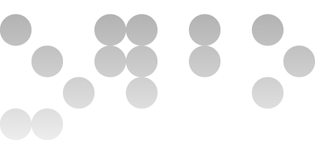
75.583 · Aprendizaje Computacional · PEC2 · 2023-24 · Ingeniería Informática · Estudios de Informática Multimedia y Telecomunicaciones



**PEC 2**

**Presentación**

Segunda actividad de evaluación continua del curso. En esta PEC se pretende conocer y desarrollar distintas técnicas de machine learning.

**Competencias**

**Competencias de grado**

- Capacidad de utilizar los fundamentos matemáticos, estadísticos y físicos y comprender los sistemas TIC.

- Capacidad para analizar un problema en el nivel de abstracción adecuado a cada situación y aplicar las habilidades y conocimientos adquiridos para abordarlo y resolverlo.

- Capacidad para conocer las tecnologías de comunicaciones actuales y emergentes y saberlas aplicar convenientemente, para diseñar y desarrollar soluciones basadas en sistemas y tecnologías de la información.

- Capacidad para proponer y evaluar diferentes alternativas tecnológicas y resolver un problema concreto.

**Competencias específicas**

- Capacidad para utilizar la tecnología de aprendizaje automático más adecuada para resolver un determinado problema.

- Capacidad para evaluar el rendimiento de los diferentes algoritmos de resolución de problemas mediante técnicas de validación cruzada.

**Objetivos**

El objetivo de esta prueba de evaluación es clasificar los objetos astronómicos del conjunto de datos proporcionado en función de su clase y evaluar el rendimiento de los modelos utilizados.

**Descripción de la PEC**

Se proporciona un conjunto de datos en formato CSV que contiene información sobre objetos astronómicos observados por el [Sloan Digital Sky Survey (SDSS)](https://www.sdss.org/). Las características incluyen medidas fotométricas en diferentes filtros, coordenadas astronómicas, información de observación, etc. La variable objetivo es la clase de objeto, que puede ser galaxia, estrella u objeto quasar

**Ejercicio 1**

a) Realiza el preprocesamiento de los datos si es necesario, justificando cada paso y mostrando el dataset tratado.

b) Utiliza un algoritmo de clasificación de tu elección para categorizar los objetos astronómicos en las tres clases mencionadas. Calcula y presenta la exactitud del modelo.

c). ¿Qué métricas son necesarias para poder tener una idea precisa del funcionamiento del modelo de clasificación?

**Ejercicio 2**

a) Construye un dendrograma utilizando un método de agrupamiento jerárquico. Muestra los pasos involucrados, incluyendo el dendrograma resultante y explica las decisiones tomadas.

b) Calcula la exactitud del dendrograma propuesto.

**Ejercicio 3**

a) Realiza el mismo tratamiento de datos que en el Ejercicio 1, utilizando herramientas que consideres adecuadas. Muestra los primeros 10 ejemplos ya tratados.

b) Utiliza un algoritmo de clasificación de vecinos más cercanos (K-Nearest Neighbors) sobre los datos tratados. Calcula y presenta la exactitud, precisión, recall, fall-out y matriz de confusión.

c) Repite el paso anterior utilizando un clasificador de árbol de decisión.

**Ejercicio 4**

a) ¿Dirías que el conjunto de datos se puede considerar balanceado?

b) ¿Cómo crees que afecta el desbalanceo a la Exactitud (accuracy)?

c) Investiga y explica brevemente otras métricas de evaluación de modelos que puedan ser más adecuadas para conjuntos de datos desbalanceados.

**Recursos**

**Básicos**

Para realizar esta PEC disponéis de los datos contenidos en *star\_classification.csv*, así como los apuntes de la asignatura.

**Criterios de valoración**

Cada uno de los cuatro ejercicios se valorará con 2.5 puntos como máximo. No

No existe una puntuación o un peso específico para cada apartado de los ejercicios, sino que cada ejercicio se valorará globalmente. Recordad que debéis justificar todas las respuestas: una respuesta sin justificar se valorará con cero puntos.

**Formato y fecha de entrega**

Tenéis que entregar la PEC en un fichero zip con el pdf de la memoria al registro de actividades de evaluación continua.

El nombre del archivo tiene que ser ApellidosNombre\_AC\_PEC1 con extensión .zip (ZIP).

Fecha límite: 0**7/06/2024**

Para dudas o aclaraciones sobre el enunciado, dirigíos al consultor responsable de vuestra aula.

| Nota: **Propiedad intelectual**  A menudo es inevitable, al producir una obra multimedia, hacer uso de recursos creados por terceras personas. Es por tanto comprensible hacerlo en el marco de una práctica de los estudios del Grado de Informática, siempre que esto se documente claramente y no suponga plagio en la práctica.  Por lo tanto, al presentar una práctica que haga uso de recursos ajenos, se presentará junto con ella un documento en el que se detallen todos ellos, especificando el nombre de cada recurso, su autor, el lugar donde se obtuvo y el su estatus legal: si la obra está protegida por copyright o se acoge a alguna otra licencia de uso (Creative Commons, licencia GNU, GPL ...). El estudiante deberá asegurarse de que la licencia que sea no impide específicamente su uso en el marco de la práctica. En caso de no encontrar la información correspondiente deberá asumir que la obra está protegida por copyright.  Deberán, además, adjuntar los archivos originales cuando las obras utilizadas sean digitales, y su código fuente si corresponde.  . |
| --- |