

1. ¿Con cuál de estas afirmaciones está de acuerdo con respecto a los problemas de datos estructurados frente a los no estructurados?

1 / 1 punto

- ☒ En general, es más fácil para los humanos etiquetar datos y aplicar aumento de datos en datos no estructurados que en datos estructurados.
- ☐ En general, es más fácil para los humanos etiquetar datos y aplicar aumento de datos en datos estructurados que en datos no estructurados.
- ☐ En general, es más fácil para los humanos etiquetar datos en datos estructurados y más fácil aplicar aumento de datos en datos no estructurados.
- ☐ En general, es más fácil para los humanos etiquetar datos en datos no estructurados y más fácil aplicar aumento de datos en datos estructurados.

✓ **Correcto**

¡Así es! Los humanos son más capaces de etiquetar datos no estructurados, como imágenes y clips de audio, que datos estructurados complejos y de alta dimensión. Además, no siempre es posible aplicar el aumento de datos a datos estructurados.

2. Tome el reconocimiento de voz. Algunos etiquetadores transcriben con "... " (como en "Um... el clima de hoy") mientras que otros lo hacen con comas ", ". El rendimiento a nivel humano (HLP) se mide según qué tan bien concuerda un transcriptor con otro. Trabajas con el equipo y logras que todos usen constantemente comas ", ". ¿Qué efecto tendrá esto en HLP?

1 / 1 punto

- ☒ HLP aumentará.
- ☐ HLP disminuirá.
- ☐ HLP permanecerá igual.

✓ **Correcto**

¡Así es! Dado que las etiquetas serán más consistentes, los etiquetadores estarán de acuerdo entre sí con mayor frecuencia, lo que aumentará el HLP.

3. Tome un problema de inspección visual del teléfono. Supongamos que incluso un inspector humano que mira una imagen no puede saber si hay un rasguño. Sin embargo, si el mismo inspector mirara el teléfono directamente (en lugar de una imagen del teléfono), entonces podría decir claramente si hay un rasguño. Su objetivo es construir un sistema que brinde decisiones de inspección precisas para la fábrica (no publicar un artículo). ¿Qué harías?

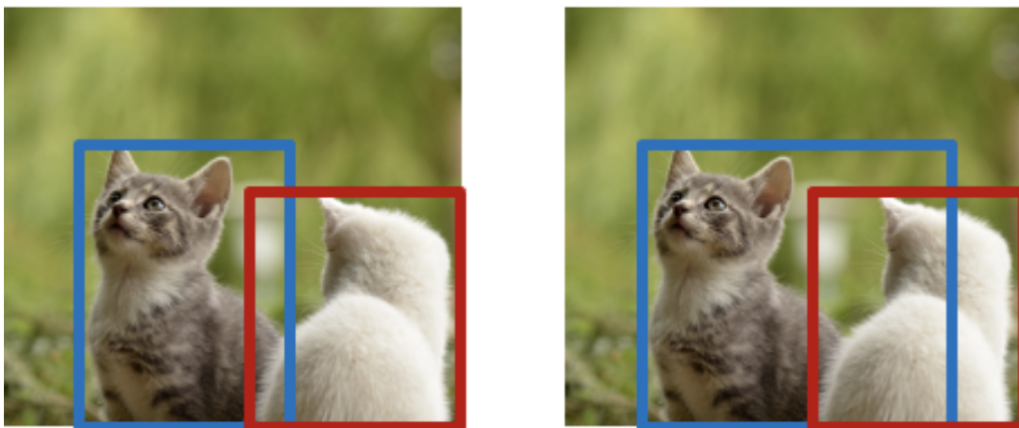
- ☒ Intentar mejorar su sistema de imágenes (cámara/iluminación) para mejorar la calidad o claridad de las imágenes de entrada, x.
- ☐ Trate de mejorar la consistencia de las etiquetas, y.
- ☐ Mida cuidadosamente HLP en este problema (que será bajo) para asegurarse de que el algoritmo pueda coincidir con HLP.
- ☐ Obtenga un gran conjunto de datos de muchos ejemplos de entrenamiento, ya que este es un problema desafiante que requerirá un gran conjunto de datos para hacerlo bien.



Correcto

¡Así es! Si incluso un ser humano que mira la imagen no puede identificar la presencia de un rasguño, deberá mejorar la calidad óptica de su cámara para mejorar el rendimiento de su sistema.

4. Estás construyendo un sistema para detectar gatos. Pide a los etiquetadores que por favor "utilicen cuadros delimitadores para indicar la posición de los gatos". Diferentes etiquetadoras etiquetan de la siguiente manera:



¿Cuál es la causa más probable de esto?

- ☐ Que esto debería haberse planteado como una tarea de segmentación más que de detección.

- ☐ Etiquetadores perezosos.
- ☒ Instrucciones de etiquetado ambiguas.
- ☐ Los etiquetadores no han tenido suficiente café.



Correcto

¡Así es! Sus trabajadores etiquetadores pueden interpretar las instrucciones ambiguas de manera diferente y etiquetar las imágenes de manera diferente. ¡Mejore sus instrucciones y sus datos etiquetados mejorarán!

