|  |  |  |  |
| --- | --- | --- | --- |
| ACTIVIDAD DIDÁCTICA CUESTIONARIO | | | |
| Generalidades de la actividad   * Las indicaciones, el mensaje de correcto e incorrecto debe estar la redacción en segunda persona. * Diligenciar solo los espacios en blanco. * El aprendiz recibe una retroalimentación cuando responde de manera correcta o incorrecta cada pregunta. * Señale en la columna Rta. Correcta con una (x) de acuerdo con las opciones presentadas. * Al final de la actividad se muestra una retroalimentación de felicitación si logra el 70 % de respuestas correctas o retroalimentación de mejora si es inferior a este porcentaje.   Para sugerir este tipo de actividad tener presente equipo de Diseño Instruccional, que solo debe haber máximo doce opciones de pregunta y que cada campo tiene un límite de palabras permitidas para garantizar el *responsive web*. | | | |
| Instrucciones para el aprendiz | | *Esta actividad le permitirá determinar el grado de apropiación de los contenidos del componente formativo Visualización, partición y automatización en el procesamiento de datos.*  *Antes de su realización, se recomienda la lectura del componente formativo mencionado. Es opcional (no es calificable), y puede realizarse todas las veces que se desee.*  *Lea la afirmación de cada ítem y luego señale verdadero o falso según corresponda.* | |
| Nombre de la Actividad | | *Modelado de datos automatizados.* | |
| Objetivo de la actividad | | *Validar la apropiación de los conceptos básicos para aplicación de técnicas de preparación**de datos en modelos de inteligencia artificial.* | |
| Texto descriptivo | | *Lea cada enunciado referente a los temas desarrollados en el componente formativo y elija entre verdadero y falso según corresponda.* | |
| PREGUNTAS | | | |
| Pregunta 1 | | **La preparación de datos no tiene impacto en el rendimiento del modelo de aprendizaje automático.** | Rta(s) correcta(s) (x) |
| Opción a) | Verdadero | |  |
| Opción b) | Falso | | X |
| Comentario respuesta correcta | | Excelente trabajo. Una preparación adecuada mejora la precisión del modelo y evita errores como el sobreajuste. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 2 | | **El proceso de codificación convierte variables categóricas en representaciones numéricas.** | |
| Opción a) | Verdadero | | X |
| Opción b) | Falso | |  |
| Comentario respuesta correcta | | Excelente trabajo. Esto permite que los algoritmos de aprendizaje automático trabajen correctamente con datos categóricos. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 3 | | **La normalización de datos busca ajustar todas las variables a una media de cero y desviación estándar de uno.** | |
| Opción a) | Verdadero | | X |
| Opción b) | Falso | |  |
| Comentario respuesta correcta | | Excelente trabajo. Esta técnica es importante para algoritmos sensibles a la escala de los datos, como SVM o regresión logística. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 4 | | **Los valores atípicos siempre deben eliminarse sin excepción.** | |
| Opción a) | Verdadero | |  |
| Opción b) | Falso | | X |
| Comentario respuesta correcta | | Excelente trabajo. Antes de eliminar, deben analizarse ya que pueden representar información importante o errores de medición. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 5 | | **El enriquecimiento de datos consiste en eliminar columnas innecesarias.** | |
| Opción a) | Verdadero | |  |
| Opción b) | Falso | | X |
| Comentario respuesta correcta | | Excelente trabajo. El enriquecimiento crea nuevas características que pueden mejorar el rendimiento del modelo. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 6 | | **El conjunto de validación se utiliza para ajustar hiperparámetros del modelo.** | |
| Opción a) | Verdadero | | X |
| Opción b) | Falso | |  |
| Comentario respuesta correcta | | Excelente trabajo. Permite mejorar el rendimiento sin usar el conjunto de prueba repetidamente. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 7 | | **La técnica *hold-out* implica dividir los datos en múltiples pliegues para entrenamiento y validación.** | |
| Opción a) | Verdadero | |  |
| Opción b) | Falso | | X |
| Comentario respuesta correcta | | Excelente trabajo. *Hold-out* divide una sola vez el conjunto de datos, típicamente en 70 % entrenamiento y 30 % prueba. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 8 | | **La validación cruzada *k-fold* mejora la estimación del rendimiento del modelo.** | |
| Opción a) | Verdadero | | X |
| Opción b) | Falso | |  |
| Comentario respuesta correcta | | Excelente trabajo. Ayuda a reducir el sobreajuste y proporciona una evaluación más robusta. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 9 | | **En un conjunto de datos balanceado, cada clase tiene aproximadamente la misma cantidad de ejemplos.** | |
| Opción a) | Verdadero | | X |
| Opción b) | Falso | |  |
