21-22

MÁSTER UNIVERSITARIO EN LENGUAJES Y SISTEMAS INFORMÁTICOS

GUÍA DE ESTUDIO PÚBLICA

MINERÍA INFORMACIÓN SOCIAL

CÓDIGO 31101324

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

21-22

MINERÍA INFORMACIÓN SOCIAL

CÓDIGO 31101324

ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

EQUIPO DOCENTE

HORARIO DE ATENCIÓN AL ESTUDIANTE

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

RESULTADOS DE APRENDIZAJE

CONTENIDOS

METODOLOGÍA

SISTEMA DE EVALUACIÓN

BIBLIOGRAFÍA BÁSICA

BIBLIOGRAFÍA COMPLEMENTARIA

RECURSOS DE APOYO Y WEBGRAFÍA

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/

DE4F445A8877C9A5BFD360FE9D825F09

Nombre de la asignatura	MINERÍA INFORMACIÓN SOCIAL
Código	31101324
Curso académico	2021/2022
Título en que se imparte	MÁSTER UNIVERSITARIO EN LENGUAJES Y SISTEMAS INFORMÁTICOS
Tipo	CONTENIDOS
Nº ETCS	6
Horas	150.0
Periodo	ANUAL
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

La asignatura "Minería de Información Social" se enmarca dentro del Máster en tecnologías del Lenguaje impartido por la Escuela Técnica Superior de Ingeniería Informática de la UNED.

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

Víctor Fresno forma parte del grupo NLP&IR de la UNED. Sus líneas de investigación se centran fundamentalmente en el estudio y propuesta de modelos de representación de textos para su procesamiento automático y su aplicación a problemas de Clasificación Automática, Agrupamiento y Recuperación de Información. Realizó una estancia de investigación post-doctoral como Visiting Faculty en el grupo BLENDER del Queens College, perteneciente a la City University of New York (CUNY).

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

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

Universidad Rey Juan Carlos y la UNED.
e.mail: raquel@lsi.uned.es

Colaboradores externos

MONTALVO HERRANZ, Mª DEL SOTO:

Soto Montalvo es profesora en la Universidad Rey Juan Carlos y forma parte del grupo NLP&IR de la UNED y del Grupo de investigación de alto rendimiento en desarrollo de software de la URJC. Su actividad investigadora se centra en diferentes campos: clustering de documentos multilingües y minería de textos aplicada a diferentes problemas, actualmente más orientada al dominio biomédico.
e.mail: soto.montalvo@urjc.es

ZUBIAGA MENDIALDUA, ARKAITZ:

Arkaitz Zubiaga es profesor en la Queen Mary University of London (Londres, Reino Unido), donde lidera la Social Data Science lab. Sus intereses de investigación se centran en torno a la minería de datos sociales y el procesamiento del lenguaje natural. Su trabajo reciente ha afrontado problemas como la detección y resolución de rumores, la detección de discurso de odio en medios sociales y el análisis de opiniones online.
Tras obtener su doctorado en la UNED en 2011, trabajó como investigador postdoctoral en la City University of New York (Estados Unidos) y posteriormente en la University of Warwick (Reino Unido).
e.mail: arkaitz@zubiaga.org

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

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 curse en paralelo las asignaturas de Procesamiento del Lenguaje Natural y Descubrimiento de Información en Textos.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

EQUIPO DOCENTE

Nombre y Apellidos	VICTOR DIEGO FRESNO FERNANDEZ (Coordinador de asignatura)
Correo Electrónico	vfresno@lsi.uned.es
Teléfono	91398-8217
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

Nombre y Apellidos	RAQUEL MARTINEZ UNANUE
Correo Electrónico	raquel@lsi.uned.es
Teléfono	
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

COLABORADORES DOCENTES EXTERNOS

Nombre y Apellidos	ARKAITZ ZUBIAGA MENDIALDUA
Correo Electrónico	azubiaga@lsi.uned.es

Nombre y Apellidos	SOTO MONTALVO HERRANZ
Correo Electrónico	soto.montalvo@lsi.uned.es

HORARIO DE ATENCIÓN AL ESTUDIANTE

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 (coodinador)

email: vfresno@lsi.uned.es

Tfno: 913988217

Horario de atención al estudiante: Martes y Miércoles de 11:30 a 13:30.

•Raquel Martínez

email: raquel@lsi.uned.es

Tfno: 913988725

Horario de atención al estudiante: Martes de 09:30 a 13.30.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

Competencias Básicas:

CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

Competencias Generales:

CPG1 - Adquirir capacidad de abstracción, análisis, síntesis y relación de ideas.

CPG2 - Adquirir capacidad crítica y de decisión

CPG3 - Adquirir capacidad de estudio y autoaprendizaje

CPG4 - Adquirir capacidad creativa y de investigación

CPG5 - Adquirir habilidades sociales para el trabajo en equipo

Competencias Específicas:

CE1 - Adquirir 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

CE3 - Adquirir capacidad de estudio de los sistemas y aproximaciones existentes y para distinguir las aproximaciones más efectivas.

