</b> Duración: 04:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 6	<b>Module 1: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Computer Lab. Module 1.</b> Duración: 02:00 OT: Otras actividades formativas		<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Work on Assignment 1</b> Duración: 04:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial

Semana 7	<p><b>Module 1: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>	<p><b>Computer Lab. Module 1.</b> Duración: 02:00 OT: Otras actividades formativas</p>		<p><b>Work on Assignment 1</b> Duración: 04:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 8	<p><b>Module 2: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>	<p><b>Computer Lab. Module 2.</b> Duración: 02:00 OT: Otras actividades formativas</p>		<p><b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Work on Assignment 2</b> Duración: 03:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Upload Assignment 1</b> Duración: 00:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad no presencial</p>
Semana 9	<p><b>Module 2: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>	<p><b>Computer Lab. Module 2.</b> Duración: 02:00 OT: Otras actividades formativas</p>		<p><b>Work on Assignment 2</b> Duración: 04:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 10	<p><b>Module 2: Lecture</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral</p>	<p><b>Computer Lab. Module 2.</b> Duración: 02:00 OT: Otras actividades formativas</p>		<p><b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 11	<p><b>Module 3: Lecture</b> Duración: 03:00 LM: Actividad del tipo Lección Magistral</p>			<p><b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>
Semana 12	<p><b>Module 3: Lecture</b> Duración: 03:00 LM: Actividad del tipo Lección Magistral</p>			<p><b>Upload Assignment 2</b> Duración: 00:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad no presencial <b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial</p>

Semana 13	<b>Module 4: Lecture.</b> Duración: 03:00 LM: Actividad del tipo Lección Magistral			<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Work on Assignment 3</b> Duración: 04:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 14		<b>Computer Lab. Module 4.</b> Duración: 03:00 OT: Otras actividades formativas		<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Work on Assignment 3</b> Duración: 04:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 15		<b>Computer Lab. Module 4.</b> Duración: 03:00 OT: Otras actividades formativas		<b>Individual Study</b> Duración: 02:00 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial <b>Work on Assignment 3</b> Duración: 04:30 OT: Otras técnicas evaluativas Evaluación continua Actividad no presencial
Semana 16				<b>Upload/Presentation Assignment 3</b> Duración: 03:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial
Semana 17				<b>Final Exam</b> Duración: 03:00 PG: Técnica del tipo Presentación en Grupo Evaluación sólo prueba final Actividad presencial

**Nota.-** El cronograma sigue una planificación teórica de la asignatura que puede sufrir modificaciones durante el curso.

**Nota 2.-** Para poder calcular correctamente la dedicación de un alumno, la duración de las actividades que se repiten en el tiempo (por ejemplo, subgrupos de prácticas") únicamente se indican la primera vez que se definen.

## Actividades de Evaluación

Semana	Descripción	Duración	Tipo evaluación	Técnica evaluativa	Presencial	Peso	Nota mínima	Competencias evaluadas
1	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
2	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
3	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
4	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
4	Work on Assignment 1	04:30	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10, CG19
5	Work on Assignment 1	04:30	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10, CG19
5	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
6	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
6	Work on Assignment 1	04:30	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10, CG19
7	Work on Assignment 1	04:00	Evaluación continua	OT: Otras técnicas evaluativas	No			
7	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
8	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
8	Work on Assignment 2	03:30	Evaluación continua	OT: Otras técnicas evaluativas	No			CB7
8	Upload Assignment 1	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	34%	3 / 10	4, CB10, CB7, CG19
9	Work on Assignment 2	04:30	Evaluación continua	OT: Otras técnicas evaluativas	No			
9	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
10	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
11	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
12	Upload Assignment 2	00:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	No	33%	3 / 10	4
12	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
13	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
13	Work on Assignment 3	04:30	Evaluación continua	OT: Otras técnicas evaluativas	No			
14	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
14	Work on Assignment 3	04:30	Evaluación continua	OT: Otras técnicas evaluativas	No			
15	Individual Study	02:00	Evaluación continua	OT: Otras técnicas evaluativas	No			CB10
15	Work on Assignment 3	04:30	Evaluación continua	OT: Otras técnicas evaluativas	No			
16	Upload/Presentation Assignment 3	03:00	Evaluación continua	TG: Técnica del tipo Trabajo en Grupo	Sí	33%	3 / 10	4

