

Máster en Ingeniería Informática (Plan 2018)

DATOS DE LA ASIGNATURA

Nombre:

Minería de Opinión en Redes Sociales

Denominación en inglés:

Opinion Mining in Social Networks

Código:

1180419

Carácter:

Optativo

Horas:

	Totales	Presenciales	No presenciales
Trabajo estimado:	75	30	45

Créditos:

Grupos reducidos				
Grupos grandes	Aula estándar	Laboratorio	Prácticas de campo	Aula de informática
1.5	0	0	0	1.5

Departamentos:

Tecnologías de la Información

Áreas de Conocimiento:

Lenguajes y Sistemas Informáticos

Curso:

1º - Primero

Cuatrimestre:

Segundo cuatrimestre

DATOS DE LOS PROFESORES

Nombre:	E-Mail:	Teléfono:	Despacho:
De la Villa Cordero, Manuel	manuel.villa@dti.uhu.es	959217639	Edif. de la ETSI, Planta 1, Dcho. 131
*Maña López, Manuel J.	manuel.mana@dti.uhu.es	959217389	ETSI, 1ª planta, Campus de El Carmen

*Profesor coordinador de la asignatura

Consultar los horarios de la asignatura

1. Descripción de contenidos

1.1. Breve descripción (en castellano):

Las redes sociales y, en general, la Web acumulan un gran volumen de información que puede ser de gran relevancia para las organizaciones. La mayor parte de esta información se encuentra en forma de texto. A las dificultades habituales del tratamiento de cualquier texto, como por ejemplo la ambigüedad presente en todos los niveles del lenguaje, se suman características que dificultan mucho más la tarea. Entre estas dificultades adicionales podemos destacar la carencia de una estructura de discurso, el uso de jergas, abreviaturas, emoticonos y etiquetas, la utilización de frases agramaticales y la se está diciendo en cada momento en Internet sobre una persona, una marca, una institución o un producto, y detectar más relevante analizar todo lo que se está diciendo en cada momento en Internet sobre una persona, una marca, una institución o un producto, y detectar automáticamente las entidades, conceptos, polaridad de opiniones o reputación.

- El objetivo de la asignatura es dotar al alumno de las técnicas necesarias para organizar, resumir y analizar grandes volúmenes de información no estructurada en forma de texto con especial atención a las redes sociales y la Web.

La asignatura estará centrada en la comprensión, el análisis y la evaluación de los siguientes contenidos:

- Procesamiento del lenguaje natural (PLN). Análisis léxico y morfológico. Análisis sintáctico. Reconocimiento de entidades con nombre. Detección de la negación y la especulación. Recursos léxico-semánticos. PLN en la Web y las redes sociales.

Análisis de la información textual. Búsqueda en la Web. Resumen automático. Agrupamiento. Categorización. Minería de opiniones. Análisis de sentimiento. Integración en sistemas empresariales para toma de decisiones. Analítica web y social media.

1.2. Breve descripción (en inglés):

Natural Language Processing (NLP). Lexical and morphological analysis. Parsing. Named entity recognition. Negation and speculation detection. Lexical-semantic resources. NLP on the Web and social networks. Opinon Mining. Sentiment analysis. Web search. Web analytics and social media. Integration in business systems for decisions support. Automatic summarization. Clustering. Categorization.

2. Situación de la asignatura

2.1. Contexto dentro de la titulación:

Las redes sociales y, en general, la Web acumulan un gran volumen de información que puede ser de gran relevancia para las organizaciones. La mayor parte de esta información se encuentra en forma de texto. A las dificultades habituales del tratamiento de cualquier texto, como por ejemplo la ambigüedad presente en todos los niveles del lenguaje, se suman características que dificultan mucho más la tarea. Entre estas dificultades adicionales podemos destacar la carencia de una estructura de discurso, el uso de jergas, abreviaturas, emoticonos y etiquetas, la utilización de frases agramaticales y la carencia de contexto, sobre todo en el caso de los microblogs, que incrementa la ambigüedad. Sin embargo, cada vez es más relevante analizar todo lo que se está diciendo en cada momento en Internet sobre una persona, una marca, una institución o un producto, y detectar automáticamente las entidades, conceptos, polaridad de opiniones o reputación. A pesar de la relevancia que tiene el texto y del auge de las redes sociales y la Web, la minería de opiniones no es una materia de la que se ocupe ninguna otra asignatura de este Máster, con la profundidad que requiere el tema. Tan solo la asignatura Computación Inteligente (asignatura obligatoria, primer cuatrimestre) incluye un tema dedicado al procesamiento del lenguaje natural. De esta forma, esta asignatura opcional viene a cubrir esa carencia.