CE4 - Adquirir capacidad para detectar carencias en el estado actual de la ciencia y la tecnología

CE5 - Adquirir capacidad para proponer nuevas aproximaciones que den solución a las carencias detectadas.

CE6 - Adquirir capacidad de especificar, diseñar, implementar y evaluar tanto cualitativa como cuantitativamente los modelos y sistemas propuestos.

CE7 - 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.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

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.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

•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.

CONTENIDOS

1. Introducción al Social Media: definiciones, terminología y conceptos

Social Media: definiciones, terminología y conceptos

2. Generación y anotación de corpus de Social Media

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

3. Análisis de contenido en Social Media

3.1 - Aplicación de técnicas de PLN
3.2 - Procesamiento semántico
3.3 - Análisis de tendencias
3.4 - Clasificación y clustering textual

4. Análisis de información de usuario en Social Media

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

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

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)

SISTEMA DE EVALUACIÓN

TIPO DE PRIMERA PRUEBA PRESENCIAL

Tipo de examen: No hay prueba presencial

TIPO DE SEGUNDA PRUEBA PRESENCIAL

Tipo de examen2: No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad: No

Descripción

No hay prueba presencial y las prácticas no requieren presencialidad.

Criterios de evaluación

Ponderación de la prueba presencial y/o los trabajos en la nota final

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

Fecha aproximada de entrega
Comentarios y observaciones

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC? Si,PEC no presencial

Descripción

En esta asignatura no se realiza una prueba presencial, la evaluación se realiza mediante evaluación continua a partir de las siguientes pruebas:
Las prácticas obligatorias a lo largo del curso.
La práctica opcional para subir nota, una vez que se han realizado las prácticas obligatorias.
Aquellos alumnos que deseen obtener una mayor calificación en la convocatoria de junio, podrán elegir uno de entre los trabajos optativos que se irán proponiendo. En estos casos la calificación final dependerá de la calidad del trabajo realizado.
Las tareas obligatorias se deberán entregar en los plazos que se vayan indicando. La no entrega de las tareas en el plazo previsto supondrá suspender la asignatura en la convocatoria de junio. El trabajo optativo para subir nota también tendrá una fecha límite de entrega. Habrá otro plazo de entrega de tareas para la convocatoria de septiembre.

Criterios de evaluación

Todos los temas del programa de la asignatura a partir del Tema 2 tienen asociada una práctica obligatoria cuya entrega es un requisito imprescindible para aprobar la asignatura. La realización correcta de todas las prácticas obligatorias asegura una nota de APROBADO, que podría llegar hasta NOTABLE (8) dependiendo de la calidad de las soluciones en su conjunto.
Aquellos estudiantes que deseen obtener una mayor calificación podrán elegir uno de entre los trabajos optativos que se proponen por parte del equipo docente. En estos casos la calificación final dependerá de la calidad del trabajo realizado.

Ponderación de la PEC en la nota final El promedio de las calificaciones obtenidas en las prácticas obligatorias y, en su caso, en la práctica opcional constituye la nota final de la asignatura.

Fecha aproximada de entrega
Comentarios y observaciones

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

Las prácticas obligatorias tienen un plazo de entrega fijo, que suele ser de unas tres semanas después de haber finalizado el tema correspondiente, de acuerdo con la temporización de la asignatura y los periodos vacacionales. Esta temporización permite al estudiante suficiente margen de tiempo para poder organizar su trabajo de acuerdo con sus circunstancias personales.

Los estudiantes que no entreguen las tareas en el plazo establecido para la convocatoria de junio tendrán otro plazo de entrega en la convocatoria de septiembre.

La práctica o tarea opcional también tiene un plazo de entrega acorde con la temporización de la asignatura.

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? No

Descripción

Criterios de evaluación

Ponderación en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

¿CÓMO SE OBTIENE LA NOTA FINAL?

El promedio de las calificaciones obtenidas en las prácticas obligatorias (hasta un máximo de 8 sobre 10) al que se podrá sumar hasta 2 puntos si se ha realizado la práctica opcional y en función de la calidad de ésta.

BIBLIOGRAFÍA BÁSICA

Bibliografía básica:

- Farzindar, A. and Inkpen, D. *Natural Language Processing for Social Media (Synthesis Lectures on Human Language Technologies)*. 2nd Edition. Morgan&Claypool Publishers, 2017.
- 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).

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09

BIBLIOGRAFÍA COMPLEMENTARIA

Las siguientes referencias tratan temas relacionados con la asignatura (se trata de bibliografía clásica en el área):

•Daniel Jurafsky and James H. Martin. *Speech and Language Processing*. (3rd. ed. draft) Pearson Prentice Hall. 2019.
Versión preprint accesible online: Draft chapters in progress, October 16, 2019

•Christopher D. Manning and Hinrich Schütze. *Foundations of Statistical Natural Language Processing*. The MIT Press. 1999..
Companion website for the following book

•Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schutze. *Introduction to Information Retrieval*. Cambridge University Press. 2008.
Online edition (c) 2009 Cambridge UP (Draft of April, 2009)

•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.

RECURSOS DE APOYO Y WEBGRAFÍA

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.

IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
DE4F445A8877C9A5BFD360FE9D825F09