



PROCESO SELECTIVO POR EL SISTEMA DE PROMOCIÓN INTERNA PARA INGRESO EN LA ESCALA DE TÉCNICOS SUPERIORES ESPECIALIZADOS DE LOS ORGANISMOS PÚBLICOS DE INVESTIGACIÓN, CONVOCADO POR RESOLUCION DE 24 DE MAYO DE 2024 (BOE N° 131 DE 30 DE MAYO)

Cuestionario del primer ejercicio

A6 – CIENCIA DE DATOS

Especialidad: D2 – PROGRAMACIÓN Y COMPUTACIÓN CIENTÍFICA

- No abra el **CUESTIONARIO** ni empiece el examen hasta que se le indique.
- Solo se calificarán las respuestas marcadas en la **HOJA DE RESPUESTAS**
- El cuestionario consta de **80 preguntas**, cada una de ellas con **cuatro respuesta alternativas**, de las cuales **sólo una de ellas es correcta**.
- 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.
- Las **contestaciones erróneas se PENALIZARÁN** con un 25 % de su valoración.
- Lea atentamente las **instrucciones** para contestar la **HOJA DE RESPUESTAS**, que figuran al dorso de la misma.
- Cumplimente los datos personales y firme la **HOJA DE RESPUESTAS**.
- El tiempo para la realización de este ejercicio será de **noventa (90) minutos**.
- **NO SEPARE** ninguna de las copias de la **HOJA DE RESPUESTAS**. Una vez finalizado el ejercicio, el personal del aula le indicará los pasos a seguir.
- El **CUESTIONARIO** se podrá utilizar como borrador y se podrá llevar por el opositor al finalizar el tiempo marcado para el ejercicio.

Barcelona 20 de septiembre de 2024

1. Indicar de las siguientes características cuál NO se corresponde con el modelo de computación en la nube:

- A) Acceso bajo demanda.
- B) Flexibilidad en la configuración de los recursos.
- C) Elasticidad y rapidez en el aprovisionamiento de recursos.
- D) Dependencia de recursos físicos locales.

2. La virtualización de recursos consiste en:

- A) El conjunto de técnicas software que permiten abstraer las características físicas de un computador, de forma que se puedan ejecutar diferentes máquinas virtuales.
- B) La utilización de máquinas virtuales hardware para proporcionar acceso al software existente en un sistema para un grupo virtual de usuarios.
- C) La explotación de organizaciones virtuales para el aprovisionamiento de recursos de computación a gran escala, especialmente en colaboraciones científicas.
- D) La sustitución de los recursos humanos de una organización por asistentes basados en inteligencia artificial.

3. Indicar cuál de los siguientes conceptos NO es una propiedad de seguridad de un activo de información en una organización.

- A) Confidencialidad.
- B) Integridad.
- C) Migralidad.
- D) Autenticidad.



4. Indicar cuál de los siguientes es un motor de bases de datos relacional:

- A) MariaDB.
- B) MongoDB.
- C) Apache Cassandra.
- D) Amazon DynamoDB.

5. ¿Cuál de las siguientes características es fundamental en el desarrollo de software abierto?

- A) Código cerrado y restrictivo.
- B) Licencias propietarias.
- C) Colaboración y transparencia.
- D) Propiedad exclusiva del autor.

6. ¿Cuál de las siguientes afirmaciones es correcta acerca del aprendizaje automático (machine learning)?

- A) El aprendizaje automático se refiere únicamente a la capacidad de las máquinas para procesar grandes volúmenes de datos.
- B) El aprendizaje automático es un enfoque que permite a las máquinas aprender y mejorar automáticamente a través de la experiencia sin ser programadas explícitamente.
- C) El aprendizaje automático solo se aplica en la investigación científica y no tiene aplicaciones en otros campos.
- D) El aprendizaje automático es una técnica exclusiva de la programación orientada a objetos.

7. ¿Qué es un Data Management Plan (Plan de Gestión de Datos)?

- A) Un plan para la eliminación de datos obsoletos y sin valor.
- B) Un plan para asegurar que los datos sean inaccesibles y no puedan ser utilizados.
- C) Un plan que establece cómo se gestionarán, organizarán, almacenarán y compartirán los datos durante un proyecto de investigación.
- D) Un plan que prioriza la recopilación de datos sin tener en cuenta su calidad o relevancia.

