



Proyecto/Guía docente de la asignatura

Asignatura	FUNDAMENTOS DE APRENDIZAJE AUTOMÁTICO		
Materia	FUNDAMENTOS DE APRENDIZAJE AUTOMÁTICO		
Módulo	BLOQUE BÁSICO		
Titulación	MÁSTER EN INVESTIGACIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES		
Plan	719	Código	55334
Periodo de impartición	1 ^{er} CUATRIMESTRE	Tipo/Carácter	OBLIGATORIA
Nivel/Ciclo	MÁSTER	Curso	1º
Créditos ECTS	4 ECTS		
Lengua en que se imparte	CASTELLANO		
Profesor/es responsable/s	ALFONSO BAHILLO MARTÍNEZ IGNACIO DE MIGUEL JIMÉNEZ		
Datos de contacto (E-mail, teléfono...)	TELÉFONO: 983 423697 / 983 18 5574 E-MAIL: alfonso.bahillo@uva.es , ignacio.miguel@tel.uva.es		
Departamento	TEORÍA DE LA SEÑAL Y COMUNICACIONES E INGENIERÍA TELEMÁTICA		



1. Situación / Sentido de la Asignatura

1.1 Contextualización

La cantidad de datos que se está generando a nivel mundial está creciendo exponencialmente. Esto se debe a la cada vez mayor penetración de las tecnologías de información y comunicaciones en nuestras vidas. Cuando navegamos por Internet o interaccionamos con nuestro dispositivo móvil estamos generando datos, las redes de sensores permiten medir y recoger parámetros en cantidades antes inimaginables, los experimentos científicos cada vez recogen y dan lugar a una mayor cantidad de datos, y la aparición del denominado Internet de las Cosas, donde (casi) todo estará conectado a Internet, hacen que esa tendencia vaya en aumento. Esto supone una oportunidad para traducir esos datos en información útil y extraer un enorme valor añadido en ámbitos muy diversos, para empresas, gobiernos, centros docentes, médicos y de investigación, organizaciones no gubernamentales o de cualquier otro tipo, y en definitiva para las personas.

El aprendizaje automático, o *machine learning*, es un conjunto de técnicas que permiten que un ordenador sea capaz de aprender de los datos. Este tipo de técnicas tiene una gran relevancia en el escenario antes mencionado. El aprendizaje automático tiene un gran número de aplicaciones, tales como los motores de búsqueda, la clasificación de imágenes y la detección de objetos en las mismas, los sistemas predictivos, los sistemas de recomendación de productos o servicios, la segmentación de clientes, los diagnósticos médicos, el análisis del mercado de valores, la detección de intrusiones y de fraudes, o el reconocimiento del habla y de la escritura, por nombrar solo algunos.

La asignatura proporciona una introducción al aprendizaje automático y analiza aspectos básicos del mismo como la teoría de la generalización, los efectos de los errores y el ruido, el compromiso entre sesgo y varianza, los riesgos que lleva asociado el aprendizaje automático (y cómo evitarlos), además de las técnicas de validación de modelos. En la asignatura también se estudian y aplican diversas técnicas, como el aprendizaje supervisado (modelos lineales, redes neuronales, SVM, etc.) y el aprendizaje no supervisado.

1.2 Relación con otras materias

No requiere conocimientos previos adquiridos en otras asignaturas del máster, pero es muy recomendable haber cursado “Lenguajes de programación en código abierto”. También está muy relacionada con el resto de las asignaturas y materias del Bloque de “Análisis de Datos y Big Data”.

1.3 Prerrequisitos

Una vez admitido en el máster, no existen requisitos previos de obligado cumplimiento para cursar la asignatura.



