| DATOS PERSONALES | FIRMA |
| --- | --- |
| |  |  | | --- | --- | | Nombre: | DNI: | |  |
| Apellidos: |

| ESTUDIO | ASIGNATURA | CONVOCATORIA |
| --- | --- | --- |
| MÁSTER UNIVERSITARIO EN INGENIERÍA MATEMÁTICA Y COMPUTACIÓN (PLAN 2016) | 4391020006.- TÉCNICAS MULTIVARIANTES | Extraordinaria |

| FECHA | MODELO | CIUDAD DEL EXAMEN |
| --- | --- | --- |
| 10-12/09/2021 | Modelo - C |  |

| Etiqueta identificativa |
| --- |
|  |

**INSTRUCCIONES GENERALES**

1. Ten disponible tu documentación oficial para identificarte, en el caso de que se te solicite.
2. Rellena tus datos personales en todos los espacios fijados para ello y lee atentamente todas las preguntas antes de empezar.
3. Las preguntas se contestarán en la lengua vehicular de esta asignatura.
4. Si tu examen consta de una parte tipo test, indica las respuestas en la plantilla según las características de este.
5. Debes contestar en el documento adjunto, respetando en todo momento el espaciado indicado para cada pregunta. Si este es en formato digital, los márgenes, el interlineado, fuente y tamaño de letra vienen dados por defecto y no deben modificarse. En cualquier caso, asegúrate de que la presentación es suficientemente clara y legible.
6. Entrega toda la documentación relativa al examen, revisando con detenimiento que los archivos o documentos son los correctos. El envío de archivos erróneos o un envío incompleto supondrá una calificación de “no presentado”.
7. Durante el examen y en la corrección por parte del docente, se aplicará el Reglamento de Evaluación Académica de UNIR que regula las consecuencias derivadas de las posibles irregularidades y prácticas académicas incorrectas con relación al plagio y uso inadecuado de materiales y recursos.

**Puntuación**

**Preguntas**

* Opción personalizada 10.00 puntos

El examen consta de cuatro preguntas.

Las puntuaciones son:

Pregunta 1: 1.5 puntos.

Pregunta 2: 2 puntos.

Pregunta 3: 1.5 puntos.