2.2. Recomendaciones:

Las competencias básicas que deberían poseer los alumnos antes de comenzar la asignatura son:

- Tener destreza en el lenguaje de programación Java.
- Poseer destreza para buscar información útil en la Red.
- Poseer conocimientos suficientes para leer documentos escritos en inglés.
- Saber manejar fuentes bibliográficas.
- Tener capacidad de lectura comprensiva.
- Saber expresarse correctamente de forma oral y escrita.

3. Objetivos (Expresados como resultados del aprendizaje):

Con la realización de esta asignatura, el estudiante será capaz de: (a) Conocer y aplicar las técnicas básicas para el procesamiento del texto y, en particular, cuando se trata de texto procedente de redes sociales, y (b) Diseñar sistemas que permitan acceder a la información textual y analizar y extraer conocimiento y polaridad de la misma, especialmente cuando se trata de grandes volúmenes de texto.

Competencias Específicas:

- Conocimiento y capacidad para aplicar las técnicas básicas para el procesamiento del texto y, en particular, cuando se trata de texto procedente de redes sociales.
- Capacidad para el diseño de sistemas que permitan acceder a la información textual y analizar y extraer conocimiento de la misma, especialmente cuando se trata de grandes volúmenes de texto.

4. Competencias a adquirir por los estudiantes

4.1. Competencias específicas:

4.2. Competencias básicas, generales o transversales:

- **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
- **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
- **CG4:** Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería en Informática.
- **CG8:** Capacidad para la aplicación de los conocimientos adquiridos y de resolver problemas en entornos nuevos o poco conocidos dentro de contextos más amplios y multidisciplinarios, siendo capaces de integrar estos conocimientos
- **CG9:** Capacidad para comprender y aplicar la responsabilidad ética, la legislación y la deontología profesional de la actividad de la profesión de Ingeniero en Informática
- **CT1:** Gestionar adecuadamente la información adquirida expresando conocimientos avanzados y demostrando, en un contexto de investigación científica y tecnológica o altamente especializado, una comprensión detallada y fundamentada de los aspectos teóricos y prácticos y de la metodología de trabajo en el campo de estudio.
- **CT3:** Desarrollar una actitud y una aptitud de búsqueda permanente de la excelencia en el quehacer académico y en el ejercicio profesional futuro.
- **CT4:** Comprometerse con la ética y la responsabilidad social como ciudadano y como profesional, con objeto de saber actuar conforme a los principios de respeto a los derechos fundamentales y de igualdad entre hombres y mujeres y respeto y promoción de los Derechos Humanos, así como los de accesibilidad universal de las personas discapacitadas, de acuerdo con los principios de una cultura de paz, valores democráticos y sensibilización medioambiental.

5. Actividades Formativas y Metodologías Docentes

5.1. Actividades formativas:

- Sesiones de Teoría sobre los contenidos del Programa.
- Sesiones de Resolución de Problemas.
- Sesiones Prácticas en Laboratorios Especializados o en Aulas de Informática.
- Actividades Académicamente Dirigidas por el Profesorado: seminarios, conferencias, desarrollo de trabajos, debates, tutorías colectivas, actividades de evaluación y autoevaluación.

5.2. Metodologías docentes:

- Clase Magistral Participativa.
- Desarrollo de Prácticas en Laboratorios Especializados o Aulas de Informática en grupos reducidos.
- Resolución de Problemas y Ejercicios Prácticos.
- Tutorías Individuales o Colectivas. Interacción directa profesorado-estudiantes.
- Evaluaciones y Exámenes.

5.3. Desarrollo y justificación:

Las clases teóricas permitirán presentar los conceptos teóricos de forma clara y concisa, ayudándose de las herramientas docentes adecuadas al alcance del profesor. Se proporcionará al alumno material para que pueda trabajar ciertos contenidos de forma personal, estimulando, así, el aprendizaje autónomo.

En las prácticas de laboratorio los alumnos tendrán que desarrollar ejercicios prácticos que permitirán afianzar los conocimientos adquiridos en teoría.

