

PROCESO SELECTIVO PARA INGRESO, POR EL SISTEMA GENERAL DE ACCESO LIBRE, EN LA ESCALA DE TÉCNICOS ESPECIALIZADOS DE LOS ORGANISMOS PÚBLICOS DE INVESTIGACIÓN. CONVOCADO POR RESOLUCION DE 21 DE FEBRERO DE 2023 (BOE Nº 47 DE 24 DE FEBRERO DE 2023)

ÁREA GLOBAL:

A6-CIENCIA DE DATOS

ESPECIALIDAD:

D3-MODELADO Y ANÁLISIS DE DATOS

PRIMER EJERCICIO

- No abra este cuestionario hasta que se le indique.
- Cada pregunta tiene 4 opciones, siendo sólo una correcta. Una contestación incorrecta se penaliza con el 25% del valor de una correcta. Las preguntas en blanco no se penalizan.
- Marque las respuestas en la hoja de examen usando bolígrafo azul o negro, siguiendo las instrucciones contenidas en ésta sobre cómo hacer correcciones, anulaciones, etc.
- El cuestionario contiene 100 preguntas: 30 de las materias comunes, 25 del área global y 45 de la especialidad. El tiempo de realización del ejercicio es de 90 minutos.
- Puede usar este cuadernillo para hacer anotaciones, marcas, etc. Nunca las haga en la hoja de examen.
- En la hoja de examen no olvide incluir apellidos, nombre, DNI y firma.
- No separe las hojas de examen. Eso lo hará el tribunal al final del ejercicio.
- No podrá ausentarse del aula durante los primeros 10 minutos ni cuando falten 10 para el final.
- Podrá llevarse este cuadernillo sólo al final del ejercicio. Si termina el examen antes de los 90 minutos, puede poner su nombre en el cuadernillo y recogerlo al final.

- 1) Son causas de pérdida de la condición de funcionario de carrera
- A) la renuncia a la condición de funcionario
 - B) la jubilación parcial
 - C) la ausencia prolongada por enfermedad
 - C) cualquier sanción o condena administrativa, penal o criminal
- 2) ¿Cuál es el fin principal de la Ley de la Ciencia?
- A) contribuir al desarrollo económico sostenible y al bienestar social
 - B) definir la organización de los organismos investigadores del estado
 - C) definir la carrera científica
 - D) armonizar la legislación española con la europea
- 3) Se considera una presencia y composición equilibrada en términos de igualdad de género
- A) una composición en que las personas de cada sexo no superen el 40% ni sea inferior al 60%
 - B) una composición en que las personas de cada sexo no superen el 60% ni sea inferior al 40%
 - C) sólo cuando la composición sea del 50% de cada sexo
 - D) una composición paritaria permitiéndose un margen del 30%
- 4) NO es responsabilidad del personal investigador
- A) obtener el máximo de publicaciones posible
 - B) rendir cuentas sobre su trabajo
 - C) adoptar medidas de protección de datos y confidencialidad
 - D) procurar que su labor sea relevante para la sociedad
- 5) La antigüedad mínima para poder ser declarado en situación de excedencia temporal para la incorporación a otros agentes públicos del Sistema Español de Ciencia, Tecnología e Innovación será de
- A) dos años
 - B) cuatro años
 - C) cinco años
 - D) diez años
- 6) Durante el periodo de excedencia temporal, el personal en esta situación
- A) seguirá percibiendo retribuciones de su puesto de procedencia
 - B) tendrá derecho a la reserva del puesto de trabajo
 - C) no se contará ese periodo a efectos de antigüedad
 - D) no se contará ese periodo a efectos de consolidación del grado personal
- 7) La duración acumulada de las autorizaciones concedidas a cada persona para estancias formativas
- A) no puede superar los diez años
 - B) no puede superar los dos años cada cinco años
 - C) no será inferior a un año cada cinco años
 - D) no será inferior a un mes

8) La prestación de servicios por parte del personal de investigación en sociedades mercantiles creadas o participadas por la entidad para que dicho personal preste servicios

- A) es incompatible con su actividad profesional
- B) se considera una actividad de interés general y por tanto deberá ser necesariamente a tiempo completo
- C) no necesita justificación previa
- D) no podrá modificar la jornada ni el horario del puesto de trabajo inicial del interesado

9) ¿Cuál de las siguientes NO es una modalidad específica de contrato de trabajo de personal investigador?

