

Guía docente de la asignatura

Fecha de aprobación por la Comisión Académica: 26/07/2023

Minería de Datos: Aprendizaje No Supervisado y Detección de Anomalías (M51/56/3/7)

Máster

Máster Universitario en Ciencia de Datos e Ingeniería de Computadores

MÓDULO

Módulo de Nivelación de Conocimientos

RAMA

Ingeniería y Arquitectura

CENTRO RESPONSABLE DEL TÍTULO

Escuela Internacional de Posgrado

Semestre

Primero

Créditos

4

Tipo

Optativa

Tipo de enseñanza

Presencial

BREVE DESCRIPCIÓN DE CONTENIDOS (Según memoria de verificación del Máster)

1. Clustering:
 1. Métodos basados en centróides.
 2. Métodos jerárquicos.
 3. Métodos de evaluación: Interna y externa.
2. Reglas de Asociación:
 1. Métodos de generación.
 2. Conjuntos maximales y cerrados.
 3. Tratamiento de valores numéricos. Reglas Multinivel.
 4. Evaluación: Medidas de interés.
3. Anomalías:
 1. Métodos Supervisados.
 2. Métodos Semi-supervisados.
 3. Métodos No supervisados.

COMPETENCIAS

COMPETENCIAS BÁSICAS

- CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser



originales en 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

- CG01 - Capacidad de acceso y gestión de la información
- CG02 - Capacidad de análisis y síntesis
- CG03 - Capacidad de organización y planificación
- CG04 - Capacidad emprendedora
- CG05 - Capacidad para tomar decisiones de forma autónoma
- CG06 - Capacidad de uso de una lengua extranjera
- CG07 - Motivación por la calidad
- CG08 - Capacidad para trabajar en equipo

COMPETENCIAS ESPECÍFICAS

- CE03 - Capacidad para la aplicación de técnicas y metodologías que permitan abordar desde nuevas perspectivas los problemas de interés, gracias a la disponibilidad de las plataformas de computación y comunicación con altos niveles de prestaciones.
- CE04 - Capacidad de análisis de aplicaciones en ámbitos de biomedicina y bioinformática, optimización y predicción, control avanzado, y robótica bioinspirada, tanto desde el punto de vista de los requisitos para una implementación eficaz de los algoritmos y las técnicas de computación que se usan para abordarlas, como de las características deseables en las arquitecturas donde se ejecutan
- CE05 - Capacidad para modelar y resolver problemas reales o académicos mediante técnicas de ciencia de datos
- CE07 - Capacidad de utilización de herramientas avanzadas para modelar y resolver problemas en los que se dispone de un tamaño enorme de datos.
- CE08 - Conocer algunas de las principales aplicaciones de la ciencia de datos y de las tecnologías inteligentes

COMPETENCIAS TRANSVERSALES

- CT01 - Ser consciente de la importancia del desarrollo sostenible y demostrar sensibilidad medioambiental.
- CT02 - Ser consciente del derecho a la no discriminación y al acceso universal al conocimiento de las personas con discapacidad.



RESULTADOS DE APRENDIZAJE (Objetivos)

El alumno aprenderá los algoritmos más importantes para realizar agrupamientos no supervisados de conjuntos de datos utilizando los métodos basados en centroides y los basados en agrupaciones sucesivas (o jerárquicos). Aprenderá a evaluar la calidad de los agrupamientos obtenidos y aplicará los conocimientos obtenidos a casos prácticos en las aulas de laboratorio.

El alumno aprenderá los algoritmos principales para obtener patrones frecuentes y reglas de asociación, conocerá los problemas inherentes asociados al tipo de conocimiento extraído y los resolverá aplicando las distintas medidas de interés, así como utilizando patrones maximales y cerrados para condensar la información obtenida. Aprenderá a poner en práctico estos conocimientos en las aulas de laboratorio.

PROGRAMA DE CONTENIDOS TEÓRICOS Y PRÁCTICOS

TEÓRICO

1. - Clustering:
 1. Métodos basados en centróides.
 2. Métodos jerárquicos.
 3. Métodos de evaluación: Interna y externa.
2. - Reglas de Asociación:
 1. Métodos de generación.
 2. Conjuntos maximales y cerrados.
 3. Tratamiento de valores numéricos. Reglas Multinivel.
 4. Evaluación: Medidas de interés.
3. - Anomalías:
 1. Métodos Supervisados.
 2. Métodos Semi-Supervisados.
 3. Métodos No Supervisados.

PRÁCTICO

- Práctica 1: Clustering.
- Práctica 2: Reglas de Asociación.
- Práctica 3: Anomalías.

BIBLIOGRAFÍA

BIBLIOGRAFÍA FUNDAMENTAL

- I. H. Witten, E. Frank Data Mining Practical Machine Learning Tools and Techniques (2 Edition). Elsevier (2005)
- P.-N. Tan, M. Steinbach, V. Kumar. Introduction to Data Mining. Addison Wesley (2006)
- M.R. Berthold, Ch. Borgelt, F. Höppner, F. Klawonn. Guide to Intelligent Data Analysis. Springer Verlag (2010)
- I. H. Witten, E. Frank, M. A. Hall, C. J. Pal. Data mining: practical machine learning tools and techniques (4ª Edition). Morgan Kaufmann (2016)
- E. Alpaydin. Introduction to Machine Learning (4ª Edición). The MIT Press (2020)