| Comentario respuesta correcta | | Excelente trabajo. Esto facilita el aprendizaje del modelo y evita sesgos hacia una clase dominante. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| Pregunta 10 | | **El F1-score combina precisión y exhaustividad en una sola métrica** | |
| Opción a) | Verdadero | | X |
| Opción b) | Falso | |  |
| Comentario respuesta correcta | | Excelente trabajo. Es especialmente útil cuando existe un desequilibrio en las clases. | |
| Comentario respuesta incorrecta | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Pregunta 11** | | **Los *pipelines* de Scikit-learn permiten automatizar el flujo de transformación y entrenamiento de modelos.** | |
| **Opción a)** | Verdadero | | X |
| **Opción b)** | Falso | |  |
| **Comentario respuesta correcta** | | Excelente trabajo. Facilitan un procesamiento estructurado y reproducible de los datos. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 12** | | **MLflow no permite registrar métricas ni hacer seguimiento del ciclo de vida de un modelo.** | |
| **Opción a)** | Verdadero | |  |
| **Opción b)** | Falso | | X |
| **Comentario respuesta correcta** | | Excelente trabajo. MLflow permite registrar modelos, métricas, parámetros y artefactos de forma automatizada. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 13** | | **La automatización de la gestión de datos favorece la reproducibilidad de los proyectos de IA.** | |
| **Opción a)** | Verdadero | | X |
| **Opción b)** | Falso | |  |
| **Comentario respuesta correcta** | | Excelente trabajo. Al estandarizar los procesos se garantiza mayor eficiencia y trazabilidad. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 14** | | **La tasa de aprendizaje alta siempre mejora el rendimiento del modelo.** | |
| **Opción a)** | Verdadero | |  |
| **Opción b)** | Falso | | X |
| **Comentario respuesta correcta** | | Excelente trabajo. Una tasa de aprendizaje alta puede causar que el modelo no converja adecuadamente. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 15** | | **El RMSE es una métrica útil para evaluar errores en problemas de regresión.** | |
| **Opción a)** | Verdadero | | X |
| **Opción b)** | Falso | |  |
| **Comentario respuesta correcta** | | Excelente trabajo. Es sensible a errores grandes y se usa para medir la precisión de las predicciones. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 16** | | **La matriz de confusión solo sirve para problemas de regresión.** | |
| **Opción a)** | Verdadero | |  |
| **Opción b)** | Falso | | X |
| **Comentario respuesta correcta** | | Excelente trabajo. La matriz de confusión es usada en clasificación para evaluar el desempeño del modelo. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 17** | | **En los *pipelines* de Scikit-learn, el paso final siempre debe ser un estimador.** | |
| **Opción a)** | Verdadero | | X |
| **Opción b)** | Falso | |  |
| **Comentario respuesta correcta** | | Excelente trabajo. El paso final debe ser el que realiza predicción, como una regresión o clasificación. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
|  | |  | |
| **Pregunta 18** | | **El uso de herramientas ETL es irrelevante en la preparación de datos para IA.** | |
| **Opción a)** | Verdadero | |  |
| **Opción b)** | Falso | | X |
| **Comentario respuesta correcta** | | Excelente trabajo. Las herramientas ETL son clave para automatizar la extracción, transformación y carga de datos. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 19** | | **Una estrategia de gobernanza de datos ayuda a garantizar la calidad y seguridad de la información.** | |
| **Opción a)** | Verdadero | | X |
| **Opción b)** | Falso | |  |
| **Comentario respuesta correcta** | | Excelente trabajo. Implementar marcos de gobernanza asegura trazabilidad, calidad y cumplimiento normativo. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |
| **Pregunta 20** | | **La automatización del flujo de datos no aporta beneficios a proyectos de aprendizaje automático.** | |
| **Opción a)** | Verdadero | |  |
| **Opción b)** | Falso | | X |
| **Comentario respuesta correcta** | | Excelente trabajo. Automatizar permite reducir errores, ahorrar tiempo y mejorar la eficiencia en proyectos de ML. | |
| **Comentario respuesta incorrecta** | | Respuesta incorrecta, revise nuevamente el contenido del componente formativo. | |

|  |  |
| --- | --- |
| MENSAJE FINAL ACTIVIDAD | |
| Mensaje cuando supera el 70 % de respuestas correctas | ¡Excelente! Felicitaciones, ha superado la actividad y demuestra sólidos conocimientos sobre el componente formativo. |
| Mensaje cuando el porcentaje de respuestas correctas es inferior al 70 % | No ha superado la actividad. Le recomendamos volver a revisar el componente formativo e intentar nuevamente la actividad didáctica. |

|  |  |  |
| --- | --- | --- |
| **CONTROL DE REVISIÓN** | | |
|  | **Responsable** | **Fecha** |
| **Revisión Experto temático** | **Deivis Eduard Ramirez Martinez** | **Abril de 2025** |
| **Revisión Evaluadora instruccional** | **Viviana Herrera Quiñonez** | **Junio de 2025** |