2. Competencias

2.1 Generales

- CG1 - Capacidad crítica hacia el conocimiento actual como medio imprescindible para la detección de nuevos retos a resolver y por eso evaluar crítica y constructivamente resultados de investigación de otros.
- CG5 - Capacidad de analizar y aplicar los conocimientos técnicos específicos de su área en nuevos entornos y contextos, teniendo en cuenta los parámetros y variables más significativas de cada nueva situación.
- CG10 - Capacidad de conocer y emplear técnicas y herramientas relacionadas con el modelado, simulación, experimentación y validación de las propuestas técnicas, así como evaluarlas mediante unos parámetros de bondad establecidos.
- CG11 - Capacidad de desarrollar la capacidad de aprendizaje y trabajo en grupo tanto en entornos conocidos y restringidos, así como en consorcios internacionales en los que intervienen factores culturales.
- CG13 - Capacidad de proseguir en un aprendizaje a lo largo de toda la vida (Life Long Learning) a través de la asimilación de las técnicas y actitudes propias del trabajo autónomo y auto-dirigido.

2.2 Específicas

- CE-AD2 - Capacidad para comprender la teoría básica del aprendizaje automático y sus implicaciones prácticas en el diseño de sistemas.
- CE-AD3 - Capacidad para describir y aplicar diversos modelos de aprendizaje automático.
- CE-AD4 - Capacidad para describir y aplicar técnicas de optimización, regularización, validación y agregación en el desarrollo de sistemas basados en aprendizaje automático.
- CE-AD5 - Capacidad para emplear los métodos y los parámetros más relevantes para la evaluación y validación de modelos de análisis de datos en cada caso concreto.
- CE-AD6 - Capacidad para describir los conceptos fundamentales asociados a las áreas de análisis de datos y Big Data.



3. Objetivos

Al finalizar la asignatura el alumno deberá ser capaz de:

- Explicar qué es el aprendizaje automático y qué tipos hay.
- Describir la teoría básica del aprendizaje automático y sus implicaciones prácticas en el diseño de sistemas.
- Describir y aplicar diversos modelos de aprendizaje automático supervisado y no supervisado.
- Describir y aplicar técnicas de regularización, validación y agregación en el desarrollo de sistemas basados en aprendizaje automático.
- Implementar sistemas basados en aprendizaje automático.

En cada una de las lecciones de la asignatura se indicarán los objetivos específicos asociados a cada una de ellas.



4. Contenidos y/o bloques temáticos

Bloque 1: Fundamentos de Aprendizaje Automático

Carga de trabajo en créditos ECTS:

a. Contextualización y justificación

Véase la contextualización y justificación de la asignatura.

b. Objetivos de aprendizaje

Véanse los de la asignatura.

c. Contenidos y materiales de aprendizaje

TEMA 1: Introducción al aprendizaje automático

TEMA 2: ¿Es factible aprender? (Primera parte)

[Desigualdad de Hoeffding y la dimensión Vapnik-Chervonenkis]

TEMA 3: El modelo lineal: Clasificación y regresión lineal

TEMA 4: ¿Es factible aprender? (Segunda parte)

[Sesgo-varianza, underfitting-overfitting, y curvas de aprendizaje]

TEMA 5: El modelo lineal: Regresión logística

TEMA 6: Regularización

TEMA 7: Validación

TEMA 8: Redes neuronales

TEMA 9: Máquinas de vectores de soporte (SVM)

TEMA 10: Árboles de decisión

TEMA 11: Algunos aspectos a tener en cuenta en el diseño de sistemas de aprendizaje supervisado

TEMA 12: Clustering



Cada **lección** irá apoyada por el siguiente material:

- Vídeos con la explicación de los distintos apartados que componen la lección.
- Documento PDF con las diapositivas de la lección.
- Foro de dudas de la lección.
- Cuestionario sobre la lección.

Cada **semana** llevará asociado el siguiente material:

- Enunciado de los ejercicios de dicha semana.
- Plantilla para la resolución de los ejercicios.
- Foro de dudas de los ejercicios de la semana.