Se realizarán actividades académicamente dirigidas con el objetivo de profundizar en aspectos más avanzados y concretos de algunos contenidos teóricos y prácticos.

Lectura de los contenidos de los temas. Para poder seguir las sesiones presenciales, el alumno habrá de haber realizado previamente un trabajo de lectura comprensiva del material proporcionado en cada tema.

Entrega de ejercicios/prácticas/trabajos evaluables. Para formalizar el trabajo autónomo e individual, se propondrán tareas evaluables en la plataforma que el alumno desarrollará de manera no presencial.

Actividades de autoevaluación. Al final de algunos temas existirán actividades de autoevaluación que permitan al alumno medir el grado de asimilación de los contenidos.

Tutorías colectivas a través de plataformas de enseñanza virtual (foros, wikis, chats). En la plataforma docente se habilitarán espacios para el debate y la discusión grupal de contenidos de la asignatura.

Actividades no presenciales con evaluación por pares. Para reforzamiento del aprendizaje, en algunas actividades los alumnos realizarán evaluación cruzada, cambiando de rol en el proceso de aprendizaje.

Desarrollo cooperativo de trabajos utilizando herramientas de discusión asíncrona (foros, wikis, ...). Se fomentará el trabajo en grupo mediante la definición de tareas en la plataforma de apoyo a la docencia no presencial.

6. Temario desarrollado:

Bloque I. Procesamiento del Lenguaje Natural.

Tema 1. Introducción. Motivación. Conceptos básicos. Características del texto.

Tema 2. Análisis del texto. Análisis léxico. Eliminación de palabras vacías. Reducción a la raíz. Análisis morfológico. Análisis sintáctico. Reconocimiento de entidades nombradas. Índice invertido. Detección de la negación y la especulación. Recursos léxico-semánticos. PLN en la Web y las Redes Sociales.

Bloque II. Minería de opiniones en Redes Sociales.

Tema 3. Minería de opiniones. Definición del problema. Extracción de opiniones. Clasificación de la polaridad. Resumen de la opinión. Extracción de aspectos. Análisis de sentimiento.

Tema 4. Otras tareas de análisis textual en la Web y Redes Sociales. Búsqueda en la Web. Analítica web y social media. SEO y SEM. Ayuda para la toma de decisiones. Resumen automático. Agrupamiento. Categorización.

NOTA: Se desea puntualizar que de acuerdo con la metodología semipresencial, sólo la mitad de la docencia y sus contenidos se impartirán presencialmente, usándose para la otra mitad técnicas no presenciales apoyadas en la plataforma docente online.

7. Bibliografía

7.1. Bibliografía básica:

Bing Liu. Sentiment Analysis: mining sentiments, opinions, and emotions. Cambridge University Press. 2015.

Daniel Jurafsky y James H. Martin. Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Second Edition. Pearson Education International. 2009.

R. Baeza-Yates y B. Ribeiro-Neto. Modern Information Retrieval: The Concepts and Technology behind Search. Second Edition. Pearson, 2011.

Christopher D. Manning, Prabhakar Raghavan y Hinrich Schütze. Introduction to Information Retrieval. Cambridge University Press, 2008.

7.2. Bibliografía complementaria:

Christopher D. Manning y Hinrich Schütze. Foundations of statistical natural language processing. MIT Press. 1999.

David A. Grossman y Ophir Frieder. Information Retrieval: Algorithms and Heuristics, 2nd edition. Springer, 2004.

8. Sistemas y criterios de evaluación.

8.1. Sistemas de evaluación:

- Examen de teoría/problemas
- Defensa de Prácticas

8.2. Criterios de evaluación y calificación:

Los principios de evaluación de la asignatura siguen unos criterios de evaluación preferentemente continua, entendiéndose por tal la evaluación diversificada que se lleva a cabo en distintos momentos del curso académico en curso. Por esta razón, la asistencia a clase es obligatoria (se considera satisfecho este requisito con una asistencia al 80% de las sesiones). Esta evaluación se realiza, para las **convocatorias ordinarias I y II** utilizando diferentes instrumentos y ponderaciones, seleccionados según la competencia que se desea evaluar:

La asignatura se califica utilizando diferentes instrumentos:

I1. Examen objetivo. En esta sección se evalúan mediante el examen teórico objetivo los fundamentos teóricos junto con la realización de casos y supuestos prácticos. Con esta actividad se evalúan las competencias CB6, CB7, CG4, CG8, CG9, CT1 y CT3. Se corresponde con el sistema de evaluación "Examen de teoría/problemas". 30% de la nota.

