



Cuestionario del primer ejercicio

Especialidad: A6 D3. Modelado y análisis de datos.

Por favor, lea detenidamente antes de comenzar:

- **NO** abra el **CUESTIONARIO** ni empiece el examen hasta que se le indique.
- Para realizar este primer ejercicio se hace entrega de dos documentos:
 1. Cuadernillo con el **cuestionario de preguntas tipo test**, con cuatro respuestas alternativas sobre las materias del programa de esta convocatoria.
 2. **Hoja de respuestas** donde se consignará la respuesta correcta a cada pregunta.
- Al finalizar la prueba se hará entrega de la hoja de respuestas. El cuadernillo con el cuestionario se podrá utilizar como borrador y se podrá llevar por el opositor al finalizar el tiempo marcado para el ejercicio.
- Sólo se calificará las respuestas marcadas en la **HOJA DE RESPUESTAS**
- Una vez abierto el cuestionario, compruebe que consta de todas las páginas y preguntas y que sea legible. En caso contrario solicite uno nuevo al personal del aula.
- Verifique que el número de la solapa donde se recogen sus **datos personales coincide con el número de la hoja** de examen donde se consignan las respuestas.
- El examen se realizará con bolígrafo azul o negro. Si no dispone de uno, solicítelo al Tribunal.
- El cuestionario consta de **100 (CIEN) preguntas**, cada una de ellas con **cuatro respuestas alternativas**, de las cuales **sólo una de ellas es correcta, más 5 (CINCO) preguntas adicionales de reserva**, que serán valoradas en el caso de que se anule algunas de las 100 anteriores. **Marque con una equis la respuesta elegida** en la celda correspondiente a la pregunta, de forma clara (ver fig. 1).
- **Las respuestas múltiples, poco claras o dudosas, se considerarán como no contestadas.** Si desea corregir una respuesta, **rodee la opción incorrecta** con un círculo (ver fig. 2) y marque con una equis la nueva opción que elige.

1	A	B	C	D
	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fig 1

1	A	B	C	D
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fig 2.

- Todas las preguntas tendrán el mismo valor y **las contestaciones erróneas se penalizarán** con 1/3 del valor de una contestación correcta.
- **NO Separe** ninguna de las copias de la **HOJA DE RESPUESTAS**. Una vez finalizado, el personal del aula le indicará los pasos a seguir.
- **Dispone de 90 minutos**, máximo, para realizar este ejercicio.



Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025
Primer Ejercicio

Fecha:
10/05/2026
Página: 2 de 28




1. ***En el contexto actual de la Ciencia de Datos aplicada a la investigación científica, ¿qué caracteriza al denominado cuarto paradigma de la ciencia?***
 - A) El reemplazo del método científico por algoritmos de inteligencia artificial.
 - B) El uso de grandes volúmenes de datos junto con técnicas computacionales para el descubrimiento científico.
 - C) El desarrollo de modelos teóricos sin necesidad de experimentación.
 - D) La automatización completa de los procesos experimentales en laboratorio.

2. ***¿Qué característica define a un modelo matemático robusto?***
 - A) Es un modelo que proporciona respuestas exactas y perfectas sin margen de error.
 - B) Es un modelo que mantiene una precisión aceptable frente a perturbaciones o errores en los datos.
 - C) Es un modelo que utiliza exclusivamente variables constantes para evitar la incertidumbre.
 - D) Es un modelo que solo puede ser resuelto mediante métodos analíticos exactos.

3. ***En el método de Euler explícito para el problema de valor inicial $x' = f(t, x)$, ¿cuál es la fórmula correcta de actualización desde el paso n hasta el paso $n+1$?***
 - A) $x_{n+1} = x_n + h \cdot f(t_n, x_n)$
 - B) $x_{n+1} = x_n + h \cdot f(t_{n+1}, x_n)$
 - C) $x_{n+1} = x_n + (h/2) \cdot [f(t_n, x_n) + f(t_{n+1}, x_n)]$
 - D) $x_{n+1} = x_n + h \cdot f(t_{n+h/2}, x_{n+1/2})$

4. ***En computación de altas prestaciones (High Performance Computing, HPC), la interfaz MPI (Message Passing Interface) se utiliza principalmente para trabajar con:***
 - A) Sistemas de memoria compartida.
 - B) Sistemas de memoria distribuida.
 - C) Procesamiento exclusivo en GPU.
 - D) Bases de datos distribuidas.

5. ***En el uso de la biblioteca pandas, ¿qué describe mejor la estructura de datos denominada DataFrame***
 - A) Una matriz estrictamente numérica sin etiquetas en filas ni columnas.
 - B) Una estructura bidimensional etiquetada con columnas potencialmente de distinto tipo.
 - C) Una lista de diccionarios utilizada para almacenar datos jerárquicos.
 - D) Un objeto optimizado exclusivamente para series temporales.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 4 de 28</p>
---	---	---

6. *¿En qué consiste el concepto de modularización dentro del diseño de programas?*
- A) Dividir un problema complejo en unidades funcionales independientes con interfaces bien definidas, que pueden desarrollarse y probarse de forma separada.
 - B) Reutilizar fragmentos de código mediante herencia entre clases, de forma que cada subclase extienda el comportamiento de su módulo padre.
 - C) Encapsular todos los datos del programa en estructuras globales accesibles desde cualquier parte del código para facilitar su reutilización.
 - D) Organizar el código en un único fichero estructurado mediante comentarios y secciones para mejorar su legibilidad.
7. *¿Qué describe correctamente una condición de contorno de Neumann en una PDE?*
- A) Fija el valor de la solución en la frontera.
 - B) Fija la derivada normal o flujo en la frontera.
 - C) Impone una combinación lineal del valor de la solución y su derivada normal en la frontera.
 - D) Establece que la solución en un extremo del dominio sea igual a la del extremo opuesto.
8. *En programación paralela, ¿cuál de las siguientes afirmaciones describe correctamente el modelo de memoria distribuida?*
- A) Todos los procesos acceden a una memoria global compartida.
 - B) Los procesos comparten memoria sin necesidad de comunicación.
 - C) Cada proceso tiene su propio espacio de memoria y se comunica mediante paso de mensajes.
 - D) No existe comunicación entre procesos.
9. *En el uso de la biblioteca NumPy, ¿qué se entiende por 'broadcasting'?*
- A) La ejecución de operaciones elemento a elemento entre arrays de distinta forma compatible.
 - B) La copia explícita de arrays para que tengan la misma dimensión.
 - C) La conversión automática de tipos de datos en operaciones aritméticas.
 - D) La paralelización de operaciones mediante múltiples hilos de ejecución.




