| DATOS PERSONALES | FIRMA |
| --- | --- |
| |  |  | | --- | --- | | Nombre: Eduard Leonardo | DNI:80542440 | |  |
| Apellidos: Sierra Ballén |

| ESTUDIO | ASIGNATURA | CONVOCATORIA |
| --- | --- | --- |
| MÁSTER UNIVERSITARIO EN INGENIERÍA MATEMÁTICA Y COMPUTACIÓN (PLAN 2016) | 4391020006.- TÉCNICAS MULTIVARIANTES | Ordinaria Número periodo 1823 |

| FECHA | MODELO | CIUDAD DEL EXAMEN |
| --- | --- | --- |
| 14-16/01/2022 | Modelo - C | Colombia Virtual |

| Etiqueta identificativa |
| --- |
|  |

**INSTRUCCIONES GENERALES**

1. Ten disponible tu documentación oficial para identificarte, en el caso de que se te solicite.
2. Si tu examen consta de una parte tipo test, indica las respuestas en la plantilla según las características de este.
3. 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. 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”.
4. 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.
5. No está permitido el uso de Internet ni ningún tipo de comunicación con otra persona.Durante todo el examen tu teléfono móvil debe estar en modo avión.
6. La parte principal de cada pregunta consiste en interpretar y comentar los resultados obtenidos. Si te limitas a hacer los cálculos no vas a poder superar el examen.
7. Es fundamental que las respuestas estén debidamente redactadas, de forma clara y precisa y sin faltas de ortografía.
8. Para hacer el examen puedes utilizar los apuntes del curso y los scripts que hayas preparado y Python para hacer los cálculos.

**Puntuación**

**Preguntas**

* Puntuación máxima 10.00 puntos

El examen constará de un ejercicio práctico (8,5 puntos) y una pregunta teórica (1,5 puntos). Los enunciados están en la página 14 y el espacio para responder el examen está entre las páginas 4 y 13.  
  
**1.** Pregunta

 (Responder en 10 caras)

SOLUCIÓN

1. Ejercicio (8.5 puntos)

Fecha de nacimiento 05/03/1978

m=0+3+3=6

d=14

Se usarán las variables ()

1. (2 puntos)

Quedando los siguientes datos:

|  |
| --- |
| x1 x2 x3 x6 x14 y |
| 0 -0.21 1.59 0.45 -0.42 -1.00 92.55 |
| 1 -0.25 -0.01 2.67 0.36 -0.24 123.16 |
| 2 0.63 1.56 -0.05 0.13 -0.14 75.20 |
| 3 0.65 1.43 -0.17 0.33 -0.06 84.69 |
| 4 0.15 -0.28 -0.81 2.10 0.29 -8.66 |
| 5 0.89 0.07 0.13 -0.19 0.86 106.57 |
| 6 0.16 0.60 1.29 0.68 -0.13 72.97 |
| 7 -0.09 -0.17 0.43 0.15 -0.01 -20.51 |
| 8 0.78 -0.60 -1.26 1.47 1.08 -43.37 |
| 9 -0.43 -0.76 -0.68 0.48 -0.05 -122.46 |

De acuerdo con el comando describe del dataframe, se tiene:

Captura de pantalla de computadora

Descripción generada automáticamente

De acá, se observa que las desviaciones de todas las variables son altas, indicando una dispersión importante en los datos.

Se obtienen las gráficas para las variables x

|  |
| --- |
|  |

Para la variable de respuesta se tiene:

|  |
| --- |
| Gráfico, Histograma  Descripción generada automáticamente |

Las gráficas de las x y de las respuesta (y) confirma la dispersión en los datos y además indican que de manera unidimensional cada variable no correspoden a una distribución normal al rededor de la media, algunas presentan sesgos hacia la derecha o hacia la izquierda. Esto también puede indicar que el dataset es muy pequeño y puede tener errores de muestreo.

División de dataset en entrenamiento (0.8) y prueba (0.2)

Para referencia de los demás métodos se realizó un **análisis de regresión lineal multivariable** que arrojó los siguientes resultados:

Se observa un valor superior a 0.8 en R^2 y unao de R^2 ajsutado inferior a 0.5, difieren bastante, no hay un buen ajuste a un modelo lineal

Captura de pantalla de computadora

Descripción generada automáticamente

**Regresión RIDGE**

Se observa un valor superior a 0.8 en R^2 y unao de R^2 ajsutado inferior a 0.5, difieren bastante, no hay un buen ajuste a un modelo lineal

Captura de pantalla de computadora

Descripción generada automáticamente

**Regresión LASSO**

Captura de pantalla de computadora

Descripción generada automáticamente

**Red Elástica**

**Comparativa de R^2 y R^2 ajustado**

**En general u hubo jasute adecaudao amodelo lienal, por la naturaleza e los datos, en cada corrida la respesta finalzai en respuesta difernetes con valorte de R que no se ajustan a un modelo lineal-**

Pregunta (1.5 puntos)

Un árbol de desición es un método que se emplea para solucionar principalmente problemas de clasificación, y tambié se puede empelar para hacer regresiones. El árbol se construye mediante comprobaciones sobre los datos que permite clasificar en diferentes variables categorícas construyendo nodos de desición a lo largo de las ramas bajando en profundidad, es posible determinar la probabilidad de cada categoría. Permite predecir datos. Su pricipal ventaja es que es sencillo de implementar, pero con respecto a la regresión lineal multivariante es sensible a rotaciones de los datos y en el entrenamiento es sensible a datos anómalos, afectando su capacidad de generalización.

