

MINERÍA INFORMACIÓN SOCIAL

Curso 2017/2018

(Código: 31101324)

1. PRESENTACIÓN

La asignatura "Minería de Información Social" se enmarca dentro del Programa de Posgrado en Sistemas Inteligentes impartido por la Escuela Técnica Superior de Ingeniería Informática de la UNED. Dentro de este posgrado se imparte en la especialidad de "Tecnologías del Lenguaje en la Web (Master en Lenguajes y Sistemas Informáticos)"

Ficha técnica:

- Tipo: Optativa
- Duración: Anual
- Créditos Totales y Horas: 6 ECTS / 150 horas
- Horas de estudio teórico: 75
- Horas de trabajo práctico: 75

Reseña del Profesorado:

FRESNO FERNÁNDEZ, VÍCTOR

Su actividad docente se ha centrado principalmente en el campo de la documentación electrónica y los lenguajes de programación, y su investigación en el campo de la representación automática de textos, así como en la clasificación y clustering de documentos. En la actualidad centra sus esfuerzos en la aplicación de estas técnicas al contexto del análisis de información en redes sociales y el contexto médico. Desde el año 2000 hasta la actualidad ha trabajado en el Instituto de Automática industrial (CSIC), la Universidad Rey Juan Carlos (URJC) y la Universidad Nacional de Educación a Distancia (UNED), colaborando en los programas de doctorado de dichas universidades.

e.mail: vfresno@lsi.uned.es

MARTÍNEZ UNANUE, RAQUEL:

Ha realizado la mayor parte de su actividad docente en el campo de la programación, los lenguajes de programación, la algoritmia y la documentación electrónica. Su actividad investigadora se ha desarrollado principalmente en dos campos: la alineación de corpus paralelos para generación de memorias de traducción y en minería de textos tanto monolingües como multilingües. Desde el año 2000 hasta la actualidad ha colaborado en programas de doctorado de tres universidades: la Universidad Complutense de Madrid, la Universidad Rey Juan Carlos y la UNED.

e.mail: raquel@lsi.uned.es

2. CONTEXTUALIZACIÓN

Esta asignatura se encuadra en la especialidad "Tecnologías del lenguaje en la web" de la titulación de posgrado "Master en Lenguajes y Sistemas Informáticos".

3. REQUISITOS PREVIOS RECOMENDABLES

Además de los requisitos generales de acceso a este programa de posgrado orientado a la investigación, se requiere conocimiento previo en los siguientes aspectos:

- Diseño e implementación de sistemas informáticos
- Lectura fluida del inglés.

Esta asignatura puede ser cursada aisladamente, aunque el estudiante se beneficiaría si hubiera cursado previamente o en paralelo las asignaturas de Procesamiento del Lenguaje Natural y Descubrimiento de información en textos.

4. RESULTADOS DE APRENDIZAJE

El objetivo del curso es proporcionar al alumno una visión global de las técnicas y tecnologías involucradas en la minería de información social.

El aprendizaje está diseñado para permitir que el alumno adquiera una serie de destrezas y competencias que se enumeran a continuación:

1. Tener una visión amplia de las áreas relacionadas con el acceso y análisis de información social. Esto incluye tanto el contenido compartido por usuarios, como los perfiles de los usuarios en sí, adquiriendo las destrezas y competencias que se enumeran a continuación:

- Ser capaz de acceder y analizar información extraída de diferentes redes sociales.
- Conocer diferentes aproximaciones de análisis de contenido y de usuario en el contexto de las redes sociales.
- Conocer diversos tipos de técnicas y herramientas específicas para el análisis de información social.

2. Hábituarse a la lectura de artículos científicos y desarrollar la capacidad para buscar información que complete el material propuesto inicialmente.

