PRUEBA DE CONOCIMIENTOS BÁSICOS

PREGUNTAS GENERALES

1. **En un proyecto de Data Science, ¿cuál es el propósito principal de ejecutar `pip-compile requirements.in` (paquete `pip-tools`)?:**

   a) Generar un `requirements.txt` con las versiones exactas de todas las dependencias, incluidas las transitivas, resueltas en el momento de la compilación.

   b) Comprobar si existen vulnerabilidades de seguridad conocidas en los paquetes declarados en `requirements.in`.

   c) Instalar automáticamente todas las dependencias listadas en `requirements.in` en el entorno virtual activo.

   d) Crear un entorno virtual aislado con las dependencias especificadas en `requirements.in`.

2. **¿Qué efecto tiene el archivo `.python-version` con contenido `3.11.5` en la raíz de un repositorio que usa `pyenv`?:**

   a) Documenta la versión de Python requerida por el proyecto de forma informativa, pero no activa esa versión automáticamente.

   b) Lista la versión de Python que `venv` debe usar al crear el entorno virtual del proyecto.

   c) Configura el alias `python` del sistema operativo para que apunte permanentemente a la versión 3.11.5.

   d) Indica a `pyenv` que debe usar automáticamente Python 3.11.5 al ejecutar comandos en ese directorio.

3. **¿Cuál es la ventaja de `df.query('edad > 30 and ciudad == "Madrid"')` frente al filtrado booleano equivalente `df[(df['edad'] > 30) & (df['ciudad'] == 'Madrid')]`?:**

   a) `query()` usa índices B-tree internos de Pandas que lo hacen siempre significativamente más rápido que el filtrado booleano.

   b) `query()` ofrece una sintaxis más concisa y legible, y puede aprovechar el motor `numexpr` para mejorar el rendimiento en DataFrames grandes si dicha librería está instalada.

   c) `query()` es el único método de Pandas que soporta filtros sobre columnas de tipo `Categorical`.

   d) `query()` evita la creación de todos los arrays booleanos intermedios, reduciendo el pico de uso de memoria a la mitad respecto al filtrado booleano estándar.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

4. Al optimizar hiperparámetros de un modelo XGBoost con 8 hiperparámetros de rango continuo, ¿por qué se prefiere `RandomizedSearchCV` sobre `GridSearchCV`?:

a) Porque `GridSearchCV` no soporta hiperparámetros continuos: solo acepta listas discretas de valores predefinidos.

b) Porque `RandomizedSearchCV` garantiza encontrar el óptimo global del espacio de búsqueda con menos evaluaciones.

c) Porque `RandomizedSearchCV` usa optimización bayesiana internamente, que converge más rápido que la búsqueda exhaustiva.

d) Porque `RandomizedSearchCV` muestrea un número fijo de combinaciones, explorando el espacio de forma más eficiente.

5. ¿Qué información adicional aporta `--cov-report=term-missing` al ejecutar `pytest --cov=src`?:

a) Muestra los números de línea exactos de cada módulo que no han sido ejecutados por ningún test.

b) Lista los tests que no contienen ninguna aserción (`assert`), identificando posibles tests vacíos o incompletos.

c) Genera un informe HTML con las líneas no cubiertas resaltadas en rojo para su revisión en el navegador.

d) Muestra las dependencias importadas en el código fuente que no están referenciadas en ningún test.

6. Su equipo de Machine Learning ejecuta consultas analíticas pesadas sobre la base de datos PostgreSQL de producción, causando picos de latencia en la aplicación. ¿Cuál es la solución más apropiada sin modificar el código de Machine Learning?:

a) Aumentar shared_buffers y work_mem en el servidor primario de PostgreSQL.

b) Dirigir las consultas de ML a read replicas configuradas con streaming replication.

c) Usar TABLESAMPLE BERNOULLI(10) en las consultas de ML para reducir el volumen procesado.

d) Crear índices específicos para cada consulta de ML sobre el servidor primario.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

7. Su sistema tiene 500 workers de inferencia que consultan PostgreSQL simultáneamente. La base de datos lanza errores 'FATAL: sorry, too many clients already'. ¿Cuál es la solución más apropiada?:

a) Aumentar max_connections en PostgreSQL a 1000.

b) Reemplazar psycopg2 por asyncpg en los workers.

c) Desplegar PgBouncer en modo transaction pooling entre los workers y PostgreSQL.

d) Implementar sharding horizontal con Citus.