8. ¿Cuál de los siguientes métodos es un ejemplo de accounting en los Sistemas de Autorización, Autenticación y Accounting?

- A) Control de acceso físico.
- B) Registro de tiempos de sesión.
- C) Cifrado de datos.
- D) Detección de intrusiones.

9. ¿Qué es GitHub?

- A) Una plataforma de desarrollo colaborativo.
- B) Un sistema operativo .
- C) Un lenguaje de programación.
- D) Un motor de búsqueda.



10. En un CPD de un centro de investigación se va a adquirir un Sistema de Alimentación Ininterrumpida (SAI) y tenemos las siguientes especificaciones: 500 VA (Voltiamperios) y 300 W (vatios), que indican:

- A) La potencial real y la potencia aparente respectivamente del SAI.
- B) La energía eléctrica y la potencia que puede suministrar el SAI.
- C) La potencia aparente y la potencia real del SAI.
- D) La potencia del SAI solo que dados en dos unidades diferentes de potencia.

11. En el contexto de seguridad de sistemas informáticos, la autenticación:

- A) Otorga certificados digitales a los recursos que están disponibles.
- B) Confirma la disponibilidad de los recursos.
- C) Confirma que los usuarios son quienes dicen ser validando su identidad.
- D) Proporciona permisos de acceso a los datos almacenados en los recursos.

12. ¿Cuál de estos lenguajes proporciona una plataforma unificada y un modelo de programación que permite la programación paralela de GPUs, CPUs, DSPs y FPGAs?

- A) OpenCL.
- B) OpenMP.
- C) CUDA.
- D) SYCL.



13. ¿Cuál de las siguientes afirmaciones sobre XML es correcta?

- A) XML es un modelo de metadatos utilizado para describir la presentación de una página web
- B) XML es un lenguaje de programación utilizado para crear páginas web interactivas.
- C) XML es un lenguaje de marcado que permite definir etiquetas personalizadas para estructurar y almacenar datos.
- D) XML es un formato de imagen utilizado para gráficos y multimedia en la web

14. Indica cuál de estos sistemas operativos es de tiempo real

- A) QNX
- B) Fedora Linux
- C) Windows RT
- D) Minix

15. ¿Cuál de las siguientes afirmaciones describe mejor el concepto de escalabilidad en la computación científica en la nube?

- A) Se refiere a la capacidad de almacenar grandes cantidades de datos científicos en la nube, en escalas mayores del Petabyte.
- B) Implica la posibilidad de acceder a recursos de cómputo adicionales según las necesidades del usuario.
- C) Significa utilizar software especializado en la nube para realizar cálculos científicos complejos.
- D) Se refiere a la colaboración y el intercambio de datos científicos entre diferentes investigadores en las diferentes escalas o etapas de la investigación.



16. De entre las siguientes funcionalidades, indica cuál NO está relacionada con la gestión ágil de proyectos:

- A) Diagramas de Gantt interactivos.
- B) Tableros Kanban.
- C) Hojas de ruta de commits automatizados.
- D) Planificación de minitrabajos Sprints.

17. ¿Qué función principal tiene un sistema de control de versiones en un repositorio digital de software?

- A) Almacenar y organizar los archivos de código fuente.
- B) Permitir a los desarrolladores colaborar y trabajar en equipo en el mismo proyecto.
- C) Gestionar y rastrear los cambios realizados en los archivos a lo largo del tiempo.
- D) Compilar y ejecutar el software en diferentes entornos de desarrollo.

18. En la programación orientada a objetos, una clase es:

- A) Una plantilla en la que se definen los atributos y métodos predeterminados de un tipo de objeto.
- B) Una propiedad que permite subdividir una aplicación en partes más pequeñas e independientes de la aplicación en sí.
- C) Una técnica por la cual el entorno de objetos se encarga de destruir automáticamente los objetos no referenciados.
- D) Una ubicación de almacenamiento abstracta asociada a un nombre simbólico, que contiene una cantidad de información conocida.



19. ¿En qué consiste el aprendizaje no supervisado?