El **curso** llevará asociado el siguiente material:

- Bibliografía, accesible en su mayor parte a través de la biblioteca virtual de la UVa.
- Enlaces a otros materiales de interés.

d. Métodos docentes

La asignatura se desarrolla a lo largo de cuatro semanas (más una de evaluación). El lunes de cada una de las cuatro primeras semanas de la asignatura se activarán en el Campus Virtual de la UVa 3 temas, según se detalla más adelante en el plan de trabajo, junto con todo el material asociado a la mismas. Concretamente, estarán disponibles:

- *Vídeos (píldoras de conocimiento) grabados por el profesorado de la asignatura y asociados a cada tema, junto con un documento PDF con las diapositivas de la misma.* Los alumnos deberán visualizarlos a lo largo de la semana.
- *Cuestionarios sobre cada tema para verificar si el alumno ha comprendido el material de los vídeos.* Aunque tendrá repercusión en la calificación, su objetivo es fundamentalmente formativo. Deberán responderse, preferiblemente, como muy tarde antes de la medianoche del lunes de la siguiente semana.
- *Ejercicios de la semana.* Se proporcionará el enunciado en formato PDF y además se dará una plantilla que deberá utilizarse para entregar la respuesta. Incluirán cuestiones y sobre todo casos prácticos que deberán resolverse mediante programación con el objetivo de ejercitarse en los conceptos y técnicas descritas en los vídeos. Los ejercicios se realizarán de forma individual, y se deberán entregar, en general y preferiblemente, antes del lunes de la siguiente semana a medianoche.

Se fomentará la colaboración entre el alumnado favoreciendo la colaboración en la resolución de dudas y también en la resolución de los propios ejercicios (aunque la entrega final debe ser individual, se permitirá la colaboración, con ciertas condiciones, a través de los foros). Se premiará al alumnado por sus aportaciones constructivas a los foros.

Para la comunicación con el profesorado y con el resto del alumnado, así como para la tutorización, se procederá de la siguiente forma. El elemento fundamental de comunicación serán los foros y las videoconferencias (aunque también se podrá utilizar el correo electrónico). Con respecto a las videoconferencias:



- Habrá 2 tutorías grupales por videoconferencia en las semanas 1 (comienzo de la asignatura) y 4 (al final de la asignatura), que serán grabadas por si algún alumno no pudiera asistir.
- Se podrá solicitar tutorías por videoconferencia con el profesorado.

Con respecto a los foros, el profesorado intentará responder a las dudas lo más rápido posible, en un plazo de 48 horas -lectivas- como máximo. Se dispondrá de los siguientes foros:

- *Foro de presentaciones (tarea inicial)*: La tarea inicial consistirá en presentarse en un foro, para conocer al profesorado y al resto de alumnos de la asignatura.
- *Tablón de anuncios*: Foro para indicar todas las novedades del curso, incluyendo la publicación de nuevos bloques de contenido, y para avisar de cualquier modificación o incidencia que ocurra durante el curso. El profesorado es el único que puede publicar en él.
- *Foro social*: Foro para comunicarse con los compañeros (y con el profesor) sobre temas relacionados con la asignatura pero que no encajen en otros foros. Por ejemplo, puede utilizarse para informar de noticias curiosas o relevantes que aparezcan en prensa sobre aprendizaje automático, para informar sobre páginas web o recursos que puedan ser útiles, para crear grupos de estudio sobre la asignatura, etc.
- *Foro de dudas*:
 - *Sobre el desarrollo de la asignatura*: Para plantear únicamente las dudas "burocráticas" sobre la asignatura o dudas muy generales (ej. problemas con el acceso a materiales, fechas de entrega, dudas sobre la bibliografía recomendada, etc.).
 - *Sobre cada bloque temático* (cada semana se activará un foro para las dudas relacionadas con los temas y ejercicios de dicha semana).



e. Plan de trabajo (Actividades Formativas y de Evaluación)