Semana	Descripción	Duración	Tipo evaluación	Técnica evaluativa	Presencial	Peso	Nota mínima	Competencias evaluadas
17	Final Exam	03:00	Evaluación sólo prueba final	PG: Técnica del tipo Presentación en Grupo	Sí	100%	3 / 10	CB7, CG19, 4, CB10

## Criterios de Evaluación

---

1. The final grade is a weighed average of three projects assessment of approximately equal weight. These reports will consist in the analysis of three datasets provided by the instructors. The analysis will require the application of the methods and techniques studied in class. Each project must obtain a grade equal or greater than 3 (0-10 scale) for the final grade to be computed. This final grade must be equal or greater than 5 to pass the course. Eventually, students may be asked to present orally the conclusions of their work.
2. If any (or many) project grade doesn't meet the requirements exposed above, the project can be improved and presented to be graded again in the Final Exam. Again, grade requirements apply.
3. If a project (or projects) was not delivery when requested, it can be presented to be graded in the Final Exam. Again, grade requirements apply.

## Recursos Didácticos

Descripción	Tipo	Observaciones
Moodle	Recursos web	Important Communications Repository: slides, scripts, data sets, other resources.
Computer Lab	Equipamiento	Computer Room for hands-on sessions.
Johnson, R.A., Whichern, D.W. (2007) Applied Multivariate Statistical Analysis. Pearson Education.	Bibliografía	Friendly exposition of the most important multivariate techniques, including clustering. They also introduce other Artificial Intelligence techniques like neural networks.
Rencher, A.C. Methods of Multivariate Analysis.	Bibliografía	Clear exposition of Multivariate Analysis Techniques, from a statistical point of view. Many examples.
Everitt, B.S. and Dunn G. (1997) Applied Multivariate Data Analysis. Arnold.	Bibliografía	Excellent exposition of multivariate techniques. They make the Generalised Linear Model easily understandable.
Hair, J.F., Black, W.C., Babin, B.J., Anderson R.E. Multivariate Data Analysis.	Bibliografía	A Global Perspective on multivariate Techniques. Very detailed examples. In almost every topic, a "Rules of Thumb" section summarizes the relevant facts.
Sharma, S (1996). Applied Multivariate Techniques. Wiley.	Bibliografía	
Multivariate Analysis of Ecological Data. Greenacre, Primicerio. Fundación BBVA.	Recursos web	<a href="http://www.fbbva.es/TLFU/tlfu/esp/publicaciones/libros/fichalibro/index.jsp?codigo=769">http://www.fbbva.es/TLFU/tlfu/esp/publicaciones/libros/fichalibro/index.jsp?codigo=769</a> Excellent recopilation and description of multivariate techniques applied to Ecological Data. Clear, educational, didactic explanations. Lots of examples
Biplots in Practice. Michael Greenacre. Fundación BBVA.	Recursos web	<a href="http://www.multivariatestatistics.org/biplots.html">http://www.multivariatestatistics.org/biplots.html</a> Excellent Monograph on Biplots.
Rawlings, J.O., Pantula, S.G., Dickey, D.A. Applied Regression Analysis.	Bibliografía	Almost everything about Regression Models.
Chatfield, C. (2003) The Analysis of Time Series: An Introduction. Chapman and Hall.	Bibliografía	
French, S. Decision Theory. Ellis Horwood Ltd.	Bibliografía	
Koller, D., Friedman, N (2009) Probabilistic Graphical Models. Principles and techniques. MIT Press.	Bibliografía	
Duda, R. Hart P.E., Stork D.G. Pattern Classification (2001). Wiley.	Bibliografía	