10. ***En el contexto de la programación orientada a objetos, ¿qué principio establece que los módulos deben ser capaces de ser utilizados por otros, pero también ser susceptibles de cambios en su implementación sin afectar a sus clientes?***
- A) El Principio de Abstracción Total.
 - B) El Principio de Abierto/Cerrado.
 - C) El Principio de Inversión de Dependencia.
 - D) El Principio de Encapsulamiento Estricto.
11. ***¿Cuál es el propósito principal de la técnica de dropout durante el entrenamiento de redes neuronales profundas?***
- A) Acelerar la convergencia eliminando de forma permanente las conexiones con pesos cercanos a cero.
 - B) Eliminar las capas que contribuyen menos a la función de pérdida.
 - C) Comprimir el modelo descartando las unidades con menor contribución al gradiente tras cada época.
 - D) Regularizar el modelo desactivando aleatoriamente unidades para reducir el sobreajuste.
12. ***En estadística, la media muestral se considera:***
- A) Un parámetro poblacional.
 - B) Un estimador puntual de la media poblacional.
 - C) Un intervalo de confianza.
 - D) Una función de distribución.
13. ***¿Qué norma ISO certifica que un repositorio digital cumple los requisitos del modelo OAIS (Open Archival Information System)?***
- A) ISO 9001, estándar de gestión de calidad para organizaciones.
 - B) ISO 16363, norma de auditoría y certificación de repositorios digitales de confianza.
 - C) ISO 27001, estándar de gestión de seguridad de la información.
 - D) ISO 14721, norma que define el modelo de referencia OAIS.



14. *¿Cuál es el error de truncamiento global de la regla del trapecio compuesta si se asume que la segunda derivada de la función está acotada (es finita) en todo el intervalo?*
- A) Es proporcional al cuadrado del tamaño del paso ($O(h^2)$).
 - B) Es proporcional al cubo del tamaño del paso ($O(h^3)$).
 - C) Es inversamente proporcional al número de segmentos ($O(1/n)$).
 - D) El error es cero, ya que el método es exacto para funciones con segunda derivada constante.
15. *¿Cuál es la forma correcta de implementar un training step personalizado con `tf.GradientTape`?*
- A) `tape.gradient(model.trainable_variables, loss)`
 - B) `tape.gradient(loss, model.trainable_variables)`
 - C) `tf.gradients(loss, model.trainable_variables)` dentro del bloque with tape
 - D) `tape.compute_gradients(loss)`
16. *En teoría de la probabilidad, la esperanza matemática de una variable aleatoria representa:*
- A) El valor máximo posible.
 - B) El valor promedio esperado a largo plazo.
 - C) La dispersión de los datos.
 - D) La mediana de la distribución.
17. *En el contexto del desarrollo de software abierto (open source), ¿qué característica es esencial?*
- A) El acceso libre al código fuente para su uso, modificación y redistribución.
 - B) La obligación de que el software sea siempre gratuito.
 - C) La prohibición de uso en entornos comerciales.
 - D) El desarrollo exclusivo por organismos públicos.
18. *En la Cuadratura de Gauss, ¿cuál es la ventaja fundamental de este método frente a las fórmulas de Newton-Cotes (como Trapecio o Simpson)?*
- A) Permite utilizar puntos equiespaciados para facilitar el cálculo manual.
 - B) No requiere el conocimiento de la función, solo de sus derivadas en los extremos.
 - C) Elige los puntos de evaluación de forma óptima para maximizar la precisión.
 - D) Siempre garantiza un error menor que el método de Romberg para cualquier tipo de función.



19. ***¿Por qué se recomienda usar `tf.keras.layers.BatchNormalization` en redes profundas?***
- A) Normaliza las activaciones intermedias, estabilizando y acelerando el entrenamiento.
 - B) Elimina la necesidad de utilizar funciones de activación no lineales.
 - C) Reescala los gradientes durante la retropropagación para mantener su magnitud constante a lo largo de todas las capas.
 - D) Fuerza que las distribuciones de entrada de cada capa coincidan con la distribución del conjunto de test, mejorando la generalización.
20. ***En teoría de la probabilidad, la función de distribución acumulada (Cumulative Distribution Function, CDF) cumple que:***
- A) Es decreciente.
 - B) Es constante.
 - C) Es no decreciente y está acotada entre 0 y 1.
 - D) Puede tomar valores negativos.
21. ***En un proyecto de investigación científica que adopta metodologías ágiles, ¿qué ventaja aporta el trabajo en iteraciones cortas (sprints) frente a enfoques tradicionales?***
- A) Permite eliminar la necesidad de planificación inicial del proyecto.
 - B) Facilita la adaptación a cambios en hipótesis o datos durante el proyecto.
 - C) Garantiza que los resultados sean correctos desde la primera iteración.
 - D) Sustituye la necesidad de validación experimental continua.
22. ***¿Cuál es la principal limitación del método de Euler simple (hacia adelante) en términos de error y estabilidad?***
- A) Su error de truncamiento global es de segundo orden ($O(h^2)$), lo que lo hace muy lento.
 - B) Es un método implícito que requiere resolver ecuaciones no lineales en cada paso.
 - C) Su error de truncamiento global es solo de primer orden ($O(h)$) y puede volverse inestable.
 - D) No puede aplicarse a ecuaciones diferenciales de primer orden.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 8 de 28</p>
---	---	---