8. Tiene una colección de MongoDB con 50 millones de documentos de transacciones. Necesita calcular el total de ventas por categoría de producto para un dashboard. ¿Cuál es el enfoque más eficiente?:

a) Exportar todos los documentos con pymongo y calcular con Pandas en Python.

b) Usar find() con proyección y procesar la agrupación en la aplicación.

c) Usar el aggregation pipeline con $match, $group y $sum en el servidor.

d) Crear una colección separada y actualizarla con un script periódico usando Python.

9. Para una aplicación de detección de fraude, se dispone de 10 millones de transacciones de las cuales el 0,01% son fraudulentas. Con estos datos, se entrena un modelo XGBoost que obtiene una precisión (accuracy) del 99,99% aunque solo detecta en fase de test el 70% de los casos de fraude. ¿Cuál es la mejor estrategia para que el modelo mejore la detección de casos fraudulentos?:

a) Optimizar el área bajo la curva Precisión-Recall y a la vez ajustar el hiperparámetro scale_pos_weight.

b) Entrenar un modelo de redes neuronales profundas usando la entropía cruzada como función de coste.

c) Reentrenar el modelo XGBoost submuestreando la clase mayoritaria hasta equilibrar el número de casos fraudulentos al número de casos no fraudulentos.

d) Optimizar el área bajo la curva ROC y emplear la técnica SMOTE para generar sintéticamente más casos fraudulentos hasta conseguir equilibrar el número de casos fraudulentos al número de casos no fraudulentos.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

10. ¿Cuál de estas técnicas NO es una técnica para corrección de sesgos de un modelo entrenado con un conjunto de datos sesgados?:

a) Remuestreado.

b) Cuantización.

c) Optimización restringida con una métrica de fairness.

d) Optimización del umbral.

11. En un sistema clasificador binario, el fairness o equidad, entendida como igualdad de oportunidades, se expresa como:

a) El número de decisiones positivas está igualmente distribuido entre los grupos.

b) La proporción de positivos verdaderos y de falsos positivos es la misma para todos los grupos.

c) El hecho de que la aplicación de una decisión, a priori neutra, tiene distribuciones dispares entre los grupos.

d) La proporción de positivos verdaderos (true positives) es la misma en todos los grupos.

12. ¿En qué caso dos variables aleatorias cuya covarianza es cero son independientes?:

a) Nunca.

b) Solo si ambas siguen una distribución normal.

c) Solo si ambas no siguen una distribución normal.

d) Siempre.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

13. Un banco dispone de una aplicación de evaluación crediticia para empresas basada en 374 indicadores financieros relacionados de manera no lineal y densa en una red neuronal profunda. El sistema deniega una solicitud de crédito a una empresa. Los abogados de la misma reclaman al banco una explicación detallada del motivo de denegación, al amparo de la legislación vigente. ¿Qué técnica se debe emplear para determinar qué variables han tenido más peso en la denegación?:

a) Calcular los valores SHAP específicos para el caso concreto.

b) Estudiar los valores de las neuronas de las capas ocultas.

c) Entrenar un modelo Random Forest equivalente a la red neuronal profunda y calcular la métrica de Gini de los distintos atributos a partir del modelo resultante.

d) Calcular la Importancia de los atributos por permutación global (Global Permutation Feature Importance) sobre el histórico de validación.

14. Un banco va a desplegar cuatro aplicaciones de Inteligencia Artificial en las próximas semanas. ¿Cuál de ellas debe ser considerada de 'Alto Riesgo' por la AI-ACT de la Unión Europea?:

a) Un sistema automático de evaluación crediticia para la concesión de crédito a personas físicas.

b) Un chatbot basado en un modelo de lenguaje en el servicio de atención al cliente para dirigir las peticiones al equipo adecuado.

c) Un modelo de aprendizaje automático que procesa los logs de los servidores para la detección de anomalías de funcionamiento.

d) Un sistema de filtrado de correo no deseado para los empleados del banco.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

15. Se propone un proyecto interno en un banco para desarrollar un modelo que rastree las transacciones de los clientes y sus registros de navegación web, así como la información pública que se puede obtener de redes sociales para crear un índice de fiabilidad del cliente, que será tenido en cuenta a la hora de conceder productos de crédito. ¿Cómo se debe proceder según la AI-ACT de la Unión Europea?:

a) Se deberá publicar el código para asegurar la transparencia del procedimiento.

b) Se deberá informar a los clientes de la existencia del procedimiento a través de la página web del banco.

c) Se trata de una aplicación interna del banco y por tanto no hace falta informar de la misma.

d) Se trata de un tipo de aplicación de alto riesgo por la AI-ACT y debe cumplir ciertos requerimientos antes de la puesta en producción del modelo.