- A) Contrato predoctoral
- B) Contrato de investigador/a distinguido/a
- C) Contrato de actividades científico-técnicas
- D) Contrato por obra y servicios

10) El contrato de acceso de personal investigador doctor

- A) será de duración indefinida y a tiempo parcial
- B) será de duración determinada y a tiempo completo
- C) será como máximo de tres años
- D) será improrrogable

11) El personal con contrato de acceso de personal investigador doctor

- A) puede dedicar todo su tiempo a realizar actividad docente
- B) no podrá realizar actividad docente, solo investigadora
- C) puede realizar hasta cien horas de actividad docente si lo desea
- D) puede realizar hasta cien horas de actividad docente con la aprobación de la entidad en la que presta servicios

12) El certificado R3 como investigador/a establecido/a

- A) carece de efecto en los procesos selectivos de personal de nuevo ingreso
- B) no se considerará para las escalas de científico titular
- C) no es válido para la escala de profesor titular
- D) permite acceder a la escala de personal laboral fijo

13) En la Oferta de Empleo Público se reservará

- A) al menos un 15% de plazas para doctores con certificado R3 o evaluación equivalente
- B) menos de un 15% de plazas para doctores con certificado R3 o equivalente
- C) la totalidad de plazas para doctores sin certificado R3 o equivalente
- D) la totalidad de plazas para doctores con certificado R3 o equivalente

14) ¿Cuál NO es una escala del personal funcionario de carrera al servicio de los Organismos Públicos de Investigación?

- A) Escala de Profesorado de Investigación de Organismos Públicos de Investigación.
- B) Escala de Personal Investigador Científico de Organismos Públicos de Investigación.
- C) Escala de Personal Científico Titular de Organismos Públicos de Investigación.
- D) Escala de Personal Científico en Formación de Organismos Públicos de Investigación

15) El personal investigador funcionario de carrera consolidará el grado personal

- A) tras la consecución de 10 publicaciones en revistas internacionales
- B) tras la obtención de dos proyectos financiados como investigador principal
- C) con arreglo a lo dispuesto en la normativa general de la función pública
- D) tras la publicación de su primera patente

16) El disfrute de los beneficios derivados de la actividad de I+D+I

- A) corresponde exclusivamente a la Administración Pública como empleador
- B) corresponde exclusivamente a la entidad financiadora del proyecto
- C) el investigador tiene derecho a participar en los beneficios de la actividad en que haya participado sin que se considere retribución o salario
- D) Ninguna de las anteriores es cierta

17) Durante el periodo de excedencia temporal, el personal de investigación

- A) seguirá percibiendo retribuciones de su puesto de origen
- B) tendrá derecho a la reserva del puesto de trabajo
- C) perderá el derecho a la evaluación de la actividad investigadora y de los méritos investigadores y técnicos
- D) no estará obligado a proteger el conocimiento de los equipos de investigación conforme a la normativa aplicable

18) El personal técnico que preste servicios en Organismos Públicos de Investigación de la Administración General del Estado tendrá el deber de

- A) Poner en conocimiento de las entidades para las que presta servicios todos los hallazgos, descubrimientos y resultados susceptibles de protección jurídica
- B) colaborar en el máximo número de publicaciones posible
- C) definir la duración de su jornada de trabajo en función de las necesidades del servicio
- D) omitir la denominación de las entidades para las que presta servicios en la realización de su actividad

19) La promoción interna entre las escalas técnicas y las científicas

- A) no es posible
- B) es automática al alcanzar un número predeterminado de patentes
- C) se podrán prever procesos de promoción interna entre las escalas del mismo subgrupo
- D) es automática si se tiene un número predeterminado de publicaciones científicas