23. *¿Bajo qué condición es correcto aplicar el algoritmo de Dijkstra para obtener caminos mínimos?*
- A) El grafo no debe contener ciclos de ningún tipo.
 - B) El grafo debe ser no dirigido para que la relajación de aristas funcione en ambos sentidos.
 - C) Los pesos de las aristas deben ser no negativos.
 - D) Todos los pesos deben ser estrictamente positivos, ya que aristas con peso cero producen bucles infinitos.
24. *En variables aleatorias continuas, una función de densidad de probabilidad (Probability Density Function, PDF) cumple que:*
- A) Puede tomar valores negativos.
 - B) Su integral en todo el dominio es igual a 1.
 - C) Solo se define para variables discretas.
 - D) Es siempre constante.
25. *¿Cómo se denomina la incertidumbre evaluada a partir del análisis estadístico de series de observaciones, según la nomenclatura del GUM (Guide to the Expression of Uncertainty in Measurement)?*
- A) Incertidumbre de tipo B.
 - B) Incertidumbre sistemática.
 - C) Incertidumbre de tipo A.
 - D) Incertidumbre expandida.
26. *¿Qué caracteriza a un sistema rígido de ecuaciones diferenciales desde el punto de vista del análisis numérico?*
- A) Es un sistema que incluye componentes con variaciones muy rápidas y muy lentas.
 - B) Es un sistema donde todas las variables dependientes cambian a la misma velocidad constante.
 - C) Es un sistema que solo puede resolverse mediante métodos analíticos exactos debido a su complejidad.
 - D) Es un sistema donde la solución siempre tiende a infinito independientemente de las condiciones iniciales.



27. *¿Qué mide principalmente la centralidad de intermediación (betweenness centrality) de un nodo en una red?*
- A) La inversa de la distancia media desde ese nodo a todos los demás nodos alcanzables.
 - B) La fracción de caminos más cortos entre todos los pares de nodos que pasan por él.
 - C) El número de aristas incidentes al nodo, ponderadas o no.
 - D) La importancia derivada de estar conectado a otros nodos que a su vez son importantes.
28. *En el diseño de sistemas multiagente, ¿cuál es la principal diferencia entre un agente reactivo y un agente deliberativo?*
- A) El agente reactivo utiliza modelos internos del entorno.
 - B) El agente deliberativo toma decisiones basadas en representaciones internas del entorno.
 - C) Ambos agentes utilizan planificación a largo plazo.
 - D) El agente reactivo siempre es más preciso.
29. *¿Cuál de las siguientes estrategias permite reducir la incertidumbre estadística pero NO la sistemática?*
- A) Recalibrar el instrumento de medida.
 - B) Utilizar un método de medición alternativo e independiente.
 - C) Corregir el sesgo conocido del instrumento.
 - D) Aumentar el número de mediciones y promediar los resultados.
30. *En las simulaciones de dinámica molecular, ¿cuál es una de las razones principales por las que el algoritmo de integración de Verlet es preferible sobre los métodos de Runge-Kutta de orden superior?*
- A) Porque permite utilizar un paso de tiempo (h) significativamente mayor sin perder precisión local.
 - B) Porque conserva con gran estabilidad la energía del sistema a largo plazo.
 - C) Porque no requiere el cálculo de las fuerzas (aceleraciones) en cada paso.
 - D) Porque garantiza que el error de truncamiento local sea de orden $O(h^6)$.



31. ***En el contexto de la criptografía de clave pública, ¿qué operación matemática sobre grandes enteros se realiza durante los procesos de cifrado y descifrado en el algoritmo RSA?***
- A) La factorización rápida de semiprimos.
 - B) La exponenciación modular.
 - C) El cálculo del máximo común divisor mediante el algoritmo extendido de Euclides.
 - D) La generación de números primos grandes mediante tests de primalidad probabilísticos.
32. ***En sistemas multiagente distribuidos, el comportamiento emergente se define como:***
- A) La suma directa de las acciones individuales.
 - B) Un comportamiento planificado centralmente.
 - C) Un patrón global que surge de interacciones locales sin control central.
 - D) La ejecución independiente de tareas.
33. ***¿Cuál es la característica fundamental que distingue el aprendizaje supervisado del no supervisado?***
- A) El aprendizaje supervisado requiere mayor capacidad computacional que el no supervisado.
 - B) El aprendizaje supervisado utiliza datos de entrenamiento etiquetados; el no supervisado trabaja con datos sin etiquetar.
 - C) El aprendizaje no supervisado solo puede aplicarse a datos numéricos continuos.
 - D) El aprendizaje supervisado no requiere datos de validación para evaluar el modelo.
34. ***¿En qué consiste la técnica de condiciones de contorno periódica utilizada frecuentemente en simulaciones de sistemas moleculares?***
- A) En aplicar una pared rígida e impenetrable en los límites de la celda de simulación.
 - B) En rodear la celda central de simulación con réplicas exactas de sí misma.
 - C) En reiniciar la posición de todas las partículas a su estado inicial cada vez que una colisiona con el borde.
 - D) En variar la temperatura de forma cíclica para evitar que el sistema se congele en un mínimo local.



35. *¿Qué técnica se emplea habitualmente para calcular $a^n \bmod m$ de forma eficiente cuando n es muy grande?*
- A) Descomponer n en sus factores primos y calcular $a^p \bmod m$ para cada primo p , combinando después los resultados con el Teorema Chino del Resto.
 - B) Aplicar el algoritmo extendido de Euclides para reducir el exponente n al menor representante módulo $\phi(m)$.
 - C) Precalcular una tabla con todos los residuos de $a^k \bmod m$ para k de 1 a \sqrt{n} y buscar el resultado por interpolación.
 - D) Exponenciación binaria, descomponiendo n en potencias de 2 y realizando cuadrados sucesivos.
36. *En inferencia estadística, un intervalo de confianza más amplio indica:*
- A) Mayor precisión en la estimación.
 - B) Mayor incertidumbre sobre el parámetro estimado.
 - C) Menor variabilidad muestral.
 - D) Ausencia de error.
37. *En el contexto del aprendizaje supervisado, ¿qué fenómeno describe la situación en la que un modelo tiene un error bajo en entrenamiento pero alto en validación?*
- A) Underfitting, causado por un modelo con excesiva capacidad representativa.
 - B) Overfitting, causado por un modelo que ha aprendido el ruido de los datos de entrenamiento.
 - C) Regularización, causada por penalizar excesivamente los parámetros del modelo.
 - D) Bias-variance tradeoff, causado por un modelo con varianza y sesgo simultáneamente altos.
38. *¿Qué problema principal de las funciones de activación saturadas (como la sigmoide o la tangente hiperbólica) ayuda a mitigar el uso de la Unidad Lineal Rectificada (ReLU) en redes neuronales profundas?*
- A) La excesiva carga computacional del cálculo de exponenciales.
 - B) La inestabilidad numérica producida por valores de salida negativos.
 - C) La necesidad de normalizar los pesos de la red en cada iteración.
 - D) El desvanecimiento del gradiente.