16. Comparamos dos palabras codificadas cada una con un embedding distinto, pero con el mismo número de dimensiones y que mantienen las direcciones en el espacio. ¿Qué métrica podemos usar para valorar si ambas palabras están próximas o lejanas?:

a) Distancia Euclídea.

b) Distancia de Manhattan.

c) Producto Escalar.

d) Distancia del Coseno.

17. Un banco ha implementado un sistema RAG de soporte a los consultores financieros para consultar los informes internos de prospectiva de mercado. Los usuarios señalan que, si bien el sistema encuentra la información de los documentos adecuados, la síntesis del modelo de lenguaje a menudo muestra datos de 2024 en lugar de los de 2026 que se encuentran en los informes. ¿Cuál es el método más efectivo en términos de coste y esfuerzo para corregir este problema?:

a) Realizar un Fine Tuning del modelo de lenguaje con los documentos de 2026.

b) Modificar el prompt del modelo de lenguaje de manera que priorice la información extraída.

c) Incrementar la temperatura del modelo de manera que consiga igualar las probabilidades entre las opciones extraídas.

d) Aumentar el número de elementos recuperados de la base de datos vectorial para asegurar que el modelo dispone de la información correcta.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

18. Un banco está desarrollando un motor de verificación de cumplimiento normativo interno que se alimenta de unos mil documentos de más de veinte páginas que contienen lenguaje legal regulatorio. El sistema debe poder analizar el complejo contexto semántico de los términos dentro de cada párrafo. ¿Cuál es la mejor estrategia de vectorización dado este contexto?:

a) Factorización matricial basada en el TF-IDF (Term Frequency-Inverse Document Frequency).

b) Word2Vec entrenado con artículos sobre aspectos legales y financieros de Wikipedia.

c) Embedding denso de un transformer preentrenado.

d) One-hot encoding sobre un vocabulario específico legal y financiero.

19. ¿Cuál de estas herramientas de automatización de flujos de trabajo NO es del tipo 'no code'?:

a) Zapier.

b) N8N.

c) LangGraph.

d) Make.

20. ¿Qué diferencia hay entre la matriz que enmascara el mecanismo de auto-atención de un transformer en una arquitectura de encoder bidireccional, como BERT, y la que enmascara un decoder causal, como GPT?:

a) No hay ninguna diferencia.

b) La matriz del encoder es triangular inferior y la del decoder triangular superior.

c) La matriz de encoder es triangular superior y la del decoder triangular inferior.

d) La matriz del encoder es completa y la del decoder triangular inferior.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

21. ¿Qué se entiende por Olvido Catastrófico en el contexto del Fine Tuning de modelos?:

a) Es el fenómeno que se da cuando el modelo pierde la capacidad general del preentrenado y solo responde correctamente para el dominio sobre el que se efectúa el Fine Tuning.

b) Es el fenómeno que se produce cuando el modelo solo mantiene la capacidad general del preentrenado y no ha capturado la información del dominio sobre el que se efectúa el Fine Tuning.

c) Es el fenómeno que se produce cuando en el entrenamiento del Fine Tuning se desborda la memoria y se borran los valores de los parámetros del modelo.

d) Es el fenómeno que se produce cuando el dominio del Fine Tuning y el del modelo preentrenado no son compatibles.

22. Se realiza el Fine Tuning de un modelo de cumplimiento (Compliance) en un servidor seguro dentro de la infraestructura de un banco. Este servidor dispone de una GPU con 24 GB de VRAM. Para evitar problemas de memoria, se ha de reducir el tamaño del batch a 2. Se estima, sin embargo, que se necesitaría un tamaño de batch de 64 para un entrenamiento efectivo. ¿Qué estrategia se ha de adoptar para poder entrenar el modelo con el hardware disponible?:

a) Multiplicar el learning rate por 32 para equilibrar el tamaño del batch.

b) Acumular los gradientes durante 32 pasos antes de modificar los pesos.

c) Modificar únicamente 1/32 parte de los pesos en cada paso del batch.

d) Reducir la dimensión del embedding un factor 32 para ahorrar memoria y poder aumentar el tamaño del batch.