- Temporalización de las actividades

Semana 1	TEMA 1: Introducción al aprendizaje automático TEMA 2: ¿Es factible aprender? (Primera parte) TEMA 3: El modelo lineal: Clasificación y regresión lineal	Tutoría virtual grupal Cuestionarios de los temas 1 a 3 Ejercicios de la semana 1
Semana 2	TEMA 4: ¿Es factible aprender? (Segunda parte) TEMA 5: El modelo lineal: Regresión logística TEMA 6: Regularización	Cuestionarios temas 4 a 6 Ejercicios de la semana 2
Semana 3	TEMA 7: Validación TEMA 8: Redes neuronales TEMA 9: Máquinas de vectores de soporte (SVM)	Cuestionarios temas 7 a 9 Ejercicios de la semana 3
Semana 4	TEMA 10: Árboles de decisión TEMA 11: Algunos aspectos a tener en cuenta en el diseño de sistemas de aprendizaje supervisado TEMA 12: Clustering	Tutoría virtual grupal Cuestionarios temas 10 a 12 Ejercicios de la semana 4
Semana 5		Entrega final de los ejercicios

- Tiempo de dedicación a cada actividad

Actividad	Dedicación	Evaluable	Tipo	Peso Evaluación
1.- Tutoría (asistencia síncrona o asíncrona)	4 h	No		
2.- Visionado de vídeos y estudio de las diapositivas	20 h	No		
3.- Evaluación y autoevaluación (y, en su caso, evaluación por pares)	6 h	Sí	Cuestionarios	20%
4.- Resolución de ejercicios, incluyendo la búsqueda de información para resolverlos	65 h	Sí	Entrega de ejercicios	75%
5.- Seguimiento y participación en los foros	5 h	Sí (la participación)		5%
TOTAL BLOQUE	100 h			

g. Bibliografía básica

- Y.S. Abu-Mostafa, M. Magdon-Ismail y H.T. Lin, *Learning from Data: A Short Course*, Amlbook.com, 2012.
- A. Géron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow*, 3rd Ed., O'Reilly, 2022.
- M. Kuhn y K. Johnson, *Applied Predictive Modeling*, Springer, 2016.



h. Bibliografía complementaria

- T. Hastie, R. Tibshirani y J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second Edition. Springer 2009.
- I.H. Witten, E. Frank, M.A. Hall y C.J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, Fourth Edition, Morgan Kaufmann, 2016.

i. Recursos necesarios

Serán necesarios los siguientes recursos:

- Campus Virtual de la Universidad de Valladolid.
 - Material de apoyo disponible en el mismo.
- Herramientas de comunicación:
 - Asíncronos: Foros (en el marco del Campus Virtual) y correo electrónico.
 - Síncronos: Videoconferencias.
- Ordenador.



j. Temporalización

CARGA ECTS	PERIODO PREVISTO DE DESARROLLO (detallar orden semanas)
0,85	Semana 1: Temas 1 a 3
0,85	Semana 2: Temas 4 a 6
0,85	Semana 3: Temas 7 a 9
0,85	Semana 4: Temas 10 a 12
0,2	Semana 5: Entrega de últimos ejercicios

5. Métodos docentes y principios metodológicos

Véase el apartado 4.d

6. Tabla de dedicación del estudiante a la asignatura

ACTIVIDADES	HORAS
Horas de tutoría síncrona o asíncrona	4
Horas de lectura y reproducción materiales	20
Horas de trabajo autónomo individual (<i>resolución de ejercicios y foros, principalmente</i>)	65
Horas de trabajo autónomo en equipo (<i>colaboración puntual con otros alumnos en la resolución de ejercicios</i>)	5
Horas de actividades de evaluación	6
Total	100



7. Sistema y características de la evaluación

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Ejercicios semanales	75%	Es condición necesaria para superar la asignatura alcanzar al menos una puntuación de 4 sobre 10 en este apartado.
Cuestionarios de los temas	20%	
Participación en los foros	5%	

CRITERIOS DE CALIFICACIÓN
<ul style="list-style-type: none">● Convocatoria ordinaria: Véase la tabla.● Convocatoria extraordinaria: Seguirá vigente la tabla, con las siguientes particularidades:<ul style="list-style-type: none">○ Se podrán entregar de nuevo aquellos cuestionarios y ejercicios semanales que no se hubieran presentado en la convocatoria ordinaria o que no hubieran alcanzado la calificación de 5 sobre 10 en la convocatoria ordinaria.○ Se mantiene la calificación de participación en los foros.

8. Consideraciones finales