|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tipo de actividad | Quiz/Cuestionario | | | |
| Título de la actividad (como se verá en plataforma)  *Aquí se pone el título que el estudiante verá en la plataforma (vista lateral y vista por semanas)* | Mixturas gaussianas | | | |
| Objetivo(s) de la semana a la que apunta esta actividad  *Estos deben coincidir con la información de la tabla que está más adelante, en la que se verifican preguntas frente a objetivos.* |  | | | |
| ¿Es sumativa?  *Seleccionar con X* | Sí [ ] | No | | |
| Número de intentos máximos (Si es formativo anote NA) |  | | | |
| Tiempo de la actividad  *Horas:Minutos* |  | | | |
| **Tipo de comentario**  *Se refiere a la retroalimentación de la pregunta y de la actividad en general. (seleccionar una opción con [X])* | Completo [ ] | | Parcial [ ] | Limitado [ ] |
| *Completo: el estudiante podrá ver la puntuación total y el porcentaje de preguntas correctas, puntaje por pregunta y estado de cada una (correcta/incorrecta), retroalimentación de las preguntas y la/s rúbricas utilizadas para la calificación.* | | *Parcial: El estudiante dejará de ver la retroalimentación de las preguntas.* | *Limitado: el estudiante únicamente podrá ver la puntuación total y el porcentaje de preguntas correctas*. |
| **Plantilla de cuestionario**  **Opciones múltiples**   1. **Múltiples respuestas correctas** 2. **Única respuesta correcta**   **Respuesta libre**   1. **Coincidencia de texto** 2. **Expresión matemática** 3. **Numérico** 4. **Expresión regular** 5. **Expresión de código**   **Reflexión**   1. **Opción múltiple de reflexión** 2. **Opción única de reflexión** 3. **Respuesta de texto de reflexión**   **Personalizado**   1. **Pregunta de complemento** | *Más adelante en este documento encuentra el formato que debe seguir para preguntas de opción múltiple con única o múltiple respuesta. Si tenemos otras preguntas, podemos revisar para indicar el formato en que debe presentarse.*  *Las preguntas que se llaman de Reflexión dan el punto al estudiante independiente de su respuesta.*  *Aspectos importantes para tener presentes en la construcción de preguntas y opciones de respuesta:*  *1. Cada opción de respuesta debe tener retroalimentación.*  *2. Todas las opciones de respuesta deberían ser lógicas y acordes con el tema estudiado, la idea con estas es propiciar cierta discriminación que permita verificar las comprensiones de los estudiantes.*  *3. La retroalimentación no debe ofrecer la respuesta correcta, sino centrarse en la razón de la equivocación. Esto es más sencillo de lograr cuando se tiene en cuenta el punto anterior. Es difícil crear una retroalimentación formativa para una opción “descabellada” o “ilógica”.*  *4. No se sugiere opciones como ninguna de las anteriores o todas las anteriores.*  *5. Por defecto, está configurado para que las opciones de respuesta se muestren en diferente orden cada vez. Si se necesita mostrar en algún orden específico debemos indicarlo.*  *6. Por defecto, todas las preguntas se ponderan con el mismo valor. Si se quiere algo diferente, debemos indicarlo.*  *7. Por accesibilidad, cuidar que si una pregunta tiene una imagen, el texto explique lo necesario para que quien no accede a la imagen pueda responder la pregunta.* | | | |

**Preguntas con el formato de la plantilla**

**Question 1**

Imagen en blanco y negro

Descripción generada automáticamente con confianza bajaA partir de los datos mostrados a continuación ¿Cómo serían los centroides de los clústeres resultantes del algoritmo de K-Medias vs Mixturas Gaussianas si K = 2?

A: Para este caso particular, los centroides de ambos segmentos serían los mismos pues ambos algoritmos identificarían los mismos clústeres.

Feedback: Los centroides corresponden a un promedio de todos los puntos. Como Mixturas Gaussianas es una técnica de soft-clustering, todas las observaciones pertenecen a todos los clústeres, por ende, el centroide de GMM del cluster de la izquierda estarían ligeramente corrido hacia la derecha con respecto al centroide de K-Medias y el centroide del cluster de la derecha de GMM estaría ligeramente más a la izquierda que su contraparte para K-Medias.