39. *¿Cuál es la complejidad algorítmica de la FFT de Cooley-Tukey para una señal de N muestras?*
- A) $O(N^2)$
 - B) $O(N)$
 - C) $O(N \log N)$
 - D) $O(N^2 \log N)$
40. *En un test t de Student, la hipótesis nula suele establecer:*
- A) La existencia de diferencias significativas.
 - B) La igualdad de medias entre grupos.
 - C) La ausencia de varianza.
 - D) La independencia de variables.
41. *¿Cuál es la diferencia fundamental entre un modelo de datos raster y un modelo vectorial en un Sistema de Información Geográfica (GIS)?*
- A) El modelo raster solo puede almacenar datos continuos; el vectorial solo puede almacenar datos categóricos.
 - B) El modelo raster representa la información mediante entidades geométricas discretas (puntos, líneas y polígonos); el vectorial mediante una cuadrícula regular de celdas.
 - C) El modelo raster representa el espacio mediante una cuadrícula regular de celdas con valores asociados; el vectorial mediante entidades geométricas discretas.
 - D) Ambos modelos son equivalentes en precisión espacial para cualquier tipo de fenómeno geográfico.
42. *¿Cuál es el efecto principal de una capa de agrupamiento, como el max-pooling, sobre las representaciones aprendidas por una CNN?*
- A) Ayudar a que la representación sea invariante a pequeñas traslaciones de la entrada.
 - B) Aumentar la resolución espacial de los mapas de características para detectar detalles más finos.
 - C) Actuar como una función de activación no lineal que sustituye a la ReLU.
 - D) Reducir el número de canales de salida manteniendo la misma anchura y altura.




43. *¿Qué relación establece el teorema de la convolución entre los dominios temporal y frecuencial?*
- A) Convolucionar en el tiempo equivale a multiplicar en frecuencia.
 - B) Multiplicar en el tiempo equivale siempre a derivar en frecuencia.
 - C) Multiplicar en el tiempo equivale a sumar las transformadas de Fourier de ambas señales.
 - D) Convolucionar en el tiempo equivale a dividir las transformadas de Fourier punto a punto.
44. *En la prueba de chi-cuadrado de independencia, la hipótesis nula establece que:*
- A) Existe relación entre variables.
 - B) Las variables son independientes.
 - C) Las medias son iguales.
 - D) Los datos siguen una distribución normal.
45. *¿Cuál es la operación matemática que implica la inversión (volteo) del kernel y su desplazamiento sobre una imagen digital para aplicar un filtro espacial?*
- A) Correlación cruzada normalizada entre la imagen y una plantilla de referencia, utilizada para detectar objetos específicos.
 - B) Transformada de Fourier bidimensional de la imagen completa.
 - C) Convolución entre la imagen y un kernel predefinido.
 - D) Producto escalar entre los histogramas de intensidad de la imagen.
46. *¿Qué estructura de datos se utiliza típicamente para gestionar los vértices descubiertos en el algoritmo de Búsqueda en Anchura (BFS) y qué propiedad garantiza en un grafo sin pesos?*
- A) Una pila (LIFO); garantiza que se visiten primero los nodos con mayor grado de salida.
 - B) Un árbol binario de búsqueda; garantiza que los nodos se visiten en orden alfanumérico.
 - C) Una tabla hash; garantiza que no se procesen nodos duplicados en grafos cíclicos.
 - D) Una cola (FIFO); garantiza que se calcule la distancia más corta a cada nodo alcanzable.




47. *Durante un recorrido de Búsqueda en Profundidad (DFS), ¿qué indica que el algoritmo ha encontrado una arista de retroceso (back edge) en un grafo dirigido?*
- A) Que la arista conecta un nodo con un descendiente en el árbol DFS.
 - B) Que la arista conecta dos nodos que pertenecen a diferentes árboles del bosque DFS.
 - C) Que la arista conecta un nodo con un antecesor en el árbol DFS.
 - D) Que el nodo destino de la arista ya ha sido finalizado.
48. *¿Qué fenómeno aparece cuando una señal contiene componentes por encima de la frecuencia de Nyquist y se muestrea sin suficiente tasa?*
- A) Fuga espectral.
 - B) Ruido de cuantización.
 - C) Efecto Gibbs.
 - D) Aliasing.
49. *¿Cuál es el primer paso característico del protocolo Contract Net (CNP) cuando un agente gestor desea delegar una tarea?*
- A) Difundir un call for proposals (cfp) a los agentes contratistas potenciales.
 - B) Solicitar al servicio de directorio (DF) que asigne automáticamente la tarea al agente más adecuado.
 - C) Enviar un mensaje inform con la especificación completa de la tarea y esperar confirmación.
 - D) Iniciar una negociación bilateral con cada agente contratista mediante mensajes propose alternados.
50. *¿Qué papel desempeña una función de utilidad en una negociación automatizada entre agentes?*
- A) Define el protocolo de turnos y tiempos límite que estructura el proceso de negociación.
 - B) Establece el espacio de acuerdos factibles eliminando las combinaciones que violan las restricciones del dominio.
 - C) Proporciona al agente un conjunto de funciones auxiliares para gestionar el intercambio de mensajes durante la negociación.
 - D) Codifica las preferencias del agente para comparar ofertas o acuerdos posibles.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 15 de 28</p>
---	---	--

51. *En regresión logística, la variable dependiente es:*
- A) Continua.
 - B) Categórica.
 - C) Exclusivamente ordinal.
 - D) Continúa transformada.
52. *En regresión lineal, el método de mínimos cuadrados ordinarios (Ordinary Least Squares, OLS) estima los parámetros minimizando:*
- A) El error cuadrático total.
 - B) El error absoluto.
 - C) La varianza.
 - D) La log-verosimilitud.
53. *¿Cuál es el principal criterio que diferencia la segmentación por umbralización (thresholding) de la segmentación por crecimiento de regiones (region growing)?*
- A) La umbralización asigna píxeles a regiones basándose en la similitud con píxeles vecinos; el crecimiento de regiones lo hace exclusivamente en función de su valor de intensidad global.
 - B) La umbralización clasifica los píxeles exclusivamente en función de su valor de intensidad; el crecimiento de regiones incorpora además la conectividad espacial entre píxeles vecinos.
 - C) La umbralización requiere definir semillas iniciales; el crecimiento de regiones opera de forma completamente automática sin parámetros.
 - D) Ambos métodos son equivalentes cuando el histograma de la imagen es bimodal.
54. *¿Cuáles son los principios FAIR, formulados en 2016, que deben guiar la gestión de datos científicos en abierto?*
- A) Free, Accessible, Interoperable, Reusable.
 - B) Findable, Accessible, Interoperable, Reusable.
 - C) Findable, Accurate, Integrated, Reproducible.
 - D) Free, Accurate, Interoperable, Reproducible.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 16 de 28</p>
---	---	--

