Curso de Postgrado

Machine learning 2023

Guía 1: Introducción a Python.

Antes de comenzar los problemas genere en su cuenta del servidor (via `ssh/putty`) un nuevo directorio `guia1` donde trabajará y guardará todos los programas y archivos que se producirán en este práctico.

Bucles, condicionales y funciones.

Problema 1: Realice un programa que nos imprima la secuencia de Fibonacci hasta un número dado N_{max}. La secuencia de Fibonacci suma los dos últimos números de la secuencia para generar uno nuevo comenzando por 0 y 1.

Problema 2: Cree un código que pida ingresar una palabra y que luego vaya incrementando el valor de una variable a, con condición inicial $a_0 = 10$, de dos en dos, tantas veces como letras tenga la palabra ingresada. El nuevo valor de la variable deberá ser mostrado en cada iteración. Por ejemplo, si se ingresa una palabra de 3 letras, al ejecutar el código deberá mostrar los valores 12, 14 y 16.

Problema 3: Un año es año bisiesto si es divisible por 4, excepto los años que son divisibles por 100, aunque los que son divisibles por 400 si lo son. Escriba un script que pregunte por un año al usuario y que el programa determine si es año bisiesto.

Numpy. Vectores, matrices, arrays. Clases y objetos.

Problema 4: La distancia entre dos vectores $\mathbf{x} = (x_1, ..., x_n)$, $\mathbf{y} = (y_1, ..., y_n)$ está definida por:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_j (x_j - y_j)^2} \tag{1}$$

Diseñe e implemente una función que calcule la distancia entre dos vectores de dimensión n (listas o numpy arrays). Genere un vector aleatorio de dimensión $(1, 3)$ y una matriz aleatoria de dimensión $(100, 3)$, ambos arrays numpy, con densidad de probabilidad normal de media 0 y desviación estándar 5. Encuentre las distancias entre el vector y los vectores fila de la matriz, y guárdelas en un array. Determine la distancia máxima.

Problema 5: Crear una función con cuatro parámetros de entrada `a1`, `a2`, `b1`, `b2` que representen los valores de una matriz de dimensión $(2, 2)$. La función debe calcular el determinante y evaluar si la matriz es singular (determinante igual a cero) o no singular (determinante distinto de cero). La salida de la función debe ser `True` si la matriz es singular.

Crear una segunda función que determine si una matriz de dimensión (n, n) es singular, utilizando la función `numpy.linalg.det`. Evalúe las dos funciones realizadas generando una matriz aleatoria de dimensión (n, n) y llamando a las funciones.

Problema 6: Desarrolle una clase de objetos que trabaje con vectores de dimensión n (no use arrays de numpy pero sí listas).

(a) Inicialice la clase definiendo el vector (la dimensión y el tipo).

(b) Implemente la función suma de vectores.

(c) Implemente la función producto interno.

(d) Implemente la función de la media.

(e) Implemente una función que determine si dos vectores son ortonormales usando la función del inciso anterior.

(f) Implemente la función rotación de vectores alrededor del eje z.

Problema 7: Desarrolle una clase de objetos que trabaje con polinomios de grado n.

(a) Inicialice la clase definiendo el polinomio.

(b) Implemente el método suma de polinomios de grado n y m.

(c) Implemente el método que evalue el polinomio (por default).

(d) Desarrolle un método derivada del polinomio que devuelva la derivada.

(e) Implemente un método que imprima el polinomio con la forma usual i.e. $p(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0$.

Lectura y escritura. Graficación.

Problema 8: Escribir un programa que cree un archivo "test.csv" y lo llene con líneas que contengan 3 columnas: la primera con un índice entero creciente (0 a 50), la segunda $f(x) = x^2$ y la tercera $f(x) = 2x$.

Resultado esperado:

x	x^2	$2x$
0	0	0
1	1	2
2	4	4
$\vdots$	$\vdots$	$\vdots$

Leer el archivo "test.csv" y graficar las columnas.

Problema 9: Realice la graficación de la función $f(x) = sin(x)/|x|$ entre -2π y 2π usando arrays para x e y. Busque la resolución adecuada. Ponga de título la función y defina los ejes como x y $f(x)$. Superponga con la función $f_2(x) = x/6$. Coloque una leyenda apropiada y que no se superponga con los gráficos. Guarde la salida del gráfico en un archivo .png.