- A) Un tipo de aprendizaje en el que no se utilizan etiquetas para entrenar el modelo.
- B) Un tipo de aprendizaje que solo se aplica a problemas de regresión.
- C) Un tipo de aprendizaje que utiliza redes neuronales recurrentes (RNN).
- D) Un tipo de aprendizaje que utiliza regresión logística.

20. ¿Cuál de las siguientes opciones describe qué es Big Data?

- A) Una cantidad de datos que es demasiado grande para ser procesada por sistemas tradicionales.
- B) Un conjunto de herramientas utilizadas para el procesamiento de datos en tiempo real.
- C) Una base de datos de gran tamaño.
- D) Una técnica para el análisis de datos utilizando algoritmos de inteligencia artificial.

21. ¿Qué es una TPU?

- A) Es una unidad de procesamiento gráfico.
- B) Es una unidad de procesamiento central.
- C) Es una unidad de procesamiento tensorial.
- D) Es una unidad de procesamiento vectorial.

22. El aprendizaje automático se ha visto revolucionado en los últimos años por la evolución de distintas formas de computación basada en redes neuronales. El factor clave para el uso extendido de estos nuevos algoritmos ha sido

- A) El rápido acceso a disco.
- B) La eficiente implementación en Unidades Centrales de Procesamiento.
- C) La eficiente implementación en Unidades Gráficas de Procesamiento.
- D) La reducción de la latencia en la transmisión de información.



23. La principal característica de un sistema de tiempo real es:

- A) Disponer de un planificador basado en prioridades.
- B) Respuesta a interrupciones con niveles de prioridad.
- C) Respuesta a los eventos dentro de un plazo determinado.
- D) Ausencia de gestor de memoria virtual.

24. El soporte y coordinación para el tratamiento de vulnerabilidades y la resolución de incidentes de seguridad que tengan las Administraciones de las Comunidades Autónomas corresponde a:

- A) INCIBE-CERT.
- B) CCN-CERT.
- C) Al CERT correspondiente de la Comunidad Autónoma.
- D) Al CERT del Ministerio de Defensa.

25. ¿Cuál es el protocolo más común para iniciar una sesión remota segura mediante línea de comandos a una máquina Linux?

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

26. ¿Cuál de las siguientes técnicas se utiliza para reducir la dimensionalidad de los datos en aprendizaje profundo?

- A) Data augmentation (aumento de datos).
- B) Reducción de ruido.
- C) Redes neuronales convolucionales (CNN).
- D) Análisis de componentes principales (PCA).



27. ¿Cuál de las siguientes bibliotecas NO es ampliamente utilizada para implementar deep learning en Python?

- A) TensorFlow.
- B) Scikit-learn.
- C) PyTorch.
- D) Wamba.

28. Un puntero es:

- A) Lo mismo que la cima de una pila.
- B) Un dato cuyo valor es la dirección en memoria de otro determinado dato.
- C) Lo contrario que la cima de una pila.
- D) Un dato cifrado.

29. En el lenguaje de programación en R, ¿cuál es la función para obtener la inversa de la matriz A?

- A) `inv(A)`
- B) `det(A)`
- C) `solve(A)`
- D) `t(A)`

30. En el lenguaje de programación en R, ¿cuál sería la instrucción para obtener ayuda sobre la función mean?

- A) `mean.help`
- B) `mean.help()`
- C) `help(mean)`
- D) `help.mean`



31. ¿Cuál de las siguientes es la extensión correcta del archivo Python?

- A) .py
- B) .python
- C) .p
- D) .pyton

32. SMI significa:

- A) Software Maturity Indicator.
- B) Software Maturity Index.
- C) Software Mature Indicator.
- D) Software Mature Index.

33. La calidad del código mide la capacidad de

- A) Mantenibilidad del código.
- B) Mantenibilidad y legibilidad del código.
- C) Mantenibilidad, legibilidad y eficiencia del código.
- D) Complejidad, mantenibilidad, legibilidad y eficiencia del código.

34. ¿Qué es un procesador multinúcleo de alto rendimiento que se puede utilizar para acelerar una amplia variedad de aplicaciones mediante computación paralela?

- A) CPU.
- B) GPU.
- C) DSP.
- D) FPGA.



35. Un versátil y potente gestor de paquetes de código abierto diseñado específicamente para entornos de computación de alto rendimiento (HPC):

- A) Spack.
- B) Spark.
- C) Yum.
- D) REST.