3. Capacidad de reflexión sobre el material estudiado, necesaria para poder realizar una síntesis de calidad.

4. Desarrollar pequeñas aplicaciones para el acceso y análisis de información social.

Con la superación del curso se espera que el alumno complete todas las siguientes competencias generales:

- Adquirir capacidad de abstracción, análisis, síntesis y relación de ideas.
- Adquirir capacidad crítica y de decisión
- Adquirir capacidad de estudio y autoaprendizaje
- Adquirir capacidad creativa y de investigación
- Adquirir habilidades sociales para el trabajo en equipo

Así como las siguientes competencias específicas:

- Adquirir la capacidad de comprender y manejar de forma básica los aspectos más importantes relacionados con los lenguajes y sistemas informáticos en general y, de manera especial, en los siguientes ámbitos: Tecnologías del lenguaje y de acceso a la información en web.
- Adquirir capacidad de estudio de los sistemas y aproximaciones existentes para distinguir las aproximaciones más efectivas.
- Adquirir capacidad para detectar carencias en el estado actual de la ciencia y la tecnología.
- Adquirir capacidad para proponer nuevas aproximaciones que den solución a las carencias detectadas.
- Adquirir capacidad de especificar, diseñar, implementar y evaluar tanto cualitativa como cuantitativamente los modelos y sistemas propuestos.
- Adquirir capacidad para proponer y llevar a cabo experimentos con la metodología adecuada como para poder extraer conclusiones y determinar nuevas líneas de actuación e investigación.

5.CONTENIDOS DE LA ASIGNATURA

La estructura y el contenido teórico de la asignatura se detalla a continuación:

Tema 1- Introducción al *Social Media*: definiciones, terminología y conceptos

Tema 2- Generación y anotación de corpus de Social Media

- Recopilación de datos mediante APIs
- Acceso a la información y almacenamiento
- Evaluación

Tema 3- Análisis de contenido en Social Media

- Aplicación de técnicas de PLN
- Procesamiento semántico
 - Entidades Nombradas
 - Minería de opinión y análisis de sentimientos
- Análisis de tendencias
- Clasificación y clustering textual

Tema 4- Análisis de información de usuario en Social Media

- Análisis de la reputación y estructura utilizando contenido de microblogs
- Recomendación basada en contenido e información social
- Folksonomías y etiquetado social

Las actividades formativas (con nº de horas por cada una):

Tema 1

T1.1. Lectura de material sugerido (5 horas teóricas)

Tema 2

T2.1 – Lectura de material sugerido (10 horas teóricas)

T2.2 – Práctica Obligatoria (10 horas prácticas)

Tema 3

T3.1 – Lectura de material sugerido (30 horas teóricas)

T3.2 – Práctica Obligatoria (20 horas prácticas)

Tema 4

T4.1 – Lectura de material sugerido (30 horas teóricas)

T4.2 – Práctica Obligatoria (20 horas prácticas)

6.EQUIPO DOCENTE

- [VICTOR DIEGO FRESNO FERNANDEZ](#)
- [RAQUEL MARTINEZ UNANUE](#)

7.METODOLOGÍA

La metodología es la general del programa de postgrado; junto a las actividades y enlaces con fuentes de información externas, puede existir material didáctico propio preparado por el equipo docente y una bibliografía básica (en inglés). Se trata de una metodología adaptada a las directrices del EEES, de acuerdo con el documento del IUED.

La asignatura no tiene clases presenciales. Los contenidos teóricos se impartirán a distancia, de acuerdo con las normas y estructuras de soporte telemático de la enseñanza en la UNED.

El material docente incluye un resumen de los contenidos de cada tema y distintos tipos de actividades relacionadas con la consulta bibliográfica, la realización de ejercicios prácticos y, en ocasiones, la utilización de herramientas asociadas a las tecnologías y técnicas presentadas en el curso. Junto a las actividades y enlaces con fuentes de información externas, existe material didáctico propio preparado por el equipo docente con ejercicios que organizan las actividades prácticas del alumno.

Las actividades de aprendizaje se estructuran alrededor del estado del arte en cada uno de los temas del curso y a la parte práctica que el alumno deberá realizar en cada sección. Se utilizarán los recursos de la UNED para las actividades complementarias que se propondrán a lo largo del curso, actividades de asistencia o visualización (on-line, off-line) de seminarios de temas avanzados, impartidos por investigadores de reconocido prestigio.