- 20)** La vigencia de los convenios de formación del personal científico y técnico de Organismos Públicos de Investigación será
- A) inferior a cinco años de duración inicial
 - B) de cinco años improrrogables
 - C) de cinco años prorrogables indefinidamente
 - D) podrá ser indefinida en función del cumplimiento de los fines del convenio
- 21)** Los Convenios de creación de infraestructuras científicas tendrán una vigencia
- A) inferior a cinco años de duración inicial
 - B) de cinco años improrrogables
 - C) de cinco años prorrogables indefinidamente
 - D) podrá ser indefinida en función del cumplimiento de los fines del convenio
- 22)** La participación de Organismos Públicos de Investigación en sociedades mercantiles
- A) podrá ser aprobada discrecionalmente por dichos organismos
 - B) deberá ser autorizada por el Consejo de Ministros
 - C) deberá ser decidida tras un concurso público abierto e igualitario
 - D) Ninguna de las anteriores es correcta
- 23)** ¿Quién ejerce la función ejecutiva según la Constitución Española?
- A) El Gobierno
 - B) El Senado
 - C) El Congreso de los Diputados
 - D) El Rey
- 24)** Las resoluciones de los Jueces
- A) siempre son de obligado cumplimiento
 - B) son de obligado cumplimiento si son firmes
 - C) solo pueden obligar a prestar colaboración a los funcionarios del Ministerio de Justicia
 - D) no dan derecho a indemnización por daños debidos a error judicial
- 25)** Según la Constitución, ¿a quién corresponde la elaboración de los Presupuestos Generales del Estado?
- A) Al Gobierno
 - B) A las Cortes Generales
 - C) Al Ministerio de Economía y Hacienda
 - D) Al Senado
- 26)** ¿A quién corresponde la competencia exclusiva sobre Legislación sobre propiedad intelectual e industrial?
- A) Al Estado
 - B) A las Comunidades Autónomas
 - C) Al Senado
 - D) A la Unión Europea

- 27)** Son causas de pérdida de la condición de funcionario de carrera
- A) la renuncia a la condición de funcionario
 - B) la jubilación parcial
 - C) la ausencia prolongada por enfermedad
 - D) cualquier sanción o condena administrativa, penal o criminal
- 28)** El silencio administrativo será desestimatorio
- A) nunca
 - B) en todos los procedimientos relativos al ejercicio del derecho de petición
 - C) en los procedimientos de impugnación de actos y disposiciones iniciados a solicitud de los interesados
 - D) siempre
- 29)** El procedimiento administrativo común comprende las siguientes fases:
- A) iniciación, ordenación, instrucción y finalización
 - B) iniciación, instrucción, resolución y ejecución
 - C) iniciación, recusación, evaluación y resolución
 - D) instrucción, evaluación y resolución
- 30)** Las Agencias Estatales
- A) carecen de patrimonio propio
 - B) carecen de facultades para ejercer potestades administrativas
 - C) poseen una personalidad jurídica pública
 - D) carecen de autonomía en su gestión
- 31)** En el contexto de las bases de datos, ¿qué se entiende por normalización?
- A) Eliminar datos redundantes dividiendo la tabla de una manera bien definida para mantener la integridad de los datos y ahorrar gran parte del espacio de almacenamiento.
 - B) Añadir datos redundantes en la tabla para acelerar las consultas complejas y lograr así un mejor rendimiento.
 - C) Re-escalar los datos de entrada para que estén todos en la misma escala y reducir los problemas de redondeo en coma flotante.
 - D) Dividir todos los datos numéricos de la tabla por su módulo para evitar problemas de desbordamiento de memoria.
- 32)** En los sistemas de información geográfica (SIG), una de las ventajas de utilizar un modelo de datos *raster* frente a un modelo vectorial para representar la realidad sería:
- A) Codificación eficiente de la topología y las operaciones espaciales.
 - B) Mayor compatibilidad con entornos de bases de datos relacionales.
 - C) Las operaciones de superposición son muy sencillas.
 - D) Las operaciones de re-escalado y reproyección son más fáciles de ejecutar.

33) ¿Cuál de las siguientes tecnologías de memoria es utilizada en la computación de altas prestaciones para acelerar el acceso a los datos?

- A) Memoria de acceso aleatorio (RAM) convencional.
- B) Memoria de solo lectura (ROM) programable.
- C) Memoria de acceso no volátil (NVRAM).
- D) Memoria caché de nivel 1 (L1).

