

ANX-PR/CL/001-02
GUÍA DE APRENDIZAJE

ASIGNATURA

Sistemas basados en aprendizaje automatico

CURSO ACADÉMICO - SEMESTRE

2015-16 - Primer semestre

Datos Descriptivos

| | |
|--|---|
| Nombre de la Asignatura | Sistemas basados en aprendizaje automatico |
| Titulación | 09TT - Grado en Ingenieria de Tecnologias y Servicios de Telecomunicacion |
| Centro responsable de la titulación | E.T.S. de Ingenieros de Telecomunicacion |
| Semestre/s de impartición | Séptimo semestre |
| Módulo | Optativas |
| Materia | Optativas |
| Carácter | Optativa |
| Código UPM | 95000244 |
| Nombre en inglés | Machine Learning Systems |

Datos Generales

| | | | |
|------------------------------|------------|-------------------------------------|------------------|
| Créditos | 3 | Curso | 4 |
| Curso Académico | 2015-16 | Período de impartición | Septiembre-Enero |
| Idioma de impartición | Castellano | Otros idiomas de impartición | |

Requisitos Previos Obligatorios

Asignaturas Superadas

El plan de estudios Grado en Ingenieria de Tecnologias y Servicios de Telecomunicacion no tiene definidas asignaturas previas superadas para esta asignatura.

Otros Requisitos

El plan de estudios Grado en Ingenieria de Tecnologias y Servicios de Telecomunicacion 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

El coordinador de la asignatura no ha definido otros conocimientos previos recomendados.

Competencias

CEB2 - Conocimientos básicos sobre el uso y programación de los ordenadores, sistemas operativos, bases de datos y programas informáticos con aplicación en ingeniería

CECT3 - Capacidad para utilizar herramientas informáticas de búsqueda de recursos bibliográficos o de información relacionada con las telecomunicaciones y la electrónica

CG2 - Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio

CG7 - Trabajo en equipo

CG8 - Comunicación oral y escrita

Resultados de Aprendizaje

RA607 - Conocimiento de los principales programas informáticos y herramientas para el análisis de datos y aprendizaje de la máquina basado en los mismos.

RA45 - Conocimientos y habilidades de las temáticas científico tecnológicas desarrolladas en las asignaturas ofertadas

RA605 - Conocimientos de los sistemas de clasificación y agrupamiento de datos y de las principales características de cada uno de los métodos principalmente utilizados

RA358 - Capacidad para el análisis de problemas, trabajo en equipo y exposición de los resultados del análisis

RA604 - Conocimientos de la problemática de análisis de grandes cantidades de datos

RA606 - Analizar y evaluar los distintos algoritmos de aprendizaje automático basado en datos

Profesorado

Profesorado

| Nombre | Despacho | e-mail | Tutorías |
|---|----------|--------------------------|--|
| Pardo Muñoz, Jose Manuel (Coordinador/a) | C-224 | josemanuel.pardom@upm.es | L - 10:00 - 13:00 M - 10:00 - 13:00 X - 10:00 - 13:00 J - 10:00 - 13:00 V - 10:00 - 13:00 Pedir cita previa |
| Ferreiros Lopez, Javier | B-110 | javier.ferreiros@upm.es | L - 10:00 - 13:00 M - 10:00 - 13:00 X - 10:00 - 13:00 J - 10:00 - 13:00 V - 10:00 - 13:00 Pedir cita previa |
| Cordoba Herralde, Ricardo De | B-108 | ricardo.cordoba@upm.es | L - 10:00 - 13:00 M - 10:00 - 13:00 X - 10:00 - 13:00 J - 10:00 - 13:00 V - 10:00 - 13:00 Pedir cita previa |

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

En esta asignatura se tratan de manera introductoria las técnicas de aprendizaje automático con una importante componente práctica que permite al estudiante entrar en contacto con la realidad de uso y diseño de este tipo de sistemas.

La asignatura cubre los distintos aspectos de diseño de los sistemas de aprendizaje automático basado en datos, desde los detalles relacionados con la entrada de datos, el análisis de los mismos, la obtención de características principales, la agrupación automática, creación de patrones, y generación de un sistema automático que los aprenda y los reconozca. Finalmente se estudia la evaluación del comportamiento del sistema al realizar la tarea asignada, estimación de su rendimiento y eficacia y los métodos de ajuste del mismo para optimizar su validez.

Estos sistemas son fundamentales en muy diversos campos de aplicación, desde la automatización inteligente de procesos industriales hasta el análisis de grandes cantidades de datos para extraer información y clasificarla en lo que se ha dado en llamar "Big Data?".