36. ¿Cuál de las siguientes afirmaciones sobre la Ley de Amdahl es verdadera?

- A) El rendimiento de una aplicación aumenta a la par que se incrementa el número de procesadores.
- B) La mejora obtenida en el rendimiento de un sistema debido a la alteración de uno de sus componentes no está limitada por la fracción de tiempo que se utiliza dicho componente.
- C) La Ley de Amdahl trata sobre las ventajas de la “paralelización” y los límites en escalabilidad, donde estos límites están asociados a la fracción de código no “paralelizable”.
- D) La ley de Amdahl no tiene relación con la la mejora máxima de rendimiento de un sistema de información.

37. ¿Qué es un clúster de computación?

- A) Es una agrupación local de recursos computacionales autónomos, normalmente de prestaciones similares, interconectados y que trabajan de manera estrecha.
- B) Un sistema paralelo distribuido por varios países (de ámbito global).
- C) Un sistema distribuido paralelo de ámbito global.
- D) Un sistema virtualizado.



38. ¿Qué metodología de gestión de proyectos se basa en la iteración y el desarrollo incremental?

- A) Cascada.
- B) Scrum.
- C) PRINCE2.
- D) PMI.

39. ¿Cuál es el principal objetivo de DevOps?

- A) Aumentar el número de servidores en la nube.
- B) Mejorar la colaboración entre equipos de desarrollo y operaciones.
- C) Reducir el costo de las licencias de software.
- D) Acelerar el proceso de ventas.

40. ¿Qué herramienta se utiliza comúnmente para la integración continua en DevOps?

- A) Microsoft Word.
- B) Adobe Photoshop.
- C) Jenkins.
- D) Slack.

41. ¿Cuál de las siguientes prácticas es fundamental en un entorno DevOps?

- A) Desarrollar en entornos aislados sin pruebas.
- B) Implementar cambios en producción sin revisión.
- C) Automatizar pruebas y despliegues.
- D) Mantener la documentación en papel.



42. ¿Cuál de las siguientes estructuras de datos permite un acceso rápido a elementos en cualquier posición?

- A) Lista enlazada.
- B) Árbol binario.
- C) Array.
- D) Cola.

43. ¿Qué algoritmo es comúnmente utilizado para ordenar una lista de elementos?

- A) Algoritmo de Dijkstra.
- B) Algoritmo de Quicksort.
- C) Algoritmo de Floyd-Warshall.
- D) Algoritmo de Huffman.

44. ¿Qué estructura de datos se utiliza para gestionar una colección de elementos en la que el último elemento añadido es el primero en ser retirado?

- A) Pila.
- B) Cola.
- C) Conjunto.
- D) Árbol.

45. ¿Cuál de los siguientes conceptos es central en la programación funcional?

- A) Uso de clases y objetos.
- B) Funciones como elementos básicos.
- C) Herencia múltiple.
- D) Manejo explícito de memoria.

46. ¿Qué término se utiliza en programación funcional para describir una función que devuelve otra función?

- A) Función pura.
- B) Función de orden superior.
- C) Función recursiva.
- D) Función anónima.

47. ¿Qué característica describe mejor las funciones puras en la programación funcional?

- A) Tienen efectos secundarios.
- B) Su salida depende solo de sus entradas.
- C) No pueden ser usadas en un entorno concurrente.
- D) Son mutables.

48. ¿Qué estructura de control se utiliza para manejar excepciones en Python?

- A) try-except
- B) catch-finally
- C) error-handling
- D) try-catch

49. ¿Qué herramienta se utiliza principalmente para el control de versiones del código fuente?

- A) Jenkins.
- B) Git.
- C) Eclipse.
- D) Maven.



50. ¿Qué técnica se utiliza para mejorar la eficiencia del código mediante la eliminación de operaciones redundantes y reorganización del mismo?

- A) Refactorización.
- B) Caching.
- C) Profiling.
- D) Compilación Just-In-Time.

51. ¿Qué es el profiling en el contexto de optimización del software?

- A) Un proceso para medir y analizar el tiempo de ejecución de las funciones del programa.
- B) Un método para proteger el código fuente.
- C) Una técnica para diseñar interfaces de usuario.
- D) Un tipo de prueba de seguridad.