34) ¿Cuál de las siguientes opciones es una característica de la metodología de desarrollo “ágil”?

- A) El equipo trabaja en un plan detallado desde el principio.
- B) El equipo trabaja en iteraciones cortas y regulares.
- C) El equipo no tiene contacto con el usuario final.
- D) El equipo no realiza pruebas hasta el final del proyecto.

35) En los lenguajes orientados a objetos, declarar y definir dos métodos con el mismo nombre pero distintos argumentos se denomina:

- A) Polimorfismo
- B) Sobrecarga
- C) Sobrescritura
- D) Herencia

36) ¿Cuál de las siguientes opciones describe mejor lo que es Zenodo?

- A) Una plataforma de almacenamiento en la nube para compartir y archivar datos de investigación.
- B) Un motor de búsqueda para artículos científicos y académicos.
- C) Una red social para científicos y académicos.
- D) Un software para el análisis de datos en el campo de la bioinformática.

37) ¿Cuál es el protocolo más común para iniciar una sesión remota segura mediante línea de comandos en una máquina Linux?

- A) SSH
- B) Telnet
- C) RDP
- D) VPN

38) ¿Cuál de las siguientes opciones corresponde a un protocolo para la recolección e intercambio de metadatos entre repositorios digitales?

- A) IPMI
- B) MPI
- C) AI-OPS
- D) OAI-PMH

39) ¿Qué es la herencia en programación orientada a objetos?

- A) Es la capacidad de una clase de heredar bucles *for*.
- B) Es la capacidad de recibir una herencia y dejar de programar.
- C) Es la capacidad de una clase de heredar sólo los métodos de otra clase.
- D) Es la capacidad de una clase de heredar los atributos y métodos de otra clase y agregar nuevos atributos y métodos.

40) En programación, ¿cuál es la estructura de control que permite repetir un conjunto de instrucciones mientras se cumpla una condición?

- A) if-else
- B) case
- C) while
- D) switch

41) ¿Cuál NO es un término de las “V” del Big Data?

- A) Variedad
- B) Velocidad
- C) Versatilidad
- C) Volumen

42) ¿Qué caracteriza un sistema en tiempo real?

- A) Que está formado por variables de tipo real.
- B) Es un sistema para controlar el tráfico en la red.
- C) Es un sistema que ofrece una respuesta acotada en el tiempo.
- D) Un sistema creado en un computador real, con programas reales.

43) ¿Qué son los discos NVMe?

- A) Discos duros magnéticos de tecnología de litio.
- B) Discos ópticos con velocidad de muestreo de 32 MHz
- C) Discos con “New Velocity Move of Data”
- D) Dispositivos de almacenamiento de memoria no volátil conectados a través de un bus PCI (expres).

44) Cuando decimos que “indexamos un fichero”, ¿Qué estamos haciendo?

- A) Construimos un índice de los datos para lograr un acceso más rápido.
- B) Eliminamos todas las líneas en blanco y pasamos de mayúsculas a minúsculas.
- C) Numeramos las líneas de datos con letras y números, como las matrículas de los coches.
- D) Medimos el índice de compresión de los datos para transmitirlos más rápido.

45) ¿Cuál de los siguientes NO es un protocolo para AAA (“Autorización, Autenticación y Accounting”)

- A) SSHD
- B) DIAMETER
- C) RADIUS
- D) OIDC

46) ¿Cuál de los siguientes lenguajes de programación NO es orientado a objetos?

- A) Python
- B) C
- C) C++
- D) Java

47) En programación orientada a objetos (OOP), ¿qué elemento se utilizaría para representar el “color de un coche”?

- A) un atributo
- B) una clase
- C) un objeto
- D) una instancia

48) ¿Cuál de las siguientes licencias es la recomendada por la UE para publicar código abierto de las administraciones públicas?

- A) EUPL
- B) BSD
- C) MIT
- D) GPL

49) ¿Para qué sirve un DOI?