23. ¿Cuál es el propósito del Key-Value Cache en la fase de inferencia de un modelo de lenguaje autorregresivo basado en transformers?:

a) Almacenar los valores de la matrices clave y valor de los tokens anteriores para evitar tener que recalcular sus valores a cada nuevo token para ahorrar tiempo de respuesta.

b) El Key-Value Cache tradicional se realiza en la fase de entrenamiento y no de inferencia. Su objetivo es predecir los valores de los tokens futuros para ahorrar tiempo de entrenamiento.

c) Almacenar un conjunto de valores de las matrices clave y valor precalculados según el historial del usuario para ahorrar tiempo de respuesta.

d) Realizar una predicción de los valores de las matrices clave y valor de los tokens futuros para avanzar el cálculo de la respuesta del modelo y así ahorrar tiempo de respuesta.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

24. ¿Cuál de estas métricas se usa para valorar la calidad de un sistema de traducción automática?:

a) BLEU.

b) BLANC.

c) ROUGE.

d) VERT.

25. Se quiere desplegar un modelo de lenguaje propietario de soporte a finanzas de 70 mil millones de parámetros en un servidor seguro de un banco con una GPU de 48 GB de VRAM. ¿Cuál es la precisión con la que se deben guardar los parámetros para que el modelo quepa en la memoria de la GPU?:

a) FP16.

b) FP32.

c) FP4.

d) FP8.

26. ¿Cuál de las siguientes afirmaciones describe una característica INCORRECTA del protocolo MCP (Model Context Protocol)?:

a) Permite la estandarización de la comunicación entre modelos y herramientas externas.

b) Utiliza JSON-RPC 2.0 como protocolo de formato de mensajes.

c) Ofrece capacidades de orquestación entre agentes.

d) Fue definido por Anthropic.

27. En el protocolo A2A (Agent-to-Agent), ¿cómo se conoce al manifiesto que expone el propio agente para proporcionar toda su información en el proceso de descubrimiento por parte de otros agentes?:

a) AgentCapabilites.

b) AgentSkills.

c) AgentCard.

d) AgentManifest.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

28. En una arquitectura RAG, con una base vectorial y modelo de embeddings adecuados, ¿cuál es la consecuencia más probable de usar fragmentos (chunks) grandes al indexar documentación heterogénea?:

a) Se garantiza una mejora simultánea de recall y precisión al contener cada fragmento más información.

b) Se elimina la necesidad de almacenar metadatos, porque el propio tamaño del fragmento aporta suficiente contexto.

c) Se evita por completo la generación alucinada, ya que el modelo siempre recibe suficiente contexto.

d) Se incrementa la probabilidad de recuperar pasajes con información mezclada, lo que puede degradar la precisión de la respuesta final.

29. En el contexto de modelos de lenguaje a gran escala (LLMs), el "Concept Drift" puede manifestarse cuando el modelo se despliega en producción durante largos periodos. ¿Cuál de las siguientes afirmaciones describe con mayor precisión este fenómeno y su impacto en el rendimiento del modelo?:

a) Ocurre cuando el modelo, debido al paso del tiempo, incrementa su tamaño de parámetros tras el despliegue, lo que provoca sobreajuste a nuevos datos.

b) Se produce cuando el significado, el contexto o la relación entre las palabras y la intención del usuario cambian en el mundo real a lo largo del tiempo, reduciendo la precisión del modelo en tareas previamente bien resueltas.

c) Ocurre cuando el modelo maneja información distinta bajo una misma nomenclatura, generando una salida con información conceptual mezclada.

d) Afecta a la degradación del hardware donde se ejecuta el modelo, aumentando la latencia.

30. Se dispone de un sistema productivo basado en agentes. Se desea poner a disposición para dichos agentes un sistema de consulta a través de lenguaje natural a una base de datos relacional. Esta base de datos contiene información compleja de negocio relacionada con productos financieros, manejando términos técnicos poco comunes en el lenguaje cotidiano. Desde el punto de vista de diseño de arquitectura, ¿cuál sería la mejor aproximación?:

a) Exponer un conjunto de “tools” mediante un servidor MCP que abstraiga las consultas sobre la base de datos, para que los agentes lo descubran y puedan llevar a cabo las consultas.

b) Construir un sistema RAG sobre la base de datos relacional.

c) Implementar un agente especializado en el dominio de esta base de datos que haga uso de un conjunto de “tools” de dicha base de datos, expuestas a través de un servidor MCP.

d) Incorporar las reglas de negocio en los agentes existentes para que sean capaces de generar consultas SQL adecuadas.