Cada vez más la industria requiere profesionales con experiencia previa en este campo del aprendizaje automático, reconocimiento automático de patrones y extracción de Información a partir de datos.

Las áreas de aplicación son innumerables siempre que se disponga de suficientes datos, por ejemplo datos de redes sociales, datos de patrones de enfermedades, optimización de procesos, optimización de recursos, análisis de comportamiento de clientes, organización operativa, detección de fallos en sistemas, análisis de contenido de textos escritos, análisis de grabaciones de audio y vídeo etc.

Temario

1. Introducción

- 1.1. Visión general del curso
- 1.2. Minería de datos y aprendizaje automático
- 1.3. Ejemplos simples: Tiempo, iris, negociaciones laborales, clasificación de granos
- 1.4. Áreas reales de aplicación: Análisis de webs, análisis de imágenes, diagnosis de enfermedades, Mercadotecnia y ventas

2. Clasificación y regresión

- 2.1. Datos de entrada: Ejemplos, atributos, clases
- 2.2. Concepto de clasificación, agrupamiento y regresión
- 2.3. 1-R
- 2.4. ZeroR, Regla de Bayes y naive Bayes
- 2.5. Clasificación lineal
- 2.6. Regresión lineal

3. PRÁCTICA 1: Manejo de la herramienta WEKA
 - 3.1. Introducción al interface ?explorer? de Weka
 - 3.2. Carga de datos, editor de datos, filtrado de datos
 - 3.3. Panel de visualización
 - 3.4. Panel de clasificación
4. Evaluación
 - 4.1. Concepto de entrenamiento/evaluación
 - 4.2. Validación cruzada
 - 4.3. Significancia estadística
 - 4.4. Sobre-entrenamiento (overfitting)
5. PRÁCTICA 2: Ejemplos de aplicación
 - 5.1. Definición de la aplicación
 - 5.2. Preparación de los datos de entrenamiento y prueba
 - 5.3. Pruebas de algoritmos simples con datos sencillos
6. Selección y transformación de rasgos característicos
 - 6.1. Basado en estimación de bondad. Selección de componentes principales
 - 6.2. Basado en clasificación. Transformación lineal discriminativa (LDA)
7. PRÁCTICA 3: Estudio y selección de rasgos para mejorar el sistema
 - 7.1. Prueba con distintos atributos
 - 7.2. Selección manual de atributos
 - 7.3. ? Selección automática de atributos: Selección de componentes principales, Transformación lineal discriminativa (LDA)
8. Clasificación avanzada y agrupamiento
 - 8.1. Árboles de decisión y regresión
 - 8.2. Regla de los K vecinos más próximos (K-NN)
 - 8.3. Perceptron multicapa (ANN)
 - 8.4. Técnicas de agrupamiento, Simple kmeans, Estimate-Maximize (EM)
9. INTRODUCCION A LA PRACTICA 4
 - 9.1. Tutorial sobre distintos datos disponibles
 - 9.2. Tutorial sobre posibles algoritmos a usar en cada problema
10. PRACTICA 4. Desarrollo del un sistema evolucionado de aprendizaje y clasificación

Cronograma

Horas totales: 36 horas y 45 minutos

Horas presenciales: 36 horas y 45 minutos (47.1%)

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 | Tema 1, Introduccion Duración: 02:00 LM: Actividad del tipo Lección Magistral | | | |
| Semana 2 | Tema 2. Clasificacion y regresion Duración: 02:00 LM: Actividad del tipo Lección Magistral | | | |
| Semana 3 | | Practica 1: Introduccion al imterface "explorer" de Weka Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio | | Entrega formulario resultado primera practica Duración: 00:15 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial |
| Semana 4 | Conceptos de entrenamiento!evaluacion Duración: 02:00 LM: Actividad del tipo Lección Magistral | | | |
| Semana 5 | | Practica 2: Ejemplos de aplicacion Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio | | Entrega de formulario de resultado Duración: 00:15 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial |
| Semana 6 | Seleccion de rasgos caracteristicos Duración: 02:00 LM: Actividad del tipo Lección Magistral | | | |
| Semana 7 | | PRÁCTICA 3: Estudio y selección de rasgos para mejorar el sistema Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio | | Entrega de formulario del resultado Duración: 00:15 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial |
| Semana 8 | Clasificacion avanzada y agrupamiento: Arboles de decision y regresion Duración: 01:00 LM: Actividad del tipo Lección Magistral | Practica con arboles de decision y regresion Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio | | |
| Semana 9 | Clasificacion avanzada y agrupamiento: Regla del vecino más próximo Duración: 01:00 LM: Actividad del tipo Lección Magistral | Practica con regla del vecino más próximo Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio | | |

| | | | | |
|-----------|---|--|--|--|
| Semana 10 | Clasificación avanzada y agrupamiento: Perceptron Multicapa Duración: 01:00 LM: Actividad del tipo Lección Magistral | Practica con perceptron multicapa Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio | | |
| Semana 11 | Clasificación avanzada y agrupamiento: Tecnicas de agrupamiento, K.means y EM Duración: 01:00 LM: Actividad del tipo Lección Magistral | Práctica con Simple k means, EM Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio | | |
| Semana 12 | Introducción a practica 4 Duración: 01:00 LM: Actividad del tipo Lección Magistral | Trabajo en proyecto final Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio | | |
| Semana 13 | | Trabajo en proyecto final Duración: 04:00 PL: Actividad del tipo Prácticas de Laboratorio | | |
| Semana 14 | | Trabajo en proyecto final Duración: 04:00 PL: Actividad del tipo Prácticas de Laboratorio | | |
| Semana 15 | | | | Entrega y breve presentacion del proyecto final Duración: 02:00 TG: Técnica del tipo Trabajo en Grupo Evaluación continua Actividad presencial |
| Semana 16 | | | | |
| Semana 17 | | | | Test sobre la parte teórica Duración: 02:00 EX: Técnica del tipo Examen Escrito Evaluación continua y sólo prueba final Actividad presencial Entrega proyecto final Duración: 02:00 TG: Técnica del tipo Trabajo 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 |
|--------|---|----------|---|---------------------------------------|------------|------|-------------|----------------------------|
| 3 | Entrega formulario resultado primera practica | 00:15 | Evaluación continua | TG: Técnica del tipo Trabajo en Grupo | Sí | 10% | | CEB2, CG7, CG8, CECT3 |
| 5 | Entrega de formulario de resultado | 00:15 | Evaluación continua | TG: Técnica del tipo Trabajo en Grupo | Sí | 10% | | CEB2, CG8, CG7, CECT3 |
| 7 | Entrega de formulario del resultado | 00:15 | Evaluación continua | TG: Técnica del tipo Trabajo en Grupo | Sí | 10% | | CG8, CG7, CEB2, CECT3 |
| 15 | Entrega y breve presentacion del proyecto final | 02:00 | Evaluación continua | TG: Técnica del tipo Trabajo en Grupo | Sí | 50% | | CEB2, CECT3, CG2, CG8, CG7 |
| 17 | Test sobre la parte teórica | 02:00 | Evaluación continua y sólo prueba final | EX: Técnica del tipo Examen Escrito | Sí | 20% | | CG2 |
| 17 | Entrega proyecto final | 02:00 | Evaluación sólo prueba final | TG: Técnica del tipo Trabajo en Grupo | Sí | 80% | | CG8, CG2, CG7, CECT3, CEB2 |

Criterios de Evaluación

Los alumnos serán evaluados, por defecto, mediante evaluación continua. Dado el carácter eminentemente práctico de la asignatura se desaconseja la evaluación mediante un único examen final.

Las prácticas se evaluarán según el cuestionario entregado. El proyecto final se evaluará según la calidad del mismo, la presentación oral, los resultados obtenidos y la calidad de la memoria.

En cumplimiento de la Normativa de Evaluación de la Universidad Politécnica de Madrid, los alumnos que lo deseen serán evaluados mediante un único examen final siempre y cuando lo comuniquen por escrito al Coordinador de la Asignatura según la normativa vigente. Esta solicitud puede realizarse en cualquier momento hasta la quinta semana incluida. El examen final consistirá en la realización y presentación oral y escrita de un proyecto de minería de datos y aprendizaje automático y la realización de un test teórico de la asignatura.

Recursos Didácticos

| Descripción | Tipo | Observaciones |
|---|--------------|--|
| Texto | Bibliografía | "Data Mining: Practical Machine Learning Tools and Techniques" Mark Hall, Ian Witten and Eibe Frank. Morgan Kaufmann Publishers. January 2011 |
| WEKA | Equipamiento | Herramientas de análisis de datos y aprendizaje de máquina |
| Manual de Weka | Bibliografía | Manual de la herramienta Weka |
| Página web de Weka. Documentación y Software | Recursos web | http://www.cs.waikato.ac.nz/ml/weka/documentation.html http://www.cs.waikato.ac.nz/ml/weka/downloading.html |
| Repositorio de datos | Recursos web | Repositorio de datos: http://archive.ics.uci.edu/ml/ |