55. *¿Qué establece el mandato de acceso abierto de la Ley de la Ciencia española respecto a las publicaciones financiadas con fondos públicos?*
- A) Recomienda voluntariamente depositar las publicaciones en repositorios abiertos en un plazo máximo de 24 meses tras su publicación.
 - B) Obliga a pagar tasas APC en revistas Gold Open Access para garantizar el acceso inmediato a todas las publicaciones financiadas públicamente.
 - C) Prohíbe la publicación en revistas de suscripción para proyectos financiados con fondos públicos, salvo autorización expresa del organismo financiador.
 - D) Obliga al depósito de la versión final aceptada en repositorios de acceso abierto de forma simultánea a la fecha de publicación, sin periodos de embargo.
56. *En un algoritmo de backtracking para resolver un problema combinatorio, ¿cuál es la función de la técnica denominada poda (pruning)?*
- A) Reordenar los elementos del conjunto de entrada para que el algoritmo termine en tiempo lineal.
 - B) Eliminar la necesidad de utilizar una pila de llamadas recursivas.
 - C) Abandonar la exploración de una rama del árbol de búsqueda que no conduce a una solución válida.
 - D) Seleccionar aleatoriamente el siguiente nodo a explorar para evitar caer en mínimos locales.
57. *¿Cuáles son los dos ingredientes fundamentales que debe tener un problema de optimización para que se le pueda aplicar la técnica de programación dinámica de forma efectiva?*
- A) Recursividad simple y ausencia de ciclos en el espacio de estados.
 - B) Subestructura óptima y subproblemas superpuestos.
 - C) Capacidad de ser resuelto mediante un algoritmo voraz (greedy) y linealidad.
 - D) Memoria limitada y una función de coste estrictamente creciente.
58. *¿Qué hipótesis estructural caracteriza al clasificador Naive Bayes aplicado a texto?*
- A) Que cada término solo puede pertenecer a una clase del vocabulario.
 - B) Que la frontera de decisión entre clases es siempre lineal y sin sesgo.
 - C) Que la presencia o frecuencia de los términos de un documento son condicionalmente independientes entre sí, dada la clase.
 - D) Que la distribución conjunta de los términos se estima maximizando la verosimilitud sobre todo el corpus sin distinguir clases.

	Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025 Primer Ejercicio	Fecha: 10/05/2026 Página: 17 de 28
---	--	--

59. *¿Qué efecto tiene TF-IDF sobre términos muy frecuentes en casi todos los documentos del corpus?*

- A) Reduce su peso discriminativo al aplicar el logaritmo del inverso de la frecuencia de documento.
- B) Le asigna un peso proporcional a su frecuencia bruta en cada documento, sin penalización.
- C) Los agrupa en un único token comodín para reducir la dimensionalidad del vocabulario.
- D) Mantiene su peso intacto, pero los excluye del cálculo de similitud coseno entre documentos.

60. *El método de Monte Carlo es una técnica computacional que se basa en:*

- A) Optimización determinista.
- B) Ajuste de modelos de regresión.
- C) Clasificación supervisada.
- D) Simulación mediante muestreo aleatorio.

61. *¿Cuál es una limitación habitual de los métodos de Monte Carlo?*

- A) No pueden aplicarse a problemas de integración.
- B) Presentan alta varianza en sus estimaciones.
- C) Son completamente deterministas.
- D) No utilizan probabilidades.

62. *En el análisis de datos de simulación ambiental, ¿qué mide el estadístico RMSE (Root Mean Square Error)?*


- A) La correlación lineal entre las series temporales simuladas y las observadas en un dominio espacial.
- B) La proporción de varianza de los datos observados explicada por el modelo de simulación.
- C) La dispersión media del ensamble de modelos respecto a la mediana de las simulaciones.
- D) La desviación cuadrática media entre los valores simulados y los observados, expresada en las mismas unidades que la variable analizada.



63. ***Si el coste de no detectar una avería es muy alto, ¿qué criterio suele ganar importancia al evaluar un modelo de predicción de fallos?***
- A) Maximizar la precisión (precision) de la clase de fallo para evitar que los técnicos pierdan tiempo con falsas alarmas.
 - B) Optimizar la exactitud global (accuracy), ya que refleja el rendimiento general del modelo sobre todas las clases.
 - C) Priorizar el recall de la clase de fallo, aunque aumenten algunos falsos positivos.
 - D) Minimizar el tiempo de inferencia del modelo para poder ejecutar predicciones con la mayor frecuencia posible.
64. ***En simulaciones sociales y de votación, ¿qué establece el teorema de imposibilidad de Arrow?***
- A) Que bajo condiciones de dictadura es posible alcanzar siempre una función de bienestar social que sea totalmente justa.
 - B) Que en sistemas con solo dos alternativas de elección, es matemáticamente imposible alcanzar un consenso de mayoría.
 - C) Que ningún sistema de votación por rangos de preferencia satisface ciertos criterios de equidad con tres o más alternativas.
 - D) Que los agentes en un sistema multiagente siempre optarán por una votación honesta para maximizar el beneficio común.
65. ***Dentro de la robótica y la teoría de juegos, ¿qué define a un equilibrio de Nash en un sistema de múltiples agentes?***
- A) Un estado donde ningún agente puede obtener un mayor beneficio cambiando su estrategia unilateralmente.
 - B) Una situación donde todos los agentes cooperan para maximizar el beneficio global del sistema.
 - C) El estado en el que todos los robots se mueven a la misma velocidad para evitar colisiones.
 - D) Un protocolo donde el agente con mayor capacidad de cómputo toma todas las decisiones.
66. ***¿Cuál es la unidad básica de representación de información en RDF?***
- A) Un documento XML con un nodo raíz.
 - B) Una tripleta compuesta por sujeto, predicado y objeto.
 - C) Una tabla relacional con clave primaria.
 - D) Un par clave-valor codificado en JSON.



