



## Syllabus:

### Algoritmos no supervisados de deep learning

David Gálvez Ruiz

[davidgalvezruiz@gmail.com](mailto:davidgalvezruiz@gmail.com)

Curso académico 2019 – 2020

## 1. Presentación de la materia

Este módulo pretende introducir los algoritmos no supervisados más importantes. Estos algoritmos se aplican esencialmente a técnicas de clusterización para afrontar problemas de categorización que pueden ser una finalidad del análisis o bien un paso intermedio ante futuros análisis. La complejidad de estos algoritmos varía según la necesidad a la que se enfrenta el científico de datos, pues el uso e interpretación de ellos requiere un conocimiento sobre el procedimiento que realmente realizan y el problema al que se aplican.

## 2. Objetivos de aprendizaje

1. Conocimiento, identificación y selección de los métodos algorítmicos de clasificación no supervisados.
2. Clasificación y elección de métodos en función del problema a tratar.
3. Interpretación y análisis de resultados y redacción de informes. Manejo de software R.

## 3. Programa de la materia: estructura y contenido

1. **Bloque 1.** Definición de algoritmo no supervisado. Tipología de algoritmos no supervisados: los métodos de clusterización. Definiciones de similaridad y distancia. Presentación algebraica del problema.
2. **Bloque 2.** Algoritmos jerárquicos. Algoritmos de amalgamiento. Algoritmos de desamalgamiento. Consideraciones de convergencia e interpretación.
3. **Bloque 3.** Algoritmos no jerárquicos. Métodos de reasignación. El problema de los puntos semilla. Métodos de densidad: densidad topológica y densidad probabilística. El método EM-MGM

## 4. Metodología y Actividades

- Sesiones online teóricas con el profesor 2/04/19; 7/04/19: Videoconferencia donde se explicarán de forma esquemática conceptos relacionados con los algoritmos no supervisados de deep learning (actividades 1 y 6).
- Sesiones prácticas: Uso de plataforma *Blackboard* donde los alumnos podrán acceder a material complementario para profundizar en los aspectos mencionados en las sesiones online (actividades 2, 3, 4 y 5).

El curso se desarrollará entre los días 2/04/19 y 7/04/19 con la siguiente distribución de actividades aproximada:

Actividades	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Lunes
Actividad 1 Videoconferencia Inicial (19:00)							
Actividad 2 Lecturas voluntarias							
Actividad 3 Ejemplos resueltos y ejercicios propuestos							
Actividad 4: Material para uso de R en problemas específicos de clustering							
Actividad 5: Foro de dudas							
Actividad 6 Videoconferencia final (19:00)							

Día de Descanso

## Descripción de las actividades:

### Actividad 1: Videoconferencia inicial

Carácter: grupal

Herramientas: la videoconferencia se realizará a través de la plataforma Blackboard.

Se usará el documento **Algoritmos no supervisados1. Videoconferencia inicial. pdf**

Desarrollo y plazo de ejecución: el primer día de comienzo del módulo se realizará una videoconferencia para dar la bienvenida a los alumnos y para presentar la estructura del módulo y los contenidos a desarrollar. Asimismo, el profesor se centrará en explicar las herramientas comunes de los métodos no supervisados de clustering e introducirá los métodos jerárquicos. Día y hora: martes, 2 de abril de 2019 a las 19:00 (GMT+1).

Duración: 1 hora y 30 minutos.

### Actividad 2: Lecturas voluntarias

Carácter: individual

Herramientas: Se adjuntará documento a través de la plataforma Blackboard donde se detallan las lecturas y vídeos adicionales que se sugieren para entender mejor la materia y para profundizar en algunos conceptos.

Desarrollo y plazo de ejecución: a lo largo de la semana que dura el módulo. Los alumnos podrán exponer al profesor en el Foro de Dudas cualquier duda o comentario que deseen hacer sobre su contenido. La lectura es voluntaria y no se tendrá en cuenta para la evaluación.

**Actividad 3: Ejercicios resueltos y propuestos**

Carácter: Individual.

Herramientas:

1. Se adjuntará documento a través de la plataforma Blackboard mostrando ejercicios resueltos sobre los temas ya tratados.
2. Por otro lado, se adjuntarán ejercicios propuestos sin resolver. El alumno tendrá que resolver los ejercicios.

**Desarrollo y plazo de ejecución:** podrán hacer el ejercicio durante desde el martes 2 hasta el domingo 7 de abril de 2019 y podrán exponer al profesor en el Foro de Dudas.

**Actividad 4: Foro de dudas**

Carácter: individual

Herramientas: Foro en plataforma Blackboard..

**Desarrollo y plazo de ejecución:** durante la duración del módulo estará abierto el foro para que cualquier alumno pueda preguntar al profesor cualquier duda del módulo.

**Actividad 5: Material para uso de R en problemas específicos de clustering**

Carácter: individual

Herramientas: Se adjuntarán documentos explicando el funcionamiento de las librerías de cluster más usadas en R para el Deep Learning.

**Desarrollo y plazo de ejecución:** durante la duración del módulo, a partir del día 3 de abril se incorporará material sobre el uso de los paquetes de R más potentes para el análisis de clusters no supervisados, con ejemplos e interpretación de las salidas del software.

**Actividad 6: Videoconferencia final**

Carácter: grupal

Herramientas: la videoconferencia se realizará a través de la plataforma Blackboard.

Se usará software R. **Algoritmos no supervisados1. Videoconferencia final. pdf**

**Desarrollo y plazo de ejecución:** el último día de impartición del módulo tendrá lugar una videoconferencia dividida en 2 partes: a) Algoritmos de clustering no jerarquizados: clases, fundamentos y aplicación en R b) Compartir conclusiones y resolver las últimas dudas que sigan abiertas. Duración: 1 h y 30 minutos. Día y hora: domingo, 7 de abril de 2019 a las 19:00.

## 5. Evaluación

- La evaluación será continua y se hará en base a los siguientes porcentajes:

- Ejercicios ..... 50%
- Participación en foros ..... 50%

## 6. Bibliografía y materiales de consulta

- Se irán incluyendo en la plataforma Blackboard

## 7. CV del Profesor

David es graduado y master en Economía y Doctor en Matemáticas (Estadística e Investigación Operativa). Profesor del Departamento de Estadística e Investigación Operativa de la Universidad de Sevilla desde 2011 y de la EOI desde 2016, ha contribuido con numerosas publicaciones en revistas de prestigio sobre campos como la teoría de la decisión en condiciones de imprecisión, microeconometría, algoritmos predictivos en escenarios competitivos y estadística aplicada. Ponente habitual en foros sobre aprendizaje automático y avances en temas de conocimiento lógico y matemático, también realiza actividad profesional en Data Science para empresas a través de Avocado Labs, empresa de la que es co-fundador y CEO, y de Agilia Center, donde es director del área de Inteligencia Artificial y Algoritmia. También es evaluador literario para narrativa, ensayo y poesía en Azur Editorial