

Proyecto 2

Usted implementará un sistema de visión artificial que detectará un número escrito a mano en un papel. Para esto, usará el dataset «digits» disponible en el paquete scikit-learn.

Se le pide:

- a) Manipular el dataset digits de sklearn para generar una matriz de 8x8 con la imagen promedio de cada uno de los 10 dígitos disponibles en este dataset. Puede generarlas todas juntas, o generarlas en base a un menú, o generarlas en base al número que el usuario ingrese por teclado.
- b) Muestre las matrices de 8x8 de dichos 10 promedios en alguna forma que sea visualmente comprensible para los seres humanos. Puede mostrar su imagen de la manera que usted desee. Aquí se le presentan algunas ideas:
 - Puede usar el formato condicional de Excel
 - Puede imprimir los números con colores en la salida de pycharm (si lo hace en pycharm, puede poner los números 0 de color negro y los números 16 de algún color elegido por usted y usar colores intermedios para los valores entre 1 y 15).
 - Puede generar una imagen nueva a partir de su matriz de 8x8 y mostrar la imagen en algún reproductor de imágenes.
 - Puede generar una imagen nueva a partir de su matriz de 8x8 y mostrar la imagen desde Pycharm.
- c) Lea un nuevo dígito. Este puede ser un dígito dibujado por usted y convertido a 8x8 con valores entre 0 y 16 donde 0 es completamente blanco y 16 es completamente negro (para eso deberá convertir los 255 a 0 y los 0 a 16), o también podría ser una entrada por teclado con los 64 valores.
- d) Una vez capturado el nuevo número en 8x8, busque los 3 dígitos más parecidos a este. Es decir, debe recorrer todos los dígitos del dataset digits y calcular la distancia euclidiana como la raíz cuadrada de los cuadrados de los 64 residuales de todos los números del dataset digits, y quedarse con los 3 dígitos más cercanos (aquellos que tengan las 3 menores distancias euclidianas con respecto a su nuevo número).

- e) Imprima los targets que corresponden los 3 dígitos más cercanos a su nuevo dígito ingresado.
- f) Intente clasificar a su nuevo dígito:
- Si 2 o 3 de estos targets corresponden a un mismo valor, o si los 3 corresponden a un mismo valor, su programa deberá concluir que su nuevo dígito ingresado corresponde a dicho target también, por lo que deberá imprimir un mensaje que diga “Soy la inteligencia artificial, y he detectado que el dígito ingresado corresponde al número X”, donde X es un número entre 0 y 9.
 - Si los 3 targets son diferentes, usted decida qué hacer para clasificar al número nuevo. La tarea de usted es clasificar a dicho número. Puede implementar el método que desee.
- g) Finalmente, calcule la distancia el nuevo dígito ingresado y los 10 dígitos promedios generados en el inciso a) de este enunciado. Identifique cuál es la menor de las 10 distancias. Su programa deberá concluir que el dígito ingresado corresponde a aquél promedio que generó la menor distancia, por lo que deberá imprimir un mensaje que diga “Soy la inteligencia artificial versión 2, y he detectado que el dígito ingresado corresponde al número X”, donde X es un número entre 0 y 9.
- h) Indique cuál de los dos métodos cree usted que es mejor, el de la versión 1 (inciso f de este enunciado) o el de la versión 2 (inciso g este enunciado).

Entregables:

Su trabajo deberá contener los siguientes elementos:

- Un zip con el proyecto de Python (debe incluir los archivos generados / usados).
- Un informe en Word en el que describa los pasos realizados, y se evidencie que está entregando todos los puntos solicitados en los párrafos anteriores. Al final de su documento debe escribir 4 conclusiones de su trabajo. En la carátula de este informe debe incluir el nombre y código de los integrantes de su grupo y el porcentaje de participación de cada integrante.
- Las conclusiones de su trabajo deben estar adecuadamente formuladas y deben poder sustentarse con el contenido de su trabajo.
- Una presentación oral.

Importante:

Durante la exposición, todos los miembros del equipo deben tener conocimiento completo sobre el funcionamiento del código que están exponiendo y conocimiento completo sobre el contenido del informe que presentaron. Si un estudiante no comprende el código que está entregando, se considerará que no tuvo participación en la elaboración de dicho código y no se le dará el puntaje del desarrollo de software (10 puntos según la rúbrica).

Rúbrica:

Criterio	EXCELENTE	ADECUADO	MÍNIMO	INSUFICIENTE
Desarrollo de software	Diseña y elabora el software para lograr una solución adecuada al problema planteado. El software debe ser ordenado, claro y óptimo. (10 p.)	Diseña y elabora el software para lograr una solución adecuada al problema planteado. El software funciona pero no es ordenado, claro y óptimo. (6 p.)	Diseña el software para lograr una solución adecuada al problema planteado. El software no se concluyó adecuadamente. (4 p.)	No logra el diseño ni la implementación correcta del software. (2 p.)
Presentación escrita	El informe contiene las secciones de Antecedentes, Fundamento Teórico, Métodos y Desarrollo y Conclusiones. Estas últimas, adecuadamente formuladas. (5 p.)	El informe contiene las secciones de Antecedentes, Fundamento Teórico, Métodos y Desarrollo, pero no pone énfasis en las conclusiones. (3 p.)	El informe contiene menos de la mitad de las secciones estipuladas, incluyendo conclusiones. (2 p.)	El informe contiene menos de la mitad de las secciones estipuladas, sin incluir conclusiones. (0 p.)
Presentación oral	El alumno presenta el proyecto en forma adecuada y responde a las preguntas del profesor en forma lógica y coherente. (5 p.)	El alumno presenta el proyecto en forma adecuada, pero no responde a todas las preguntas del profesor en forma lógica y coherente. (3 p.)	El alumno no presenta el proyecto en forma adecuada, pero responde a las preguntas del profesor en forma lógica y coherente. (2 p.)	El alumno no presenta el proyecto en forma adecuada ni responde a las preguntas del profesor en forma lógica y coherente. O no se presenta a la presentación oral. (0 p.)