BIBLIOGRAFÍA COMPLEMENTARIA

- F. Provost, T. Fawcett. Data Science for Business. O'Reilly Media (2013)
- S. Shalev-Shwartz, S. Ben-David. Understanding machine learning: from theory to algorithms. Cambridge University Press (2014)
- G. Grolemund, H. Wickham. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media (2016)

ENLACES RECOMENDADOS

- Sitio web del Máster Universitario Oficial en Ciencia de Datos e Ingeniería de Computadores: <http://masteres.ugr.es/datcom/>
- Departamento de Ciencias de la Computación e I.A.: <https://decsai.ugr.es/>
- R: <http://www.r-project.org/>
- RStudio: <https://posit.co/download/rstudio-desktop/>
- Python: <https://www.anaconda.com/download>
- KDNuggets: <http://www.kdnuggets.com>

METODOLOGÍA DOCENTE

- MD01 Lección magistral/expositiva
- MD02 Resolución de problemas y estudio de casos prácticos
- MD03 Prácticas de laboratorio
- MD04 Seminarios
- MD05 Análisis de fuentes y documentos
- MD06 Realización de trabajos en grupo
- MD07 Realización de trabajos individuales

EVALUACIÓN (instrumentos de evaluación, criterios de evaluación y porcentaje sobre la calificación final)

EVALUACIÓN ORDINARIA

Todo lo relativo a la evaluación se regirá por la normativa de evaluación y de calificación de los estudiantes de la Universidad de Granada. El sistema de calificaciones se expresará mediante calificación numérica de acuerdo con lo establecido en el art. 5 del R. D 1125/2003, de 5 de septiembre, por el que se establece el sistema europeo de créditos y el sistema de calificaciones en las titulaciones universitarias de carácter oficial y validez en el territorio nacional.

La metodología de evaluación se compone de los siguientes elementos:

- SE1. Evaluación de la Parte Teórica: exámenes finales o parciales, sesiones de evaluación y entregas de ejercicios sobre el desarrollo y los resultados de las actividades propuestas. Ponderación 60%
- SE2. Evaluación de la Parte Práctica: se realizarán prácticas de laboratorio, resolución de problemas y desarrollo de proyectos (individuales o en grupo), y se valorarán las entregas de los informes/memorias realizados por los alumnos, o en su caso las entrevistas personales con los alumnos y las sesiones de evaluación. Ponderación 40%



Más en concreto:

- Examen escrito global de la asignatura. Se propone este examen debido a que se trata de una asignatura de nivelación y debe garantizarse que el alumno adquiere todos los conceptos básicos que serán necesarios en otras materias.
- Un trabajo práctico a desarrollar por el alumno de aquella parte de la asignatura (clustering, anomalías, asociación) que el alumno elija. Dicho trabajo consistirá en desarrollar el guión de prácticas de la parte correspondiente con un dataset distinto al utilizado en clase, a elegir por el alumno. **El trabajo debe incluir un análisis descriptivo de los resultados obtenidos** y el profesor podrá solicitar la defensa del trabajo presentado.
- La ponderación será: 0,6 examen+0,4 trabajo práctico. Será necesario obtener una puntuación mayor a 3,5 sobre 10 en cada una de las partes para hacer la media. En caso de que no se llegue al mínimo en alguna de las partes, se le asignara al estudiante como calificación final el mínimo entre 4,9 y la nota obtenida en la asignatura.

EVALUACIÓN EXTRAORDINARIA

El artículo 19 de la Normativa de Evaluación y Calificación de los Estudiantes de la Universidad de Granada establece que los estudiantes que no hayan superado la asignatura en la convocatoria ordinaria dispondrán de una convocatoria extraordinaria. A ella podrán concurrir todos los estudiantes, con independencia de haber seguido o no un proceso de evaluación continua. De esta forma, el estudiante que no haya realizado la evaluación continua tendrá la posibilidad de obtener el 100% de la calificación mediante la realización de un examen de teoría y la entrega de un trabajo de prácticas.

La evaluación sigue el mismo esquema que el examen de la convocatoria ordinaria, salvo que, si el estudiante ya se hubiese presentado al examen escrito o entregado la práctica, se le mantiene la nota de la parte que el alumno decida (siempre que no se presente a esa parte en esta convocatoria extraordinaria).

EVALUACIÓN ÚNICA FINAL

El artículo 8 de la Normativa de Evaluación y Calificación de los Estudiantes de la Universidad de Granada establece que podrán acogerse a la evaluación única final, el estudiante que no pueda cumplir con el método de evaluación continua por causas justificadas. Para acogerse a la evaluación única final, el estudiante deberá solicitarlo al Coordinador del Máster dentro del plazo establecido.

En la evaluación única final, el 60% de la nota será un examen de teoría y el otro 40% de la nota se contabiliza mediante un examen práctico en las aulas de ordenadores, en el que el alumno deberá resolver varias cuestiones relacionadas con los guiones de prácticas de las tres partes de la asignatura. La ponderación será: 0,6 examen teoría+0,4 examen práctico. Será necesario obtener una puntuación mayor a 3,5 sobre 10 en cada una de las partes para hacer la media. En caso de que no se llegue al mínimo en alguna de las partes, se le asignara al estudiante como calificación final el mínimo entre 4,9 y la nota obtenida en la asignatura.

INFORMACIÓN ADICIONAL

Información de interés para estudiantado con discapacidad y/o Necesidades Específicas de Apoyo Educativo (NEAE): [Gestión de servicios y apoyos](https://ve.ugr.es/servicios/atencion-) (<https://ve.ugr.es/servicios/atencion->





social/estudiantes-con-discapacidad).