I2. Trabajos de laboratorio. Los alumnos realizarán un conjunto de prácticas guiadas, con el fin de asimilar conceptos vistos en las clases teórico-prácticas. Con esta actividad se evalúan las competencias CB6, CB7, CB10, CG4, CG8, CT1. CT3. Se corresponde con el sistema de evaluación "Defensa de prácticas". 40% de la nota.

I3. Pruebas de evaluación mediante plataformas de enseñanza virtual. Los alumnos podrán realizar test de verificación del grado de adquisición de conocimientos. Con esta actividad se evalúan las competencias CB6, CB10, CG4, CT1. CT4. 20% de la nota.

I4. Participación en las actividades propuestas. Con esta actividad se evalúan las competencias CB6, CB7, CG4, CG8, CG9, CT1, CT3 y CT4. 10% de la nota.

Para aprobar la asignatura es necesario aprobar por separado I1 e I2 y sumar 5 entre las 4 calificaciones.

CRITERIOS DE EVALUACIÓN Y CALIFICACIÓN EN CONVOCATORIAS I Y II

Para aprobar la asignatura en estas convocatorias es necesario aprobar por separado I1 e I2 y sumar 5 puntos entre las 4 calificaciones. Si se aprueba la asignatura, la calificación final en estas convocatorias se obtendrá sumando los resultados obtenidos en los cuatro instrumentos de evaluación.

La nota en acta del alumno se obtendrá del siguiente modo, teniendo en cuenta que las evaluaciones de los conocimientos teóricos (I1), prácticos (I2), pruebas de evaluación online y participación en actividades, con una valoración sobre 10:

Si $\text{Nota}_{I1} \geq 5$ y $\text{Nota}_{I2} \geq 5$ Entonces $\text{Nota}_{\text{Final}} = \text{Nota}_{I1} \times 0,3 + \text{Nota}_{I2} \times 0,4 + \text{Nota}_{I3} \times 0,2 + \text{Nota}_{I4} \times 0,1$

En caso contrario $\text{Nota}_{\text{Final}} = \text{Mínimo}(\text{Nota}_{I1}, \text{Nota}_{I2})$

De manera que estará aprobada la asignatura sólo si la $\text{Nota}_{\text{Final}}$ es ≥ 5 .

A los alumnos que no alcancen el aprobado en la asignatura, se les mantendrá la nota de la Parte_I1, Parte_I2 hasta la última convocatoria anual, siempre y cuando su puntuación en las mismas sea ≥ 5 y la nota de los trabajos.

Para la **convocatoria ordinaria III**, se utilizarán los siguientes instrumentos y ponderaciones:

La asignatura se califica utilizando diferentes instrumentos (la evaluación de competencias es la misma que en convoc. I y II):

I1. Examen objetivo. Se corresponde con el sistema de evaluación "Examen de teoría/problemas". 40% de la nota.

I2. Trabajos de laboratorio. Se corresponde con el sistema de evaluación "Defensa de prácticas". 40% de la nota.

I3. Pruebas de evaluación mediante plataformas de enseñanza virtual. 20% de la nota.

Para aprobar la asignatura es necesario aprobar por separado I1 e I2 y sumar 5 entre las 3 calificaciones.

CRITERIOS DE EVALUACIÓN Y CALIFICACIÓN EN CONVOCATORIA III

Para aprobar la asignatura en esta convocatoria es necesario aprobar por separado I1 e I2 y sumar 5 puntos entre las 2 calificaciones. Si se aprueba la asignatura, la calificación final en estas convocatorias se obtendrá sumando los resultados obtenidos en los dos instrumentos de evaluación.

La nota en acta del alumno se obtendrá del siguiente modo, teniendo en cuenta que las evaluaciones de los conocimientos teóricos I1 y prácticos I2 se realizarán sobre 10:

Si $\text{Nota}_{I1} \geq 5$ y $\text{Nota}_{I2} \geq 5$ Entonces $\text{Nota}_{\text{Final}} = \text{Nota}_{I1} \times 0,4 + \text{Nota}_{I2} \times 0,4 + \text{Nota}_{I3} \times 0,2$

En caso contrario $\text{Nota}_{\text{Final}} = \text{Mínimo}(\text{Nota}_{I1}, \text{Nota}_{I2})$

De manera que estará aprobada la asignatura sólo si la $\text{Nota}_{\text{Final}}$ es ≥ 5 .