52. ¿Qué biblioteca se utiliza comúnmente en entornos HPC para realizar cálculos paralelos en C++?

- A) MPI.
- B) OpenCL.
- C) CUDA.
- D) Pthreads.

53. ¿Qué caracteriza a los entornos HTC en comparación con los HPC?

- A) Se enfoca en el procesamiento paralelo intensivo.
- B) Se enfoca en la ejecución de grandes cantidades de tareas independientes.
- C) Se centra en la minimización del tiempo de ejecución de tareas individuales.
- D) Utiliza exclusivamente hardware especializado.



54. ¿Cuál es una herramienta comúnmente utilizada para gestionar y coordinar trabajos en entornos HTC?

- A) Slurm.
- B) Docker.
- C) GitLab.
- D) TensorFlow.

55. ¿Qué son las librerías POSIX Threads?

- A) Unas librerías de lenguaje R para visualizar datos.
- B) Unas librerías de Python para procesar bases de datos.
- C) Unas librerías de Linux utilizables en C para generar hilos de ejecución paralelos.
- D) Unas librerías para la gestión de datos bioinformáticos.

56. Un portátil tiene 4 cores. ¿Es un computador paralelo?

- A) No. Es un único computador y no puede paralelizar.
- B) No. Puede ser utilizado para conectarse a un computador que sea realmente paralelo, pero no lo es.
- C) Sí. Podemos ejecutar varios programas simultáneamente o lanzar varios hilos de ejecución de un mismo programa.
- D) Sí. Pero solo si tienes permiso del administrador.

57. ¿Qué estándar de programación se utiliza para desarrollar aplicaciones que pueden ejecutarse en diferentes tipos de aceleradores, incluyendo GPUs y CPUs?

- A) CUDA.
- B) OpenCL.
- C) Vulkan.
- D) DirectX.



58. ¿Qué característica es fundamental para la integración de una base de datos relacional en una aplicación científica?

- A) Soporte para transacciones y ACID (Atomicidad, Consistencia, Aislamiento, Durabilidad).
- B) Capacidad para almacenar datos en formato XML.
- C) Funcionalidad para la visualización de datos.
- D) Capacidad de ejecución en tiempo real.

59. ¿Cuál es una ventaja de utilizar una base de datos relacional para aplicaciones científicas?

- A) Permite la modelización de datos no estructurados.
- B) Ofrece robustas capacidades de análisis y consulta estructurada.
- C) Facilita la integración con plataformas de redes sociales.
- D) Maximiza la velocidad de procesamiento de imágenes.

60. ¿Qué tipo de base de datos NoSQL se basa en un modelo de documentos?

- A) MongoDB.
- B) Redis.
- C) Neo4j.
- D) Cassandra.

61. ¿Qué tipo de base de datos NoSQL es adecuada para aplicaciones que necesitan gestionar relaciones complejas entre datos?

- A) Base de datos de documentos.
- B) Base de datos de columnas.
- C) Base de datos de grafos.
- D) Base de datos de clave-valor.



62. ¿Qué ocurre si se aplica un filtro de suavizado o promediado antes de aplicar un filtro de detección de bordes en una imagen?

- A) Se reduce la intensidad de los bordes detectados.
- B) Se resaltan los bordes en la imagen.
- C) No hay ningún efecto en la detección de bordes.
- D) Los bordes nunca se detectan.

63. ¿Qué tipo de filtro se utiliza comúnmente para resaltar los bordes en una imagen?

- A) Filtro de media (average filter).
- B) Filtro de mediana.
- C) Filtro de Sobel.
- D) Filtro gaussiano.

64. ¿Qué tipo de arquitectura se utiliza para gestionar y coordinar múltiples servicios en un entorno distribuido?

- A) Microservicios.
- B) Monolítica.
- C) Cliente-servidor.
- D) Cliente-punto a punto.

65. ¿Cuál es el objetivo principal de la aritmética computacional?

- A) Resolver problemas de álgebra lineal.
- B) Realizar cálculos precisos y eficientes con números en una computadora.
- C) Diseñar algoritmos de búsqueda.
- D) Crear gráficos interactivos.



66. ¿Qué problema puede ocurrir debido a la representación finita de números en una computadora?

- A) Desbordamiento.
- B) Sobrecarga.
- C) Redundancia.
- D) Estática.