31. ¿Cuál de las siguientes bases de datos vectoriales NO puede escalar horizontalmente de forma nativa?:

a) Chroma.

b) Milvus.

c) Pgvector.

d) Qdrant.

32. ¿Cuál es la principal finalidad de LangGraph dentro del desarrollo de aplicaciones basadas en modelos de lenguaje?:

a) Optimizar el rendimiento de los modelos.

b) Permitir la orquestación de flujos complejos de agentes y estados mediante grafos dirigidos.

c) Sustituir los modelos de lenguaje por reglas deterministas basadas en árboles de decisión.

d) Generar grafos de relación semántica entre vectores de una base vectorial.

33. En el protocolo A2A (Agent-to-Agent), según su especificación, ¿qué tipo de información NO forma parte de los aspectos que un agente declara de manera explícita en su manifiesto?:

a) Esquema de autenticación soportado en la comunicación con el agente.

b) Modelo utilizado por el agente.

c) Capacidades de respuesta asíncrona (push notifications).

d) Protocolo de comunicación del endpoint de inferencia.

34. Se desea implementar un sistema RAG en producción con una base vectorial de un tamaño medio (aproximadamente 1 millón de vectores), con una gran riqueza en metadatos, donde la estrategia de retrieval debe aplicar un "pre-filter" sobre estos metadatos, garantizando rendimiento. ¿Cuál sería la base de datos vectorial idónea en base a sus características nativas?:

a) Annoy.

b) Milvus.

c) FAISS.

d) QDrant.

35. ¿Cuál de los siguientes frameworks de código abierto NO es una plataforma de observabilidad aplicable a una arquitectura de agentes?:

a) AutoGen.

b) Langfuse.

c) Opik.

d) TruLens.

36. En Kubernetes, ¿qué recurso mantiene el estado deseado de réplicas y gestiona actualizaciones declarativas (rolling update) de un conjunto de Pods?:

a) ReplicaSet.

b) Deployment.

c) DaemonSet.

d) Job.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

37. En un despliegue blue/green, ¿qué característica lo define mejor?:

a) Mantiene dos entornos y cambia el enrutamiento del tráfico al nuevo cuando éste está validado.

b) Actualiza instancias una a una reemplazándolas gradualmente con una estrategia rolling.

c) Divide el tráfico por porcentaje (p. ej., 10%/90%) entre versiones.

d) Requiere detener completamente el servicio para evitar inconsistencias.

38. ¿Qué comando se utiliza para previsualizar los cambios que Terraform aplicaría sin ejecutarlos realmente?:

a) terraform init.

b) terraform show.

c) terraform preview.

d) terraform plan.

39. ¿Qué diferencia fundamental existe entre una imagen y un contenedor en Docker?:

a) La imagen es una instancia en ejecución y el contenedor es una plantilla.

b) La imagen solo existe en memoria y el contenedor solo en disco.

c) No existe diferencia real. Ambos términos son equivalentes.

d) La imagen es inmutable y el contenedor es una instancia ejecutable basada en ella.

40. En Git, ¿qué efecto tiene el uso de git reset --hard <commit>?:

a) Revierte únicamente el último commit manteniendo los cambios en el working directory.

b) Mueve el puntero de la rama y descarta los cambios en el índice y en el directorio de trabajo.

c) Crea un nuevo commit que deshace los cambios anteriores.

d) Reinicia por completo el repositorio eliminando todo el histórico de commits.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

41. En Kubeflow, ¿qué componente está más alineado con MLOps para búsqueda de hiperparámetros y AutoML experimental?:

a) Kubeflow Trainer.

b) Kubeflow ML.

c) Kubeflow Katib.

d) Kubeflow Kserve.

42. ¿Qué sección de un Jenkinsfile declarativo se utiliza para definir las diferentes fases del pipeline?:

a) agent.

b) when.

c) steps.

d) stages.

43. ¿Qué componente de MLflow se utiliza para registrar parámetros, métricas y artefactos durante un experimento?:

a) MLflow Experiments.

b) MLflow Tracking.

c) MLflow Models.

d) MLflow Registry.

44. ¿Cuál de las siguientes opciones NO es un motor de inferencia de LLMs?:

a) Bento.

b) VLLM.

c) SGLang.

d) TGI (Text Generation Inference).