A los alumnos que no alcancen el aprobado en la asignatura, se les mantendrá la nota hasta la última convocatoria anual, siempre y cuando su puntuación en las mismas sea ≥ 5 y la nota de los trabajos.

EVALUACIÓN ÚNICA FINAL

Aquellos estudiantes que así lo consideren pueden acogerse a la realización de una evaluación única final. En este caso deberá presentar una solicitud por correo electrónico (con acuse de recibo) al coordinador de la asignatura, que tendrá validez tras la respuesta confirmando la recepción por el profesor. La evaluación única final consistirá, para todas las convocatorias, en un solo acto académico que estará formado por las siguientes pruebas:

La asignatura se califica utilizando diferentes instrumentos:

I1. Examen objetivo. Se corresponde con el sistema de evaluación "Examen de teoría/problemas".

I2. Trabajos de laboratorio. Se corresponde con el sistema de evaluación "Defensa de prácticas".

I3. Pruebas de evaluación mediante plataformas de enseñanza virtual.

I4. Participación en las actividades propuestas.

Para aprobar la asignatura es necesario aprobar por separado I1 e I2 y sumar 5 entre las 4 calificaciones.

Prueba 1. 70% de la nota. Esta prueba cubre los sistemas de evaluación Examen de Teoría/Problemas (60%), Pruebas de evaluación mediante plataformas y participación en actividades propuestas (10%) y Seguimiento Individual del Estudiante (0%). Consistirá en un examen teórico con preguntas a desarrollar. No se podrá utilizar ningún material docente.

Prueba 2. 30% de la nota. Esta prueba cubre los sistemas de evaluación Defensa de Prácticas (30%). Consistirá en un examen práctico en aula de informática. No se podrá utilizar ningún material docente salvo el que proporcione el profesor.

La duración total de ambas pruebas será como máximo de 4 horas.

Para aprobar la asignatura mediante la evaluación única final será necesario aprobar por separado ambas pruebas y sumar 5 puntos entre las 2 calificaciones. Si se aprueba la asignatura, la calificación final en esta convocatoria se obtendrá sumando los resultados obtenidos en las dos pruebas.

CRITERIOS SOBRE MATRÍCULAS DE HONOR

En el caso de haber más candidatos que posibilidades de **matrículas de honor** por número de estudiantes en la asignatura, y con el objetivo de discriminar situaciones de equidad en la calificación final, se seguirán los siguientes criterios: nota del examen de teoría (I1 o PRUEBA 1, según el método de evaluación) y, en caso de persistir el empate, nota de las prácticas de laboratorio (I2 o PRUEBA 2, según el método de evaluación).

9. Organización docente semanal orientativa:

	Semanas	Grupos Grandes	Grupos Reducidos Aula Estándar	Grupos Reducidos Aula de Informática	Grupos Reducidos Laboratorio	Grupos Reducidos prácticas de campo	Pruebas y/o actividades evaluables	Contenido desarrollado
#1	2	0	0	0	0		Evaluación Plataforma Aprendizaje	Tema 1
#2	2	0	0	0	0			Tema 2
#3	2	0	0	0	0			Tema 2
#4	2	0	0	0	0		Evaluación Plataforma Aprendizaje	Tema 2
#5	0	0	2	0	0			Práctica Parte I: PLN
#6	0	0	2	0	0			Práctica Parte I: PLN
#7	0	0	2	0	0		Evaluación Práctica Parte I	Práctica Parte I: PLN
#8	2	0	0	0	0			Tema 3
#9	2	0	0	0	0		Evaluación Plataforma Aprendizaje	Tema 3
#10	2	0	0	0	0			Tema 4
#11	1	0	1	0	0		Evaluación Plataforma Aprendizaje	Tema 4. Práctica Parte II: PLN
#12	0	0	2	0	0			Práctica Parte II: PLN
#13	0	0	2	0	0			Práctica Parte II: PLN
#14	0	0	2	0	0			Práctica Parte II: PLN
#15	0	0	2	0	0		Evaluación Práctica Parte II	Práctica Parte II: PLN
	15	0	15	0	0			