67. ***¿Qué papel desempeña la negociación de contenido (content negotiation) en un entorno Linked Data?***
- A) Obliga a crear una URI distinta para cada formato de serialización.
 - B) Sustituye el uso de RDF por HTML cuando hay agentes automáticos.
 - C) Redirige al cliente a un endpoint SPARQL donde puede formular consultas sobre el recurso solicitado.
 - D) Permite devolver distintas representaciones del mismo recurso según el formato solicitado por el cliente.
68. ***JADE (Java Agent DEvelopment Framework) es una plataforma utilizada principalmente para:***
- A) Simulación física.
 - B) Comunicación y desarrollo de agentes software.
 - C) Clustering.
 - D) Análisis de componentes principales.
69. ***MASON es una plataforma utilizada principalmente para:***
- A) Bases de datos.
 - B) Simulación multiagente.
 - C) Procesamiento de lenguaje natural.
 - D) Redes neuronales.
70. ***¿Cuál es la diferencia fundamental entre la licencia CC BY y la licencia CC0 en el contexto de datos científicos?***
- A) CC0 exige atribución al autor original; CC BY permite la renuncia completa a los derechos de autor.
 - B) CC BY exige atribución al autor original como única condición de uso; CC0 implica la renuncia completa a todos los derechos, incluyendo el de atribución.
 - C) Ambas licencias son equivalentes pero CC0 está restringida a uso no comercial.
 - D) CC BY solo es aplicable a publicaciones científicas; CC0 se aplica exclusivamente a conjuntos de datos.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 20 de 28</p>
---	---	--


71. *¿Qué implica la cláusula **ShareAlike (SA)** presente en algunas licencias **Creative Commons** cuando se aplica a un dataset?*
- A) Que el dataset solo puede compartirse de forma privada entre investigadores acreditados.
 - B) Que cualquier uso comercial del dataset debe ser previamente autorizado por el autor.
 - C) Que el dataset no puede ser modificado ni adaptado bajo ninguna circunstancia.
 - D) Que cualquier obra derivada del dataset debe distribuirse bajo la misma licencia o una compatible con el original.
72. *¿Cuál es la función principal de la métrica **TF-IDF (Term Frequency-Inverse Document Frequency)** en la recuperación de información y minería de texto?*
- A) Medir la similitud gramatical entre dos oraciones complejas.
 - B) Evaluar la importancia de una palabra en un documento en relación con un corpus.
 - C) Calcular la probabilidad de que una palabra sea seguida por otra en un modelo de lenguaje de n-gramas.
 - D) Normalizar la longitud de los documentos para que todos tengan el mismo número de tokens.
73. *En el contexto de la extracción de tópicos, ¿qué representa el algoritmo **Latent Dirichlet Allocation (LDA)**?*
- A) Un clasificador supervisado que asigna una etiqueta única a cada documento basada en su frecuencia de palabras.
 - B) Una red neuronal recurrente que predice la siguiente palabra de un texto basándose en el contexto previo.
 - C) Un modelo probabilístico que asume que cada documento es una mezcla de varios temas y cada tema es una distribución de palabras.
 - D) Una técnica de reducción de dimensiones que utiliza únicamente la descomposición en valores singulares.
74. *En aprendizaje automático, ¿cuál de las siguientes afirmaciones es correcta?*
- A) Añadir más variables siempre mejora el modelo.
 - B) El overfitting implica aprender únicamente patrones reales.
 - C) La regularización puede mejorar la capacidad de generalización.
 - D) La validación del modelo no es necesaria.



75. *¿Para qué tipo de datos está específicamente diseñado el modelo LASSO (Least Absolute Shrinkage and Selection Operator)?*
- A) Clustering de datos no etiquetados mediante medidas de distancia.
 - B) Regularización en modelos de regresión para reducir el sobreajuste y seleccionar variables.
 - C) Clasificación supervisada mediante máquinas de vectores de soporte.
 - D) Optimización sin restricciones de funciones no convexas.
76. *¿Cuál es la función principal de Zenodo en el ecosistema de datos científicos abiertos?*
- A) Actuar como motor de búsqueda bibliográfica para artículos científicos de acceso abierto.
 - B) Gestionar el proceso de revisión por pares de publicaciones en revistas Open Access.
 - C) Proporcionar un repositorio generalista de acceso abierto para depositar y compartir datos, software y publicaciones científicas, asignando un DOI a cada depósito.
 - D) Centralizar la financiación europea de proyectos de investigación en datos abiertos.
77. *¿Qué identificador persistente se utiliza de forma estándar en los principales portales de datos científicos abiertos (Zenodo, Figshare, Dryad) para garantizar la citabilidad y localización permanente de un dataset?*
- A) ORCID (Open Researcher and Contributor ID).
 - B) ISBN (International Standard Book Number).
 - C) Handle System (HDL).
 - D) DOI (Digital Object Identifier).
78. *¿Cuál es la función principal del RDF Schema en la creación de metadatos estructurados para la Web Semántica?*
- A) Proporcionar un formato de compresión para reducir el tamaño de los archivos XML.
 - B) Definir una sintaxis obligatoria basada exclusivamente en el uso de paréntesis.
 - C) Reemplazar el uso de URIs por etiquetas de texto plano para facilitar la lectura humana.
 - D) Ofrecer un vocabulario para describir grupos de recursos y las relaciones entre ellos.