- A) Para proporcionar una URL más corta y estandarizada
- B) Para generar un código QR que identifique un servidor web
- C) Para identificar un documento en papel en una biblioteca física
- D) Para identificar un documento digital de forma independiente de su localización

50) ¿Qué es la “hiperpaginación”?

- A) una situación en que un exceso de consumo de memoria virtual degrada el rendimiento
- B) una situación en que un exceso de uso de CPU degrada el rendimiento
- C) una situación en que un exceso de tráfico de red degrada el rendimiento
- D) una situación en que una saturación del bus de datos degrada el rendimiento

51) ¿Para qué sirve la memoria virtual?

- A) Para obtener un acceso más rápido a los datos
- B) Para usar dispositivos de tipo disco como si fueran memoria de acceso aleatorio
- C) Para correr máquinas virtuales sin usar memoria de acceso aleatorio
- D) es un sustituto de la memoria de solo lectura

52) ¿Pueden considerarse una infraestructura computacional basada en GPUs como de Altas Prestaciones?:

- A) No, sólo son infraestructuras computacionales de Altas Prestaciones aquellas que tienen un número muy elevado de CPUs.
- B) No, sólo se pueden considerar los centros de supercomputación con titularidad estatal o internacional
- C) Sí, las GPUs son dispositivos especializados que cuentan con una arquitectura altamente paralela por lo que les permite realizar cálculos intensivos de forma eficiente
- D) No, porque las GPUs sólo se aplican a problemas de tratamiento de imágenes y no a toda la variedad de problemas científicos y tecnológicos

- 53) ¿Cuál es la afirmación INCORRECTA sobre los computadores exaescala?
- A) Está previsto que aparezca el primer computador exaescala en 2024.
 - B) La ciencia se beneficiará de los computadores exaescala.
 - C) La tecnología se beneficiará de los computadores exaescala.
 - D) Un computador exaescala es capaz de realizar 10^{18} FLOPS.
- 54) ¿Cuál de las siguientes afirmaciones sobre las GPUs es correcta?
- A) Las GPUs son exclusivamente utilizadas en la industria del entretenimiento para renderizar gráficos en videojuegos.
 - B) Las GPUs sólo se utilizan para acelerar cálculos científicos en supercomputadores.
 - C) Las GPUs son componentes hardware diseñados específicamente para procesamiento gráfico y computación paralela.
 - D) Las GPUs son incompatibles con los sistemas operativos modernos y solo funcionan en versiones antiguas de Windows.
- 55) ¿Cuál de las siguientes NO es una base de datos de aplicación científica?
- A) Protein Data Bank (PDB)
 - B) EGA (European Genome-Phenome Archive)
 - C) PubMed
 - D) GitHub
- 56) ¿Cuáles son los elementos que caracterizan una neurona artificial?
- A) Dendritas, cuerpo, axón.
 - B) Entradas, núcleo, salidas.
 - C) Entradas, pesos sinápticos, regla de propagación, función de activación.
 - D) Entradas, función de activación, salidas.
- 57) ¿Cuándo decimos que hay *overfitting* en el contexto de modelado?
- A) Cuando el modelo se ajusta muy bien a los datos de entrenamiento y no generaliza bien a otros.
 - B) Cuando el modelo no logra reproducir los datos de entrenamiento.
 - C) Cuando el modelo se ajusta muy bien a los datos de entrenamiento y generaliza bien.
 - D) Ninguna de las opciones anteriores.
- 58) ¿Qué es el método Runge-Kutta clásico?
- A) Es un método numérico para calcular puntos críticos de una curva.
 - B) Es un método para resolver ecuaciones diferenciales ordinarias implícito de paso adaptativo.
 - C) Es un método para resolver ecuaciones diferenciales ordinarias explícito.
 - D) Es un método para resolver ecuaciones diferenciales implícito de paso fijo.
- 59) ¿Para qué se utiliza el método de elementos finitos?
- A) Para resolver sistemas de ecuaciones algebraicas.
 - B) Para resolver ecuaciones en derivadas parciales unidimensionales exclusivamente.
 - C) Para resolver ecuaciones en derivadas parciales.
 - D) Es un método de cálculo de series finitas.

