



# UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

Facultad de Economía y Planificación

Departamento de Estadística e Informática

Semestre 2014 - II

## SILABO

### I. DATOS GENERALES

Curso:	TECNICAS MULTIVARIADAS AVANZADAS
Código:	EP - 4069
Créditos:	3-2-4
Prerrequisito:	Técnicas Multivariadas
Profesor:	Ms. Carlos López de Castilla Vásquez

### II. SUMILLA

Introducción al Statistical Learning. Regresión lineal. Clasificación. Métodos de remuestreo. Selección de modelos y regularización. Más allá de la linealidad. Métodos basados en árboles. Support vector machines. Aprendizaje no supervisado. Redes Bayesianas.

### III. OBJETIVO

Proporcionar los conceptos y fundamentos de la aplicación de los métodos multivariados de aprendizaje supervisado y no supervisado con la finalidad que el estudiante tenga la capacidad de aplicarlos e interpretarlos.

### IV. CONTENIDO ANALÍTICO

#### CAPÍTULO I INTRODUCCIÓN

##### Semana 1:

- Introducción. Statistical Learning. Problemas a resolver. Aprendizaje supervisado y no supervisado. Statistical Learning vs Machine Learning.

#### CAPÍTULO II REGRESIÓN LINEAL

##### Semanas 2:

- Introducción. Regresión lineal. Regresión múltiple. Interpretando los coeficientes de regresión. Estimación y predicción. Predictores cualitativos. Efectos no lineales. Interacción.

#### CAPÍTULO III CLASIFICACIÓN

##### Semanas 3:

- Introducción. Regresión logística vs Regresión lineal. Regresión logística con más de dos clases. Análisis Discriminante. Curva ROC. Naive Bayes. Regresión logística vs Análisis Discriminante.

#### CAPÍTULO IV MÉTODOS DE REMUESTREO

##### Semanas 4 y 5:

- Introducción. Validación Cruzada y Bootstrap. Proceso de validación. Validación Cruzada K – fold. Validación Cruzada para clasificación. Bootstrap vs Pruebas de permutación.

#### CAPÍTULO V SELECCIÓN DE MODELOS Y REGULARIZACIÓN

##### Semanas 6 y 7:

- Introducción. Subset selection. Selección Forward y backward. Cp, AIC, BIC y R2 ajustado. Validación cruzada. Regresión Ridge. Lasso. Regresión por Componentes Principales. Mínimos cuadrados parciales.

##### Semana 8: EXAMEN PARCIAL



# UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

Facultad de Economía y Planificación

Departamento de Estadística e Informática

Semestre 2014 - II

## CAPÍTULO VI MÁS ALLA DE LA LINEALIDAD

### Semana 9:

- Regresión Polinomial. Funciones paso. Polinomios por tramos. Splines lineales y cúbicos. Splines cúbicos naturales. Modelos aditivos generalizados. GAM para clasificación.

## CAPÍTULO VII MÉTODOS BASADOS EN ÁRBOLES

### Semanas 10 y 11:

- Introducción. Métodos de clasificación basados en árboles. Árboles de clasificación por Bagging. Random Forest. Clasificación por Boosting.

## CAPÍTULO VIII SUPPORT VECTOR MACHINES

### Semanas 12:

- Introducción. Clasificador por Support Vector Machines. No linealidad y Kernels. Support Vector Machines vs Regresión logística.

## CAPÍTULO IX APRENDIZAJE NO SUPERVISADO

### Semana 13:

- Introducción. Objetivos del aprendizaje no supervisado. Componentes principales. Clúster K-Means. Clúster Jerárquico.

## CAPÍTULO X REDES BAYESIANAS

### Semana 14:

- Introducción. Redes Bayesianas. Algoritmos de construcción de la estructura de la red. Aplicaciones.

### Semana 15: PRESENTACIÓN Y EXPOSICIÓN DE TRABAJOS

### Semana 16: EXAMEN FINAL

## V. METODOLOGÍA

El curso se desarrollará mediante exposiciones en clase, prácticas dirigidas, calificadas, tareas encargadas y un trabajo final que debe ser expuesto.

## VI. EQUIPOS Y MATERIALES

Computadora personal con el software estadístico RStudio. Multimedia en aula de clase.

## VII. EVALUACIÓN

La nota final es el promedio ponderado de lo siguiente:

Examen parcial	25%
Examen final	25%
Tareas encargadas	10%
Trabajo final	10%
Promedio prácticas calificadas (4)	30%

## VIII. FUENTES DE INFORMACIÓN

### Bibliografía Base

1. Gareth James, Daniela Witten, Trevor Hastie and Rob Tibshirani (2013) **An Introduction to Statistical Learning, with Applications in R**, Springer.

### Bibliografía complementaria

1. Britos, Paola & Otros (2005). **Minería de Datos**, Editorial Nueva Librería España.
2. Hernandez Orallo, Jose & Otros (2004). **Introducción a la Minería de Datos**. Editorial Pearson 2004. España.