This data sets consists of 3 different types of irises' (Setosa, Versicolour, and Virginica) petal and sepal length, stored in a 150x4 numpy.ndarray

Problema 10: El dataset Iris consiste en un conjunto de mediciones del tamaño de los pétalos y sépalos de tres tipos de flores de Iris (setosa, versicolor y virgínica). Utilice la librería Matplotlib para realizar los siguientes puntos:

- Cree un boxplot para comparar el ancho y largo de los sépalos y pétalos de las flores. Este gráfico permitirá observar las variaciones en las medidas de estas partes de la flor para cada tipo.

- Genere dos scatterplots para explorar la relación entre la longitud y el ancho de los sépalos en un primer gráfico y en un segundo gráfico lo mismo para pétalos. Asigne colores distintos a cada especie de flor para visualizar mejor la diferencia.

Problemas Desafio.

Problema 11: Realizar un programa que, dada una fecha, determine qué día de la semana cae. Tenga en cuenta que el 1 de Enero de 1900 fue lunes. Preguntar por una fecha posterior a esta y determinar el día de la semana. Reutilice la función desarrollada en la guía. Corrobore con la fecha de hoy y con la fecha de su nacimiento.

Problema 12: La población de una determinada especie se regula de acuerdo a la ecuación logística:

$$x_{i+1} = x_i + rx_i(1 - x_i) \tag{2}$$

donde $i+1$ es la generación siguiente de la especie, i la generación actual y r es la tasa de reproducción de la especie.

(a) Realice una función que actualice el valor de la población. Defina el valor de r como una variable global.

(b) Realice una función que simule el crecimiento de la población de la especie en función de la generación $i = 1, 2, ..., T$ utilizando la función previamente desarrollada.

(c) Implemente un programa que simule el crecimiento de la población de la especie en función de la generación $i = 1, 2, ..., T$ para $r = 0.1$ y $x_0 = 0.1$.

(d) Calcule las poblaciones para dos especies, una con $r = 0.1$ y otra con $r = 0.2$ y $x_0 = 0.1$ en ambas.

(e) Adapte la función para que en el mismo ciclo de $i = 1, 2, ..., T$ se calculen ambas poblaciones.

Problema 13: Se requiere realizar evaluaciones de la función de Bessel de orden 0 a través de la serie

$$J_0(x) = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2} - \frac{x^6}{2^2 4^2 6^2} + \cdots \tag{3}$$

(a) Diseñar una rutina tal que, dados n (número de términos) y x, calcule la serie de la función de Bessel.

(b) Realice una función para determinar cuántos términos son necesarios para obtener una precisión δ_p arbitraria.

(c) Grafique entre $n = 0$ y $n = 20$ la función de Bessel aproximada y la función de Bessel provista por scipy.

(d) Determinar a partir del gráfico cuál es la primer raíz positiva de la función de Bessel de orden 0 con una resolución de $delta_p = 0.5, 0.1, 0.01$.

(e) Utilice el método de bisección para hallar la primer raíz positiva de la función de Bessel de orden 0 con una precisión $\epsilon = 10^{-3}$.

Ayuda: Para realizar el método de bisección, programe una función que tenga de entrada la función, el intervalo $[a, b]$ donde se quiere encontrar el cero y la precisión requerida, y que tenga de salida la raíz obtenida. El método de la bisección consiste en definir $c = \frac{a-b}{2}$. Si $f(c)f(b) < 0$, la función corta al eje x entre c y b, por lo que se define $a = c$. Si no, la función corta al eje x entre a y c, por lo que se define $b = c$. Repetir este proceso las veces que sea necesaria hasta obtener la precisión deseada.

Problema 14: Interferencia de ondas.

Suponga que se tiran un par de piedras en una pileta, las cuales generan ondas circulares que se desparraman con el tiempo, produciendo interferencia entre las dos fuentes. Cada onda se puede representar por $\zeta_i = \zeta_0 \sin(kr_i)$, con $i = 1, 2$ y $r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$, donde asumimos que ambas ondas tienen la misma amplitud y longitud de onda pero difieren en su centro (lugar donde cayó la piedra). El patrón resultante de la pileta es la suma $\zeta = \zeta_1 + \zeta_2$. Notar que ζ representa el desplazamiento vertical de la superficie del agua.

(a) Realice una función que determine el patrón ondulatorio de una piedra.

(b) Realice una función que permita evaluar el patrón ondulatorio de una piedra en un dominio o grilla cuadrada, dados δx y δy y un largo del estanque de $L_x = L_y$.

(c) Evalúe el patrón de interferencia, asumiendo que las ondas tienen una longitud de onda $\lambda = 5cm$ y número de onda $k = 2\pi/\lambda$, una amplitud de $A = 1cm$ y que los centros están separados por $20cm$. El estanque tiene un metro de largo y requerimos una resolución cuadrada de $1cm$.

(d) Realice un gráfico de densidad del patrón de interferencia.