- 60)** ¿Cuál de los siguientes métodos NO es un método de integración?
- A) Método de Newton-Raphson.
 - B) Método de Simpson.
 - C) Método de Adams-Bashforth-Moulton.
 - D) Método de Euler.
- 61)** ¿Cómo son las incógnitas en una ecuación diferencial en derivadas parciales?
- A) Siempre estocásticas.
 - B) Dependientes sólo de variables espaciales.
 - C) Siempre definidas positivas.
 - D) Dependientes de diversas variables independientes.
- 62)** En una distribución normal...
- A) la media, la moda y la mediana coinciden.
 - B) sólo la media y la mediana coinciden.
 - C) sólo la moda y la media coinciden.
 - D) la media, la moda y la mediana son distintas.
- 63)** ¿Cuál de las siguientes herramientas NO se usa normalmente para el análisis de series temporales?
- A) Autocorrelación
 - B) Transformada de Fourier
 - C) Análisis espectral
 - D) Flujo óptico
- 64)** Los principales tipos de parámetros usados en estadística descriptiva son...
- A) centralidad, posición y dispersión
 - B) centralidad, límites e intervalo
 - C) forma, intervalo y tendencia
 - D) límite, tendencia y dispersión
- 65)** ¿Cuál de los siguientes NO es un test estadístico?
- A) t-test
 - B) Unit test
 - C) ANOVA
 - D) Chi-cuadrado
- 66)** ¿Cuál de las siguientes es una ventaja de las redes neuronales respecto a los árboles de decisión?
- A) Capacidad de manejo conjunto de datos numéricos y categorizados.
 - B) Se pueden extraer las reglas generadas para clasificar/predecir ("caja blanca").
 - C) Puede trabajar con valores que faltan (valores "en blanco", "missing values").
 - C) Mayor capacidad de generalización para casos distintos a los del conjunto de entrenamiento.

- 67) ¿Para qué puede usar un modelo ARIMA?
- A) para separar la tendencia, periodicidades y residuos de una serie temporal de datos.
 - B) para identificar las variables más importantes en un modelo lineal.
 - C) para realizar comparaciones múltiples después de un ANOVA positivo.
 - D) para corregir los valores de q en comparaciones múltiples.
- 68) ¿Qué es más útil para valorar un experimento con muestras de tamaños masivos?
- A) el tamaño del efecto
 - B) el valor p
 - C) el valor q
 - D) el pFDR
- 69) ¿Cuál de las siguientes afirmaciones es cierta sobre la cota de Cramér-Rao?
- A) Nos da el valor máximo de una distribución de probabilidad.
 - B) Nos da el valor mínimo de una distribución de probabilidad.
 - C) Nos da la cota inferior para la varianza de un estimador insesgado.
 - D) Nos da la cota inferior para la *skewness* de un estimador insesgado .
- 70) ¿Cuál de las siguientes afirmaciones es cierta?
- A) La Transformada de Fourier y la Transformada Z son la misma.
 - B) La Transformada de Fourier y la Transformada Z no están relacionadas.
 - C) La Transformada de Fourier es una generalización de la Transformada Z.
 - D) La Transformada Z es una generalización de la Transformada de Fourier.
- 71) ¿Cuál de las siguientes afirmaciones sobre el algoritmo de Metrópolis-Hastings es FALSA?
- A) Es un método de Montecarlo basado en cadenas de Markov.
 - B) Sirve para calcular integrales.
 - C) Obtiene una secuencia de muestras aleatoria de una distribución de probabilidad.
 - D) No usa ningún criterio de aceptación del valor propuesto.
- 72) ¿Cuál de las siguientes afirmaciones sobre la matriz de covarianza de una señal física es cierta?
- A) Nunca tiene inversa.
 - B) No es cuadrada.
 - C) Es simétrica.
 - D) La diagonal puede tener valores negativos.
- 73) ¿Cuál de las siguientes afirmaciones es cierta?
- A) La Transformada de Wavelet conserva información temporal y frecuencial.
 - B) La Transformada de Wavelet conserva información temporal pero no frecuencial.
 - C) La Transformada de Wavelet conserva información frecuencial pero no temporal.
 - D) La Transformada de Wavelet no conserva información ni temporal, ni frecuencial.