67. ¿Qué tipo de problemas puede resolver el método de Monte Carlo?

- A) Problemas deterministas simples.
- B) Problemas estocásticos y probabilísticos.
- C) Problemas lineales de programación.
- D) Problemas de optimización convexa.

68. ¿Qué algoritmo es conocido por su uso en la compresión de datos y la reducción de redundancia?

- A) Algoritmo de Huffman.
- B) Algoritmo de Dijkstra.
- C) Algoritmo de Bellman-Ford.
- D) Algoritmo de Prim.

69. ¿Cuál es el propósito de la segmentación en el procesamiento de imágenes?

- A) Mejorar la calidad del color.
- B) Identificar y separar diferentes regiones o objetos en una imagen.
- C) Reducir el tamaño del archivo de imagen.
- D) Aumentar el contraste de la imagen.

70. ¿Qué técnica se utiliza para mejorar el contraste y la nitidez de una imagen?

- A) Ecuilización del histograma.
- B) Compresión JPEG.
- C) Transformada de Hough.
- D) Interpolación bilineal.



71. ¿Cuál de las siguientes opciones describe mejor el concepto de “licencia de código abierto” en el contexto del software en ciencia abierta?

- A) Permite el uso exclusivo del software por parte de una institución o individuo.
- B) Otorga derechos exclusivos sobre el software al autor o al desarrollador original.
- C) Permite el acceso, uso, modificación y redistribución del software por parte de cualquier persona.
- D) Limita el acceso y el uso del software solo a investigadores que participan en proyectos específicos de ciencia abierta.

72. ¿Cuál de las siguiente técnicas es utilizada para optimizar el rendimiento de un código al reducir la cantidad de veces que se ejecutan operaciones que son costosas o repetitivas?

- A) Paralelización.
- B) Memorización.
- C) Compresión de código.
- D) Refactorización.

73. ¿Cuál de estos paquetes de visualización para python NO permite representar en 3D?

- A) Matplotlib.
- B) Plotly.
- C) Mayavi.
- D) Seaborn.

74. ¿Cuál de las siguientes herramientas NO está diseñada para el despliegue y/o configuración automatizada de aplicaciones en la nube?

- A) Ansible.
- B) Puppet.
- C) Terraform.
- D) NumPy.



75. La librería científica BLAS es una librería de:

- A) Procesado de imágenes.
- B) Comunicaciones.
- C) Álgebra lineal.
- D) Aprendizaje automático.

76. En el modelo de programación de CUDA, un "kernel" es:

- A) Una función que al ser llamada se ejecuta N veces en paralelo por N hilos distintos.
- B) Una función que se ejecuta con prioridad alta.
- C) Una función que sólo tiene acceso a la memoria de sólo lectura.
- D) Una función que no requiere el acceso a la memoria compartida.

77. Una de éstas NO es una metodología "ágil" de desarrollo de software:

- A) Tatum.
- B) Lean.
- C) Scrum.
- D) Kanban.

78. ¿Qué ventaja aporta el uso de técnicas Map-Reduce para realizar consultas sobre datasets masivos frente a las técnicas de bases de datos relacionales?

- A) Un programa Map-Reduce puede ejecutarse sobre un cluster distribuido de nodos.
- B) Un código Map-Reduce es más corto y legible que una consulta SQL.
- C) Un código Map-Reduce siempre se ejecuta más rápido que con cualquier otra técnica, sobre cualquier tipo de dataset.
- D) Map-Reduce ofrece la posibilidad de hacer checkpointing automático de aplicaciones.



79. ¿Cuál de las siguientes afirmaciones es correcta?

- A) Un objeto hereda los atributos de la clase a la que pertenece.
- B) Las clases se definen a partir de los objetos.
- C) Todos los atributos de un objeto son públicos.
- D) Todos los métodos de un objeto son privados.

80. ¿Qué es la sobrecarga de métodos en programación orientada a objetos?

- A) La capacidad de una clase para tener múltiples constructores.
- B) La capacidad de una clase para heredar métodos de otras clases.
- C) La posibilidad de definir múltiples métodos con el mismo nombre, pero diferente número tipo de parámetros.
- D) La habilidad de una clase para ocultar sus propiedades a otras clases.