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

45. ¿Cuál de las siguientes afirmaciones sobre CI/CD es CORRECTA?:

a) Continuous Delivery y Continuous Deployment son conceptos equivalentes.

b) La integración continua (Continuous Integration) suele incluir fase de UAT (User Acceptance Testing).

c) En un pipeline de CI/CD, el artefacto generado en la fase de build debe ser inmutable y reutilizarse en todas las etapas posteriores.

d) La entrega continua (Continuous Delivery) permite promover despliegues directos a producción.

46. En el contexto de LLMOps, ¿cuál de los siguientes elementos NO suele considerarse un artefacto que deba versionarse directamente para garantizar la reproducibilidad del comportamiento del sistema?:

a) El modelo utilizado en inferencia.

b) Los prompts.

c) La configuración de inferencia (temperatura, top-k, etc.).

d) El conjunto de datos de referencia (Golden Dataset).

47. ¿Cuál de las siguientes opciones NO está incluida en el Top 10 de OWASP de riesgos asociados al desarrollo e integración de modelos de lenguaje de gran escala (LLMs)?:

a) Prompt Injection.

b) Exposición de datos sensibles (Sensitive Information Disclosure).

c) Configuración de seguridad incorrecta (Security Misconfiguration).

d) Generación de contenido inseguro (Insecure Output Handling).

48. ¿Cuál de las siguientes opciones NO corresponde a un framework de orquestación de contenedores?:

a) Apache GridKube.

b) Docker Swarm.

c) Nomad.

d) Kubernetes.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

49. En un enfoque de responsabilidad compartida en la nube, ¿qué elemento suele seguir siendo responsabilidad directa del cliente en un servicio SaaS?:

a) La seguridad física del centro de datos.

b) La gestión de recursos.

c) La asignación de permisos y control de accesos a usuarios.

d) Los parcheos de seguridad de la solución.

50. ¿Cuál es la característica principal de un DAG en Apache Airflow dentro de la definición de flujos de trabajo?:

a) Permite la ejecución de tareas en ciclos repetitivos sin restricción de dependencias.

b) Define un conjunto de tareas con dependencias dirigidas sin permitir ciclos.

c) Es un sistema de almacenamiento de datos estructurados utilizado por Airflow.

d) Representa únicamente tareas independientes sin relaciones entre ellas.

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

PREGUNTAS DE RESERVA

51. ¿Cuál de las siguientes opciones describe una característica de un despliegue "serverless" en nube?:

a) Ofrece la posibilidad de administrar y gestionar el hardware donde es ejecutado.

b) El escalado debe configurarse manualmente mediante reglas definidas por el usuario.

c) Puede generar latencias en arranques en frío ("cold start").

d) Obliga a mantener instancias activas permanentemente para garantizar disponibilidad.

52. ¿Qué sistema de ficheros utiliza Apache Hadoop?:

a) NTFS.

b) AHFS.

c) YARN.

d) HDFS.

53. ¿Cuál de las siguientes opciones acerca de OpenAPI es CORRECTA?:

a) Su propósito principal es definir la interfaz de usuario de una aplicación web.

b) OpenAPI es una herramienta de desarrollo de APIs REST.

c) Swagger fue el nombre original de la especificación de la actual OpenAPI.

d) OpenAPI es un portal que ofrece diferentes APIs REST para su uso libre.

54. En MLflow, ¿cuál es el propósito principal de los "artefactos" registrados durante un experimento?:

a) Definir la arquitectura de red neuronal utilizada en el modelo.

b) Almacenar resultados intermedios y finales como modelos, gráficos o archivos generados.

c) Empaquetar el código fuente del entrenamiento del modelo.

d) Almacenar el dataset de entrenamiento.

55. Se ejecuta el siguiente comando en un clúster de Kubernetes:

kubectl describe pod nginx-pod

¿Cuál es el propósito principal de este comando y qué salida se espera obtener?:

a) Una lista de todos los pods dentro del namespace nginx-pod.

b) Información detallada del pod nginx-pod.

c) El estado resumido de todos los pods del namespace actual que contengan en su nombre “nginx-pod”.

d) El archivo YAML original con el que se creó el pod.

ESTA PÁGINA SE ENCUENTRA EN BLANCO INTENCIONADAMENTE

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS

ESTA PÁGINA SE ENCUENTRA EN BLANCO INTENCIONADAMENTE

CONTESTE A LAS PREGUNTAS EN LA HOJA DE RESPUESTAS