74) ¿Qué es el método Montecarlo?

- A) Un método no determinista o estadístico numérico, usado para aproximar expresiones matemáticas complejas y costosas de evaluar con exactitud.
- B) Una herramienta financiera para la predicción de la evolución de la deuda pública.
- C) Una ecuación determinista para evaluar el área bajo la curva de ecuaciones de segundo grado.
- D) Un método de desarrollo de software para el análisis de datos de física de Altas Energías.

75) ¿Cuál de los siguientes NO es un conocido algoritmo probabilístico?

- A) Las Vegas
- B) Montecarlo
- C) Atlantic City
- D) AKS

76) El método de *jackknife*...

- A) se basa en repetir el análisis muchas veces tomando cada vez una submuestra de la muestra original.
- B) se basa en repetir el análisis tantas veces como el tamaño de la muestra eliminando cada vez una observación de la muestra original.
- C) se basa en repetir el análisis muchas veces cambiando cada vez el orden en que se evalúan las muestras.
- D) se basa en repetir el análisis extrapolando valores no disponibles a partir de otros valores de la muestra.

77) Una red neuronal debe de tener...

- A) menos de 40 capas.
- B) al menos dos capas.
- C) exactamente cuatro capas.
- D) exactamente tres capas.

78) La normalización y centrado son técnicas usadas para...

- A) poner en la misma escala variables de diferente magnitud.
- B) aumentar las diferencias entre variables.
- C) obtener valores más extremos.
- D) uniformizar las varianzas.

79) En el contexto de *machine learning*, los hiperparámetros son...

- A) variables con más de tres dimensiones.
- B) las variables que se estiman durante el proceso de aprendizaje.
- C) los parámetros que se usan para controlar el proceso de aprendizaje.
- D) Todas las anteriores son correctas.

80) En Biomedicina, RCSB es el principal repositorio de datos de...

- A) secuencias de DNA.
- B) secuencias de proteínas.
- C) estructuras de macromoléculas.
- D) drogas y compuestos químicos.

- 81)** ZINC es una base de datos de...
- A) drogas de interés farmacológico.
 - B) compuestos químicos disponibles comercialmente.
 - C) estructura de moléculas pequeñas resueltas por cristalografía.
 - D) características moleculares calculadas por mecánica cuántica.
- 82)** ¿Dónde buscaría información sobre demografía y población de la sociedad española?
- A) En el Centro de Investigaciones Sociológicas del CSIC.
 - B) en ISOC.
 - C) En MARCATHLON.
 - D) en el Instituto Nacional de Estadística.
- 83)** En el contexto de la estimación de parámetros, ¿cuál de las siguientes afirmaciones describe mejor el método de los mínimos cuadrados?
- A) Es un método que utiliza la estadística Bayesiana en su formulación para encontrar los valores óptimos de los parámetros.
 - B) Es un método que minimiza la suma de los residuos al cuadrado entre los valores observados y los predichos por el modelo.
 - C) Es un método que busca maximizar la función de verosimilitud de los parámetros.
 - D) Es un método que utiliza la regresión lineal para estimar los parámetros del modelo.
- 84)** ¿Cuál de los siguientes métodos numéricos utiliza una fórmula de interpolación para aproximar la integral de una función?
- A) Regla de Simpson.
 - B) Método de la cuadratura gaussiana.
 - C) Método de Romberg.
 - D) Regla del punto medio.
- 85)** ¿Cuál de los siguientes métodos se utiliza para la ingestión de datos en tiempo real en un sistema de procesamiento de *streaming*?
- A) Kafka
 - B) ETL tradicional
 - C) Almacenamiento en disco duro
 - D) Bases de datos relacionales
- 86)** ¿Cuál debe ser el criterio definitivo para aceptar o rechazar un modelo?
- A) La evolución del valor AIC.
 - B) La evolución del valor BIC.
 - C) El orden de adición/eliminación de variables.
 - D) Su adecuación a la realidad.
- 87)** ¿Qué tipo de datos NO son relevantes para el modelado y análisis de datos en química?
- A) Coordenadas geográficas.
 - B) Coordenadas atómicas.
 - C) Parámetros físicos.
 - D) Datos espectroscópicos.