79. *En el estándar Simple Knowledge Organization System (SKOS), comúnmente usado en repositorios científicos, ¿qué propiedad se utiliza para indicar que un concepto es más genérico que otro?*
- A) skos:related
 - B) skos:broader
 - C) skos:narrower
 - D) skos:altLabel
80. *¿Qué diferencia fundamental existe entre `rdfs:Class` y `owl:Class`?*
- A) `rdfs:Class` solo admite instancias con una única propiedad, mientras que `owl:Class` permite múltiples propiedades por instancia.
 - B) `rdfs:Class` requiere que todas las instancias estén explícitamente declaradas; `owl:Class` permite inferir pertenencia mediante un razonador.
 - C) `rdfs:Class` permite definir jerarquías con `rdfs:subClassOf` pero no restricciones sobre propiedades; `owl:Class` añade constructores lógicos como intersección, unión y complemento.
 - D) No existe diferencia operativa; `owl:Class` es simplemente un alias de `rdfs:Class` mantenido por compatibilidad.
81. *En el modelado de ontologías, ¿qué función cumple un razonador (reasoner) como Pellet o HermiT?*
- A) Traducir la ontología OWL a un esquema relacional para su almacenamiento en bases de datos SQL.
 - B) Generar automáticamente una interfaz web de consulta a partir de las clases y propiedades declaradas.
 - C) Reescribir consultas SPARQL en planes de ejecución optimizados para triple stores distribuidos.
 - D) Inferir conocimiento implícito, verificar la consistencia y clasificar la jerarquía de clases.
82. *¿Para qué tipo de datos está específicamente diseñado el modelo SARIMA (Seasonal AutoRegressive Integrated Moving Average)?*
- A) Series temporales estacionarias sin patrones cíclicos, donde solo se requiere modelar la autocorrelación a corto plazo.
 - B) Datos de panel con múltiples variables explicativas y efectos fijos entre grupos de observaciones.
 - C) Series temporales que presentan patrones que se repiten con una frecuencia regular (diaria, mensual, anual).
 - D) Variables continuas con distribución no gaussiana que requieren transformaciones logarítmicas previas al ajuste.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 23 de 28</p>
---	---	--

83. *La ley de Amdahl establece que:*

- A) El speedup puede ser infinito.
- B) La parte secuencial limita la aceleración total del programa.
- C) Todo problema es completamente paralelizable.
- D) El rendimiento es independiente del hardware.

84. *En el contexto del versionado de datasets científicos, ¿qué esquema de numeración basado en la convención Semantic Versioning se recomienda para distinguir cambios mayores que alteran la estructura del dataset de cambios menores que solo corrigen errores puntuales?*


- A) Versionado por fecha (YYYYMMDD), que identifica unívocamente cada versión por su fecha de creación.
- B) Versionado semántico (MAJOR.MINOR.PATCH), donde cambios en MAJOR indican alteraciones estructurales y cambios en PATCH indican correcciones menores.
- C) Versionado alfabético secuencial (v.A, v.B, v.C), estándar en repositorios institucionales europeos.
- D) Versionado por hash MD5 del contenido del archivo, que garantiza la integridad pero no la interpretabilidad del cambio.

85. *Dentro del proceso de modelado, ¿cuál es el propósito fundamental del análisis de sensibilidad?*


- A) Determinar si el modelo es capaz de resolver cualquier problema narrativo.
- B) Eliminar todas las variables innecesarias antes de formular las ecuaciones.
- C) Investigar cómo afectan las variaciones en los datos a la confianza en la respuesta obtenida.
- D) Traducir la solución matemática a un lenguaje no especializado para el cliente.

86. *En una consulta SPARQL, ¿qué cláusula se utiliza para devolver resultados incluso si un patrón de datos específico no existe en el grafo?*


- A) OPTIONAL
- B) UNION
- C) FILTER
- D) SELECT

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 24 de 28</p>
---	---	--


87. *¿Cuál de las siguientes afirmaciones define mejor el propósito de Protégé en el ecosistema de la Web Semántica?*
- A) Es un motor de búsqueda que indexa archivos RDF en la web pública.
 - B) Es un lenguaje de programación para desarrollar aplicaciones cliente-servidor.
 - C) Es un entorno de edición y modelado para crear y gestionar ontologías en OWL.
 - D) Es un protocolo de red que sustituye a HTTP para la transmisión de datos conectados.
88. *¿Qué problema resuelven los mappings de ontologías (ontology mapping) en la interoperabilidad semántica entre repositorios científicos?*
- A) La fusión automática de dos ontologías en una sola ontología canónica que reemplaza a ambas.
 - B) La traducción sintáctica entre formatos de serialización como RDF/XML, Turtle y JSON-LD.
 - C) La alineación de conceptos equivalentes o relacionados entre ontologías distintas.
 - D) La validación de que cada instancia de una ontología cumple las restricciones de cardinalidad declaradas.
89. *¿Cuál es la finalidad principal de OAI-PMH?*
- A) Sincronizar bases de datos relacionales mediante transacciones ACID.
 - B) Editar en remoto registros bibliográficos con bloqueo persistente.
 - C) Recolectar metadatos de repositorios para construir servicios agregados.
 - D) Transferir siempre el contenido binario completo de cada objeto digital, incluidos adjuntos y miniaturas.
90. *En una API REST, ¿qué significa que las interacciones sean stateless?*
- A) Que el servidor no puede usar cachés ni balanceadores de carga.
 - B) Que las respuestas deben devolverse siempre en JSON.
 - C) Que los recursos publicados son inmutables y no pueden modificarse tras su creación.
 - D) Que cada petición debe incluir la información necesaria para ser entendida sin depender del estado de sesión guardado en el servidor.
91. *En bioinformática, el escenario $p \gg n$ (muchas variables y pocas muestras) implica:*
- A) Baja dimensionalidad.
 - B) Gran número de muestras.
 - C) Alto riesgo de sobreajuste.
 - D) Datos deterministas.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 25 de 28</p>
---	---	--