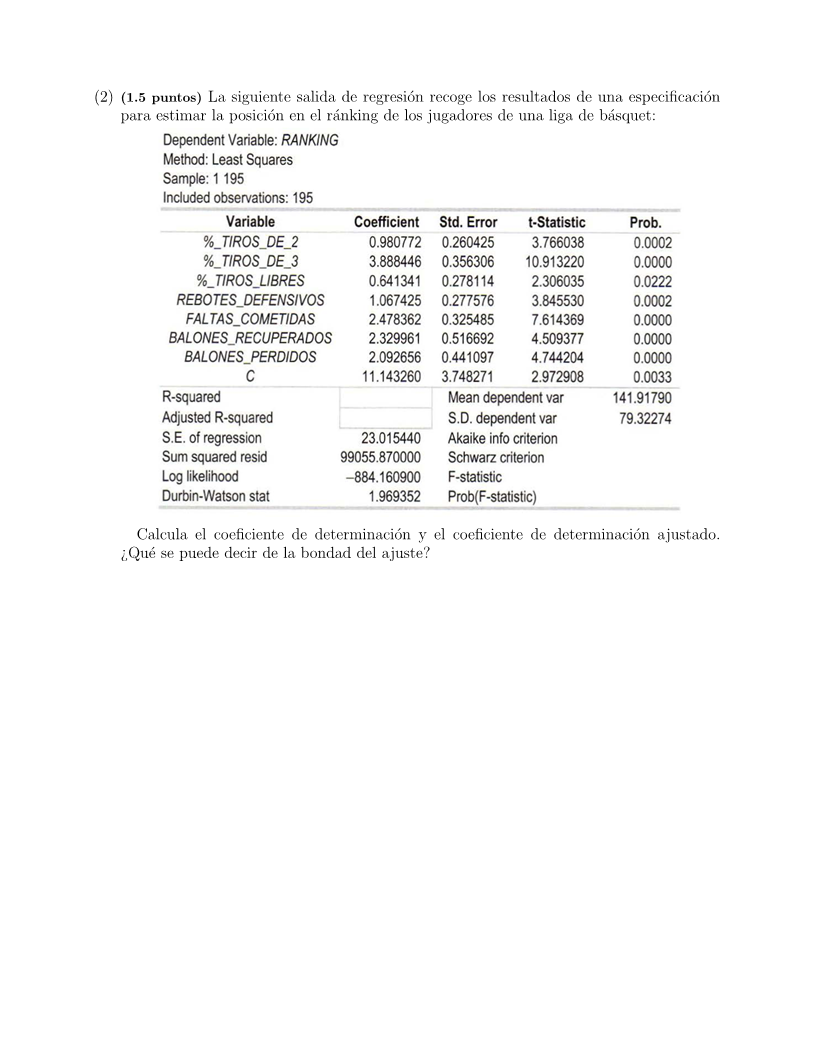
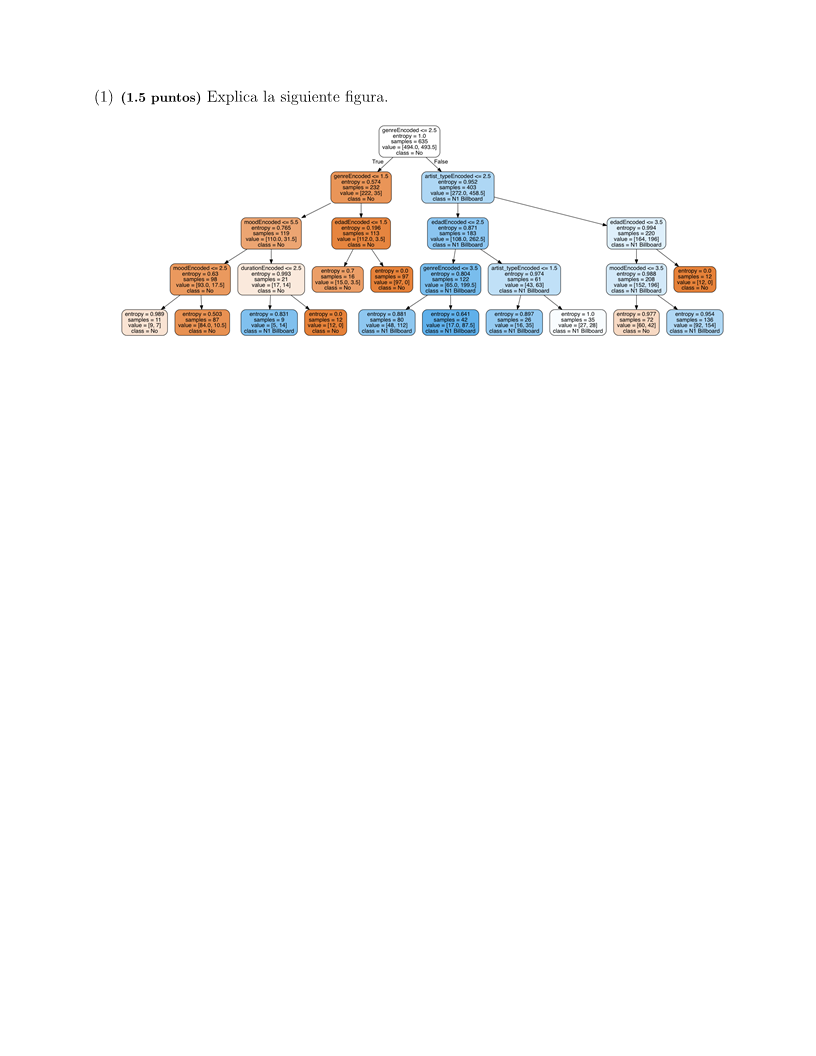
Pregunta 4: 5 puntos.