5. Está construyendo un sistema de inspección visual. HLP se mide según qué *tan bien un inspector está de acuerdo con otro* . El análisis de errores encuentra:

1 / 1 punto

tipo de defecto	Precisión	HLP	% de datos
Rascar	95%	98%	50%
Descoloramiento	90%	90%	50%

Decide que puede valer la pena *verificar la consistencia de la etiqueta* tanto en defectos de rayado como de decoloración. Si tuvieras que elegir uno para empezar, ¿cuál elegirías?

- ☒ Es más prometedor verificar (y potencialmente mejorar) la consistencia de la etiqueta en los defectos de decoloración que en los defectos por rayado. Dado que HLP tiene menos decoloración, es posible que haya instrucciones de etiquetado ambiguas que afecten a HLP.
- ☐ Es más prometedor verificar (y potencialmente mejorar) la consistencia de la etiqueta en los defectos por rayado que en los defectos por decoloración, ya que HLP es mayor en los defectos por rayado y, por lo tanto, es más razonable esperar una alta consistencia.



Correcto

¡Así es! HLP es más bajo para los defectos de decoloración, por lo que existe la oportunidad de mejorar esta métrica mejorando la consistencia de la etiqueta.

6. Para implementar el ciclo de iteración de datos de manera efectiva, la clave es tomarse todo el tiempo necesario para construir primero el conjunto de datos correcto, de modo que todo el desarrollo se pueda realizar en ese conjunto de datos sin necesidad de dedicar tiempo a actualizar los datos.

1 / 1 punto

- ☒ Falso
- ☐ Verdadero

✓ **Correcto**

¡Tocar el asunto exacto! Recopilar y etiquetar datos es un proceso iterativo, ingrese al ciclo de iteración de datos lo más rápido posible.

7. Tiene una canalización de datos para recomendaciones de productos que (i) limpia los datos eliminando las entradas duplicadas y el spam, (ii) hace predicciones. Un equipo de ingeniería mejora el sistema utilizado para el paso (i). Si el modelo entrenado para el paso (ii) sigue siendo el mismo, ¿qué podemos concluir con confianza sobre el rendimiento del sistema en general?

1 / 1 punto

- ☐ Empeorará porque cambiar una etapa anterior en una canalización de datos siempre da como resultado un peor rendimiento de las etapas posteriores.
- ☐ Definitivamente mejorará ya que los datos ahora están más limpios.
- ☐ Empeorará porque la etapa (ii) ahora está experimentando una desviación de datos/conceptos.
- ☒ No es posible decirlo, puede funcionar mejor o peor.

✓ **Correcto**

¡Así es! Es realmente difícil saberlo, ya que depende de cómo se cambiaron los datos y cómo se comporta su modelo.

8. ¿Cuál es el objetivo principal de construir un sistema PoC (prueba de

1 / 1 punto

concepto)?

- ☐ Para seleccionar la arquitectura ML más apropiada para una tarea.
- ☒ Para verificar la viabilidad y ayudar a decidir si una aplicación es viable y vale la pena implementarla.
- ☐ Para recopilar suficientes datos para construir un sistema robusto para la implementación.
- ☐ Construir un sistema de implementación robusto.

☒ **Correcto**

¡Así es! Un sistema de prueba de concepto es una forma sencilla de determinar si vale la pena el tiempo y el esfuerzo para desarrollar el producto.

9. Las herramientas de MLOps pueden almacenar metadatos para realizar un seguimiento de la procedencia y el linaje de los datos. ¿Qué significan los términos procedencia y linaje de los datos?

1 / 1 punto

- ☐ La procedencia de los datos se refiere a la tubería de datos y el linaje de los datos se refiere a la antigüedad de los datos (es decir, qué tan recientemente se recopilaron).
- ☐ La procedencia de los datos se refiere a la secuencia de pasos de procesamiento aplicados a un conjunto de datos, y el linaje de los datos se refiere a la procedencia de los datos.
- ☐ La procedencia de los datos se refiere a la entrada x, y el linaje de los datos se refiere a la salida y.
- ☒ La procedencia de los datos se refiere a la procedencia de los datos y el linaje de los datos, la secuencia de pasos de procesamiento que se le aplican.

☒ **Correcto**

¡Así es!

10. Está trabajando en la inspección visual del teléfono, donde la tarea es usar una imagen de entrada, x, para clasificar los defectos, y. Ha almacenado metadatos para todo su sistema ML, como de qué fábrica proviene cada imagen. ¿Cuáles de los siguientes son usos razonables de los metadatos?

1 / 1 punto

- ☐ Como alternativa a tener que comentar su código.
- ☒ Para sugerir etiquetas o generar información durante el análisis de errores.

✓ **Correcto**
¡Eso es correcto!

- ☒ Seguimiento de la procedencia y el linaje de los datos.

✓ **Correcto**
¡Así es! Los metadatos contendrán información sobre el origen de los datos y qué pasos de procesamiento se les aplicaron. Esto puede ser útil cuando se realiza un análisis de errores.

- ☐ Como otra entrada proporcionada a los etiquetadores humanos (además de la imagen x) para impulsar HLP.