92. *En el algoritmo de vecinos más cercanos (K-Nearest Neighbors, KNN), el aumento de dimensionalidad provoca que:*
- A) Mejore el rendimiento del modelo.
 - B) Las distancias entre observaciones sean menos informativas.
 - C) Se reduzca el ruido.
 - D) Disminuya la varianza.
93. *¿Cuál de los siguientes elementos NO forma parte del contenido esencial de un Data Management Plan según las directrices de Horizonte Europa?*
- A) La descripción de las medidas de seguridad y control de acceso aplicadas a datos sensibles o confidenciales durante el proyecto.
 - B) La identificación de los conjuntos de datos que quedarán excluidos del acceso abierto y la justificación de dicha exclusión.
 - C) La especificación del modelo de gobernanza interna del consorcio para la toma de decisiones sobre autoría y orden de firma en publicaciones científicas.
 - D) La indicación de los repositorios de datos previstos para el depósito y preservación a largo plazo de los datasets generados.
94. *¿Cuál es la función del paquete de información de archivo (AIP, Archival Information Package) en el modelo de referencia OAIS (Open Archival Information System)?*
- A) Contener los datos originales enviados por el productor antes de su validación por el repositorio.
 - B) Exportar metadatos hacia agregadores externos como OpenAIRE para su indexación.
 - C) Agrupar datos y metadatos necesarios para la preservación y acceso a largo plazo dentro del repositorio.
 - D) Registrar las condiciones de licencia y acceso de cada objeto digital depositado.
95. *¿Cuál es la principal ventaja del uso de perfiles Application-Level Profile Semantics (ALPS) en la integración de APIs científicas?*
- A) Cifrar los datos de la investigación para que solo sean legibles por el autor.
 - B) Definir la semántica de los datos de forma independiente al formato de serialización (JSON, XML).
 - C) Sustituir el protocolo HTTP por uno más rápido basado en UDP.
 - D) Eliminar la necesidad de utilizar métodos de autenticación como OAuth.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025 Primer Ejercicio</p>	<p>Fecha: 10/05/2026 Página: 26 de 28</p>
---	--	---

96. *Al seleccionar un formato para la preservación a largo plazo, ¿por qué se prefieren los formatos de especificación abierta frente a los propietarios?*
- A) Porque los formatos abiertos requieren por norma un menor espacio de almacenamiento físico.
 - B) Porque los formatos propietarios tienen una resolución de imagen inferior.
 - C) Porque su documentación pública permite crear software compatible de forma independiente.
 - D) Porque los formatos abiertos no permiten la inclusión de metadatos estructurados.
97. *Al ingerir datos de múltiples fuentes heterogéneas, ¿qué técnica se aplica cuando el mismo registro aparece con valores distintos en fuentes diferentes?*
- A) Deduplicación exacta basada en hash, que elimina automáticamente los registros con menor marca temporal.
 - B) Resolución de entidades con estrategias de fusión como prioridad por fuente, votación o selección del valor más reciente.
 - C) Normalización de esquema, que unifica los tipos de datos y con ello resuelve las discrepancias de valores.
 - D) Particionamiento horizontal de las tablas, asignando cada fuente a un fragmento independiente sin reconciliación.
98. *En un modelo de predicción de vida útil restante basado en la condición, ¿qué tipo de datos de entrada se utiliza de forma principal y dinámica?*
- A) Resultados de inspecciones visuales periódicas, codificados como categorías ordinales por técnicos de planta.
 - B) Registros de intervenciones correctivas previas junto con el coste asociado a cada reparación.
 - C) Series temporales de sensores operacionales, como vibración, temperatura, y presión.
 - D) Parámetros estáticos de diseño del equipo extraídos de su ficha técnica y hoja de materiales.
99. *En una aplicación web de visualización interactiva de datos, ¿qué técnica permite al servidor enviar actualizaciones al cliente en tiempo real sin que el cliente las solicite repetidamente?*
- A) Polling activo mediante peticiones GET periódicas.
 - B) WebSockets.
 - C) Caché del navegador con expiración prolongada.
 - D) Redireccionamiento HTTP 301 hacia un endpoint de datos actualizados.

	<p>Proceso selectivo por el sistema de acceso libre para ingreso en la Escala de Tecnólogos de los Organismos Públicos de Investigación, convocado por resolución de 22 de diciembre de 2025 (BOE N°314 30 de diciembre) – OEP 2023-2024-2025</p> <p>Primer Ejercicio</p>	<p>Fecha: 10/05/2026</p> <p>Página: 27 de 28</p>
---	---	--

100. *¿Qué es la validación cruzada temporal (temporal cross-validation) en el contexto de modelos de simulación ambiental con datos de series temporales?*

- A) La división del dominio espacial en subregiones independientes para evaluar la capacidad de generalización del modelo.
- B) La técnica de reservar un período temporal completo como conjunto de test, entrenando el modelo exclusivamente con datos anteriores a dicho período.
- C) El proceso de recalibrar los parámetros del modelo cada vez que se incorporan nuevas observaciones al sistema.
- D) La comparación sistemática entre las salidas de distintas versiones del mismo modelo a lo largo del tiempo.

PREGUNTAS DE RESERVA

101. *¿Cuál es el propósito primordial de aplicar el Análisis de Componentes Principales (PCA) a un conjunto de espectros estelares o de galaxias?*

- A) Aumentar el ruido de la señal para probar la robustez de los algoritmos.
- B) Reducir la dimensionalidad de los datos.
- C) Determinar la edad exacta de cada estrella basándose en su composición química.
- D) Predecir la posición futura de los objetos en la esfera celeste.

102. *En las máquinas de vectores de soporte (Support Vector Machines, SVM), el objetivo principal es:*

- A) Minimizar el error cuadrático.
- B) Maximizar el margen entre clases.
- C) Reducir la dimensionalidad.
- D) Clasificar sin datos etiquetados.

103. *En un modelo de regresión aplicado a datos socioeconómicos (por ejemplo, renta y nivel educativo), ¿qué indica la presencia de multicolinealidad entre variables explicativas?*

- A) La existencia de una relación lineal significativa entre dos o más variables independientes.
- B) Que la variable dependiente no depende de los factores socioeconómicos.
- C) Que los errores del modelo tienen varianza constante.
- D) Que el modelo predice correctamente todos los casos observados.



104. En el contexto de la simulación ambiental, ¿qué ventaja principal ofrece un ensamble de modelos (model ensemble) frente a una única simulación determinista?

- A) Reduce el coste computacional al distribuir la carga entre múltiples modelos más simples.
- B) Elimina completamente la incertidumbre asociada a las condiciones iniciales del sistema simulado.
- C) Permite cuantificar la incertidumbre de las predicciones mediante la dispersión estadística de los resultados del ensamble.
- D) Garantiza que al menos una de las simulaciones converge a la solución analítica exacta del sistema.

105. Al trabajar con APIs REST, ¿qué significa que el método GET sea seguro (safe) según la especificación HTTP?

- A) Que los datos viajan obligatoriamente encriptados mediante HTTPS.
- B) Que el servidor garantiza que el usuario tiene permisos de administrador.
- C) Que su ejecución no produce cambios en el estado del recurso.
- D) Que la petición siempre devuelve un código de estado 200 OK.

*****FIN DEL CUESTIONARIO*****