

- Muchos modelos de ML sufren la disminución de las capacidades de predicción con el tiempo. Una solución común utilizada para superar este deterioro es volver a entrenar su modelo con nuevos datos. Como parte de este proceso, puede encontrarse con un fenómeno llamado aparición de conceptos. ¿Cuál de las siguientes afirmaciones describe con precisión este fenómeno emergente?

1 / 1 punto

- ☐ La pérdida de la calidad de la predicción a lo largo del tiempo.
- ☐ La falta de cambio de covariable.
- ☒ La aparición de nuevos patrones en la distribución de datos que antes no estaban presentes en su conjunto de datos.
- ☐ La aparición persistente de datos estacionarios que permanecen inmutables en el tiempo.
- ☒ **Correcto**  
¡Así es! Estos nuevos patrones son los nuevos conceptos que emergen en sus datos. La evaluación y el seguimiento continuos te ayudarán a abordarlos adecuadamente.

2. El control estadístico de procesos es una técnica que detecta la deriva de conceptos asumiendo que los errores siguen una distribución binomial. ¿Activaría el sistema una alarma si  $pag_t = pag_t = 0.3$  y  $pag_{min} = pag_{min} = 0.12$ ?

1 / 1 punto

- ☐ No
- ☒ Sí
- ☒ **Correcto**  
¡Así es! Los valores proporcionados cumplen la regla de alarma:  $pag_t + pag_t \geq pag_{min} + 3p_{min}$

3. En el análisis secuencial, detecta la desviación del concepto mediante el cálculo del valor predictivo negativo, la precisión, la recuperación y la especificidad del sistema en función de una tabla de contingencia estándar. Si los datos son estacionarios, estas cantidades no deberían cambiar con el tiempo. Este análisis es tedioso ya que requiere volver a calcular todas estas métricas cada vez que obtenemos una nueva muestra. ¿Cuál de los siguientes enfoques se suele implementar para superar este problema?

1 / 1 punto

- ☐ Monte Carlo sampling
- ☒ Regla de actualización incremental
- ☐ Computación recursiva y almacenamiento en caché
- ☐ Ventana adaptativa

- ☒ **Correcto**  
¡Así es! La regla de actualización incremental es  $PAG_*^t \leftarrow el_* PAG_*^{t-1} + (1 - el_*) y_{o_{y_t}=\hat{y}_t}$

4. Las técnicas de detección de deriva en entornos no supervisados suelen sufrir la maldición de la dimensionalidad. ¿Cuál de las siguientes técnicas es una solución adecuada para mitigar los efectos de esta maldición? (Marque todo lo que corresponda)

1 / 1 punto

☒ PCA (Análisis de componentes principales)

☒ **Correcto**

¡Así es! El análisis de componentes principales es la herramienta adecuada para reducir la cantidad de características para detectar la deriva de manera más eficiente.

☒ NMF (factorización de matriz no negativa)

☒ **Correcto**

¡Así es! NMF es una técnica de reducción de dimensionalidad muy útil cuando las características están restringidas para que no sean negativas.

☐ SVD (descomposición de valores singulares)

☐ K-medias

5. En entornos no supervisados, la agrupación en clústeres es un método muy útil para detectar novedades en sus datos. En este método, agrupa los lotes de datos entrantes en una de las clases conocidas. Si observa que las características de los nuevos datos están muy alejadas de las características de las clases conocidas, puede llamarlo un concepto emergente. La desventaja de este método es que solo detecta la deriva \_\_\_\_\_ y no los cambios \_\_\_\_\_.

1 / 1 punto

☐ característica, basado en clúster

☐ basado en la población, basado en grupos

☒ basado en conglomerados, basado en la población.

☐ característica basada en clúster

☒ **Correcto**

¡Así es! La deriva se detecta solo en una vista centrada en el clúster,

6. Es una triste verdad que la mayoría de los modelos de aprendizaje automático se entrenan con un conjunto fijo de datos estacionarios. Es muy probable que en este proceso haya sesgado ligeramente su modelo a favor de sus datos limitados en el entrenamiento. En consecuencia, a medida que pasa el tiempo, el rendimiento de su modelo ML se deteriorará con el tiempo. El monitoreo ayuda a prevenir este deterioro del rendimiento ¿de qué manera? (Marque todo lo que corresponda)

1 / 1 punto

☒ Reduce las tasas de falsas alarmas

☒ **Correcto**

¡Correcto! Si se aplica sabiamente, el monitoreo puede detectar la desviación de datos desde el principio y ajustar el modelo en consecuencia para adaptarse a estos cambios y, por lo tanto, mejorar el rendimiento del modelo.

☒ Le permite identificar cambios de distribución cerca de los límites de clasificación

☒ **Correcto**

¡Sí! La deriva de datos puede cambiar los límites de clasificación de manera bastante drástica y el monitoreo lo ayudará a detectar y mitigar este comportamiento no deseado.

☐ Reentrenando a tu modelo constantemente

☒ Identificar regiones en el espacio latente donde el modelo funciona mal

☒ **Correcto**

¡Así es! El monitoreo lo ayudará a identificar áreas en el espacio latente donde su modelo tiene dificultades para la clasificación. Puede seguir utilizando este conocimiento para refinar su modelo.

☐ Al realizar la reducción de dimensionalidad

☐ Le permite establecer etiquetas de verdad del terreno