El plan de trabajo marcará unas pautas adecuadas para que el alumno medio alcance los objetivos al final del curso. Incluye una descripción ordenada y secuenciada de los contenidos de cada tema y orientaciones sobre las actividades a realizar, tanto de estudio teórico como práctico. Las actividades planteadas son para apoyo y refuerzo del aprendizaje, siendo algunas de ellas de obligada entrega para la supervisión de los profesores.

Las actividades pueden ser de los siguientes tipos:

- Lectura y estudio del material correspondiente
- Ejercicios prácticos de entrega obligatoria (exige una entrega en el aula virtual)

8.BIBLIOGRAFÍA BÁSICA

Comentarios y anexos:

Bibliografía básica:

- Farzindar, A. and Inkpen, D. *Natural Language Processing for Social Media (Synthesis Lectures on Human Language Technologies)*. Morgan&Claypool Publishers, 2013.
- Zafarani, Reza, Mohammad Ali Abbasi, and Huan Liu. *Social media mining: an introduction*. Cambridge University Press, 2014.

La bibliografía básica no incluye algunos contenidos del curso. Por ello, en el entorno virtual de la asignatura se pondrá a disposición de los alumnos material de estudio complementario (artículos, recopilaciones y referencias a otro material disponible en la web).

9.BIBLIOGRAFÍA COMPLEMENTARIA

Comentarios y anexos:

Las siguientes referencias tratan temas relacionados con la asignatura:

- Daniel Jurafsky and James H. Martin. Speech and Language Processing. Pearson Prentice Hall. 2008. Versión preprint accesible online: Draft of September 28, 1999
- Christopher D. Manning and Hinrich Schütze. Foundations of Statistical Natural Language Processing. The MIT Press. 1999.
- Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. Introduction to Information Retrieval. Cambridge University Press. 2008.
- (DITPM) Data-Intensive Text Processing with MapReduce. Jimmy Lin and Chris Dyer, Morgan & Claypool Publishers. 2010.

Como bibliografía complementaria se recomendarán también otras referencias desde el curso virtual.

10.RECURSOS DE APOYO AL ESTUDIO

La plataforma de e-Learning Alf proporcionará el adecuado interfaz de interacción entre el alumno y sus profesores. Alf es una plataforma de e-Learning y colaboración que permite impartir y recibir formación, gestionar y compartir documentos, crear y participar en comunidades temáticas, así como realizar proyectos online.

Se ofrecerán las herramientas necesarias para que, tanto el equipo docente como el alumnado, encuentren la manera de compaginar tanto el trabajo individual como el aprendizaje cooperativo.

11.TUTORIZACIÓN Y SEGUIMIENTO

La tutorización de los alumnos se llevará a cabo a través de la plataforma de e-Learning Alf, por teléfono y por correo electrónico:

- Víctor Fresno (coordinador)

email: vfresno@lsi.uned.es

Tfno: 913988217

- Raquel Martínez

email: raquel@lsi.uned.es

Tfno: 913988725 (horario de atención al estudiante: Jueves de 11:30 a 13:30, y de 14:30 a 16:30 horas.)

El horario de tutoría será: Jueves de 15:00 a 19:00h.

12.EVALUACIÓN DE LOS APRENDIZAJES

La evaluación final de la asignatura se realizará como:

1. Evaluación continua a través del seguimiento del alumno.
2. Evaluación continua a través de la realización de trabajos.
3. Evaluación global del proceso de aprendizaje y adquisición de competencias y conocimientos.
4. Calificación numérica de 1 a 10 según legislación vigente (RD 1125/2003).

Habrà un conjunto de tareas obligatorias cuya entrega en plazo es un requisito imprescindible para aprobar la asignatura. La realización correcta de todas las tareas obligatorias asegura una nota de aprobado.

Aquellos alumnos que deseen una calificación mayor pueden elegir uno de entre una serie de trabajos optativos que se propongan.

13.COLABORADORES DOCENTES

- ARKAITZ ZUBIAGA MENDIALDUA
- MARÍA DEL SOTO MONTALVO HERRANZ