- 88)** Los portales de acceso a datos en abierto...
- A) son repositorios de datos con diferentes tipos de restricciones.
 - B) contienen datos públicos disponibles solo para uso académico.
 - C) permiten reutilizar y redistribuir los datos depositados sin ninguna restricción.
 - D) no permiten el uso comercial de los datos depositados.
- 89)** ¿Cuál de estas métricas NO es relevante para la evaluación de modelos predictivos?
- A) Sensibilidad
 - B) Mediana
 - C) Precisión
 - D) AUC
- 90)** El análisis de agrupamiento o *clustering* se refiere a...
- A) las técnicas más usadas actualmente en el tratamiento de imágenes digitales.
 - B) un conjunto de técnicas estadísticas para clasificar datos individuales en grupos homogéneos.
 - C) un método determinista basado en cálculo numérico.
 - D) la aplicación de ecuaciones diferenciales a modelos científicos.
- 91)** De los siguientes identificadores de contenidos digitales, indicar cuál el que tiene menor nivel de persistencia:
- A) URL
 - B) PURL
 - C) DOI
 - D) Handle
- 92)** La iniciativa de carácter internacional surgida hacia el año 2000 cuyo principal objetivo es garantizar la rapidez y eficiencia, tanto en el acceso a la información existente en archivos y servicios astronómicos, como en el análisis de dicha información es:
- A) Observatorio Virtual (VO)
 - B) Astronomía Global (GA)
 - C) Archivo Astronómico (AA)
 - D) Astrofísica Internacional (IA)
- 93)** Indicar cuál de las siguientes opciones NO se corresponde con alguno de los experimentos coordinados de proyección climática global y regional más recientes integrados en el Interactive Atlas del sexto informe del IPCC WGI:
- A) CMIP5
 - B) CORDEX
 - C) CMIP6
 - D) MPI

94) ¿Cuál de las siguientes afirmaciones describe correctamente la secuenciación de alto rendimiento o secuenciación masiva?

- A) Es una técnica que permite la lectura de un único fragmento de ADN a la vez.
- B) Es una técnica que se utiliza para estudiar la estructura tridimensional de las proteínas.
- C) Es una técnica que permite la lectura simultánea de millones de fragmentos de ADN.
- D) Es una técnica que analiza la expresión de genes en muestras de tejido para identificar alteraciones genéticas.

95) ¿Cuál de las siguientes opciones describe correctamente la metagenómica?

- A) La metagenómica es el estudio de la función de los genes en un solo organismo.
- B) La metagenómica es una técnica que permite la clonación de genes específicos en bacterias.
- C) La metagenómica es el estudio de la expresión génica.
- D) La metagenómica es el estudio del material genético de una muestra ambiental compleja, sin necesidad de cultivar los microorganismos individuales presentes.

96) Los modelos de redes neuronales LSTM...

- A) no se aplican al análisis de series temporales.
- B) son un ejemplo de redes recurrentes.
- C) mejoran, en general, los resultados de Transformers.
- D) modelizan sólo memoria a corto plazo.

97) La regularización permite típicamente...

- A) incrementar la complejidad del modelo.
- B) acelerar los tiempos de entrenamiento.
- C) reducir el sesgo del modelo.
- D) Reducir la varianza del modelo a cambio de un ligero incremento del sesgo.

98) ¿Qué propiedades intentamos detectar al diagnosticar un modelo lineal?

- A) Posible correlación de los términos de error.
- B) Varianza no constante de los términos de error.
- C) Colinealidad.
- D) Todas las anteriores.

99) ¿Cuál de los siguientes no es un método Montecarlo basado en cadenas de Markov?

- A) Montecarlo Hamiltoniano
- B) Muestreador de Gibbs
- C) Método de inversión
- D) Algoritmo de Metropolis

100) ¿Cuál de las siguientes aproximaciones estadísticas modeliza probabilísticamente la incertidumbre a priori?

- A) Estimación de máxima verosimilitud.
 - B) Inferencia bayesiana.
 - C) Inferencia fiducial.
 - D) Inferencia fisheriana.
-