Responde a las preguntas en el espacio indicado entre las páginas 3 y 15.

Encontrarás las preguntas del examen a partir de la página 16.

¡Suerte!

**1.** Pregunta 1 (Responder en 1 caras)   
  
  
**2.** Pregunta 2 (Responder en 2 caras)   
  
  
**3.** Pregunta 3 (Responder en 1 caras)   
  
  
**4.** Pregunta 4 (Responder en 5 caras)





**Pregunta 1:**

La figura mostrada es un arbol de decision y es utilizada para problemas de clasificacion. Los arboles de decision son binarios, es decir, tiene solamente 2 ramificaciones por cada nodo. El arbol mostrado en la figura es desarrollado a traves de herramientas de SKlearn mediante el objeto DecisionTreeClassifier.

El arbol de decision, tiene varias partes A lo que se logra visualizar(la imagen tiene baja resolucion), se cuenta con genneEncoded, artist\_typeEncoded, edadEndoced, moodEncoded y durationEncoded.

En el primer nivel, tenemos 635 valores de genreEncoded con entropia 1, que se dividen en 232 muestras de genreEncoded menor a 1.5 con entropia 0.574, y 403 muestras de artist\_typeEncoded menor a 2.5 con entropia 0.952. Si sumamos ambas, da el valor de las muestras del nivel superior(403+232 =635). Si bajamos nivel a nivel, vemos que se va ramificando el arbol de decision, considerando que en el ultimo nivel(en este caso son 4 niveles a partir del inicial, 5 en total), se obtienen todos los valores de las muestras en diferentes categorias. Sumando 11+87+9+12+80+42+26+35+72+136 del ultimo nivel, y sumando 16+97+12 del nivel anterior, donde finaliza esa categoria, da un total de 635.

Resumiendo, el arbol de decision, lo que genera son niveles de categorias diferentes, donde un lado es el lado verdadero donde la condicion anterior se cumple, y el otro lado es falso, donde la condicion anterior no se cumple(normalmente a la izquierda es verdadero, y a la derecha es falso). En los arboles de decision no es necesario escalar las variables predictoras.

**Pregunta 2:**

**Pregunta 3:**

Las 3 primeras lineas del codigo son importaciones de numpy(libreria de uso numerico), matplotlib(graficos matematicos y estadisticos), y Sklearn(uso para estadistica y machine learning)**.**

La linea numero 4, genera una semilla para iniciar numeros aleatorios siempre con la misma semilla, y asi, al correr el codigo en diferentes situaciones, siempre generara los mismos valores aleatorios.

La siguiente linea, genera 2 variables, una de samples(muestras), y una de features(caracteristicas), ingresando en muestras un valor de 50, y en caracteristicas un valor de 100, generando en X los valores aleatorios correspondientes. Luego hasta la linea 14, es desarrollo de codigo de coeficientes, y de valores de entrenamiento y test para vatiables X y Y, obteniendo un arreglo X y un arreglo Y, donde el arreglo X es una distribucion uniforme, mientras el arreglo Y es una distribucion normal de valores aleatorios configurados con las variables de muestras y caracteristicas generadas con anterioridad.

A partir de este punto, se importa de SKLearn la libreria que hace referencia a ElasticNet.

Una red elastica funciona con regresiones Lasso y Ridge al mismo tiempo, en la cual actuan las penalizaciones de ambas regresiones a la vez.

En este caso, es ajustado el hiperparametro en 0.1 y el argumento l1\_ratio con valor de 0.198, se entrena el modelo de ElasticNet por medio de “fit” con los datos de entrenamiento, se hace un test por medio de predict con los datos de test, y se obtiene el coeficiente de determinacion R^2, el cual indica si se tiene asociacion entre la variable predictora y la variable de respuesta. En este caso el R^2 score muestra un valor de 0.70455, lo cual significa que la variable es una buena predictora, tiene buena asociacion entre la variable de entrenamiento y la variable de prueba.

**Pregunta 4:**

La pregunta 4 esta descrita en el notebook adjunto como .ipynb y como pdf.

