



Proyecto docente

Asignatura	Aprendizaje sobre Flujos de Datos		
Materia	Ciencia de Datos / Data Science		
Titulación	Máster Universitario en Inteligencia de Negocio y Big Data en Entornos Seguros		
Plan		Código	
Periodo de impartición	Segundo semestre	Tipo/Carácter	Obligatoria
Nivel/Ciclo	Máster	Curso	1
Créditos ECTS	3		
Lengua en que se imparte	Castellano		
Profesor/es responsable/s	Juan José Rodríguez Díez		
Datos de contacto (e-mail, teléfono...)	jjrodriguez@ubu.es 947 11 21 10		
Horario de tutorías			
Coordinador	Juan José Rodríguez Díez		
Departamento	Área de Lenguajes y Sistemas Informáticos, Ingeniería Civil, Universidad de Burgos.		
Web			
Descripción General	Conocer, aplicar e implementar los métodos de aprendizaje para el análisis de flujos de datos, posiblemente cambiantes.		



1. Situación / Sentido de la asignatura

1.1 Contextualización

La asignatura se engloba dentro de la materia de Ciencia de Datos. En esta asignatura se considera concretamente el aprendizaje sobre flujos de datos: cuando los datos sobre los que aprender no están disponibles todos inicialmente sino que van llegando de manera continua. Además, estos datos pueden evolucionar debido a cambios en el entorno.

1.2 Relación con otras asignaturas

La asignatura está relacionada principalmente con las de la materia a la que pertenece: Ciencia de Datos. Estas asignaturas son:

- Técnicas de Aprendizaje Automático Escalables
- Knowledge Discovery / Aprendizaje No Supervisado

La relación se debe a que en estas asignaturas se consideran tareas y métodos para las que existen sus correspondientes versiones sobre flujos de datos.

También hay relación con asignaturas de programación, principalmente con

- Modelos de Programación para el Big Data

La relación en este caso viene motivada porque para realizar el aprendizaje sobre flujos se programará sobre bibliotecas que implementan este tipo de métodos.

1.3 Prerrequisitos

Nivel de inglés medio, dado que la mayoría de los materiales complementarios estarán en dicho idioma.

Se necesitan conocimientos de programación, ya que el aprendizaje se puede llevar a cabo realizando programas o scripts que usan bibliotecas de análisis de datos sobre flujos. También hay que ser capaz de implementar estos métodos de aprendizaje.

No es recomendable cursar esta asignatura sin haber cursado las que se indican en el apartado previo de asignaturas relacionadas: "Modelos de Programación para el Big Data", "Técnicas de Aprendizaje Automático Escalables" y "Knowledge Discovery / Aprendizaje No Supervisado".



2. Competencias

2.1 Generales del título

- CG1. Adquisición de competencias teóricas y prácticas para el análisis y diseño de soluciones empresariales en Big Data (almacenamiento y procesamiento de grandes volúmenes de información heterogénea).
- CG3. Capacidad de diseñar e implementar sistemas capaces de extraer conocimiento práctico de grandes volúmenes de datos aplicado al mundo de la empresa (Inteligencia de Negocio/Business Intelligence)

2.2 Específicas materia

- CDS1. Capacidad de aplicar, validar y evaluar métodos de Ciencia de Datos/Data Science e Inteligencia Artificial sobre conjuntos y flujos de datos masivos y complejos.
- CDS2. Capacidad de dirigir proyectos para la extracción de conocimiento basados en métodos eficientes de análisis de datos.



3. Resultados de aprendizaje

Al finalizar la asignatura, el alumno será capaz de:

- Clasificar automáticamente flujos de datos.
- Predecir valores futuros en series temporales.
- Determinar la similitud de datos en flujos y/o agruparlos en función de su similitud.
- Resumir o sintetizar flujos de datos de gran volumen.
- Implementar algoritmos y técnicas de aprendizaje automático adaptadas a grandes flujos de datos.



4. Contenido / Programa de la asignatura

4.1 Unidades docentes (bloques de contenidos)

- Introducción al análisis sobre flujos.
- Detección de cambios.
- Clasificación sobre flujos.
- Combinaciones de clasificadores.
- Regresión sobre flujos.
- Clustering sobre flujos.
- Patrones frecuentes en flujos.
- Herramientas y bibliotecas para el análisis sobre flujos de datos.

4.2 Bibliografía

- Albert Bifet, Ricard Gavaldà, Geoff Holmes and Bernhard Pfahringer. Machine Learning for Data Streams. The MIT Press (2018) <https://moa.cms.waikato.ac.nz/book-html/>
- Minos Garofalakis, Johannes Gehrke, Rajeev Rastogi, (Eds.) Data Stream Management: Processing High-Speed Data Streams. Springer (2016).
- Joao Gama. Knowledge Discovery from Data Streams. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series (2010).



5. Metodología de enseñanza y dedicación del estudiante a la asignatura

Actividad Formativa	Competencias relacionadas	Horas	Presencialidad (%)
Clases, conferencias y técnicas expositivas		12	0
Actividades autónomas y en grupo (trabajos y lecturas dirigidas)		45	0
Pruebas de seguimiento y exposición de trabajos		10	50
Tutoría individual, participación en foros y otros medios colaborativos		8	0



6. Temporalización (por bloques temáticos)

BLOQUE TEMÁTICO	CARGA ECTS	PERIODO PREVISTO DE DESARROLLO



7. Evaluación

Instrumento / Procedimiento	Peso primera convocatoria	Peso segunda convocatoria
Evaluación sumativa, que incluye pruebas parciales individuales y prueba final	40%	40%
Realización de trabajos, proyectos, resolución de problemas y casos	40%	40%
Participación en foros y otros medios participativos	20%	20%

Criterios / Comentarios a la evaluación

- **Convocatoria ordinaria:** La calificación final será la media ponderada al porcentaje indicado en las tablas. Para la superación de la asignatura se exigirá un mínimo de 4 puntos sobre 10 en el procedimiento de "Evaluación sumativa...". En el procedimiento "Realización de trabajos..." será necesario un 4 sobre 10 en cada uno de los trabajos.
- **Convocatoria extraordinaria:** La calificación final será la media ponderada al porcentaje indicado en las tablas. Para la superación de la asignatura se exigirá un mínimo de 4 puntos sobre 10 en el procedimiento de "Evaluación sumativa...". En el procedimiento "Realización de trabajos..." será necesario un 4 sobre 10 en cada uno de los trabajos.
- Es posible que el procedimiento "Participación en foros y otros medios participativos" no sea recuperable en su totalidad en 2ª convocatoria. La evaluación se basa en la interacción entre los alumnos y es posible que esta no pueda organizarse de forma satisfactoria por restricciones de tiempo o de número de alumnos en ese período. En ese caso, se conservará la nota obtenida en la 1ª convocatoria.



8. Recursos de aprendizaje y apoyo tutorial del curso online

Transparencias.
Enunciados de ejercicios.
Cuestionarios de autoevaluación.
Páginas Webs relacionadas
Bibliografía disponible en la Biblioteca
Tutorías individualizadas o en grupo a demanda de los alumnos.



9. Consideraciones / Comentarios adicionales