\*B: Para este caso particular, los centroides de ambos segmentos serían diferentes, pero ambos algoritmos identificarían los mismos clústeres.

Feedback: Los centroides corresponden a un promedio de todos los puntos. Como Mixturas Gaussianas es una técnica de soft-clustering, todas las observaciones pertenecen a todos los clústeres, por ende, el centroide de GMM del cluster de la izquierda estarían ligeramente corrido hacia la derecha con respecto al centroide de K-Medias y el centroide del cluster de la derecha de GMM estaría ligeramente más a la izquierda que su contraparte para K-Medias.

C: En general, como son dos algoritmos diferentes, se encontrarán soluciones diferentes. Los centroides y los clústeres no van a ser similares.

Feedback: En este caso, ambos algoritmos arrojarían agrupaciones muy similares.

D: Para K menor o igual a 2, ambos algoritmos siempre arrojan la misma solución.

Feedback: Que la solución sea similar o diferente depende de la estructura de los datos, no de la técnica per se.

**Question 2**

¿Cuál es el número de parámetros necesarios para especificar un modelo de Mixturas Gaussianas con 4 clústeres, 5 variables en los datos y covarianzas diagonales?

\*A: 44

Feedback: Para describir cada clúster se necesita 1 parámetros para la probabilidad, 5 parámetros para las medias y 5 parámetros para la matriz de covarianzas ya que es una matriz diagonal. Como tenemos 4 segmentos, entonces serían necesarios 11 x 4 = 44 parámetros.

B: 84

Feedback: Esto sería cierto si la matriz de covarianzas estuviera completa y no fuera diagonal.

C: 11

Feedback: Esta es la cantidad de parámetros por componente.

D: 3

Feedback: Se deben especificar parámetros de probabilidad, de centralidad y dispersión para cada segmento. Revise sus cuentas.

**Question 3**

¿Cuál es el número de parámetros necesarios para especificar un modelo de Mixturas Gaussianas con 4 clústeres, 5 variables en los datos y covarianzas completas?

A: 44

Feedback: Esto sería cierto si la matriz de covarianzas fuera diagonal.

\*B: 84

Feedback: Para describir cada clúster se necesita 1 parámetros para la probabilidad, 5 parámetros para las medias y 15 parámetros para la matriz de covarianzas. Como tenemos 4 segmentos, entonces serían necesarios 21 x 4 = 84 parámetros.

C: 21

Feedback: Esta es la cantidad de parámetros por componente.

D: 3

Feedback: Se deben especificar parámetros de probabilidad, de centralidad y dispersión para cada segmento. Revise sus cuentas.

**Question 4**

Imagen que contiene Diagrama

Descripción generada automáticamenteConsidere aplicar EM para entrenar un modelo de mezcla gaussiana (GMM) para agrupar los datos a continuación en dos grupos. Los puntos '+' indican los centroides actuales de dos componentes de la mezcla gaussiana después de la k-ésima iteración de EM. ¿En qué dirección se van a mover los centroides?

A: y a la izquierda.

Feedback: Como se quiere encontrar dos segmentos y los datos están divididos en dos extremos, cada centroide se moverá para el extremo que tiene más cerca.

\*B: a la izquierda y a la derecha.

Feedback: Como se quiere encontrar dos segmentos y los datos están divididos en dos extremos, cada centroide se moverá para el extremo que tiene más cerca.

C: a la derecha y a la izquierda.

Feedback: Como se quiere encontrar dos segmentos y los datos están divididos en dos extremos, cada centroide se moverá para el extremo que tiene más cerca.

D: y a la derecha.

Feedback: Como se quiere encontrar dos segmentos y los datos están divididos en dos extremos, cada centroide se moverá para el extremo que tiene más cerca.

**Question 5**

¿Qué sucede con la log verosimilitud marginal de los datos en cada iteración del EM?

\*A: Aumenta

Feedback: Cada iteración del algoritmo EM aumenta la probabilidad de los datos, a menos que esté exactamente en un óptimo local.

B: Decrece

Feedback: Cada iteración del algoritmo EM aumenta la probabilidad de los datos, a menos que esté exactamente en un óptimo local.

C: Se buscan los parámetros que la mantengan constante.

Feedback: Cada iteración del algoritmo EM aumenta la probabilidad de los datos, a menos que esté exactamente en un óptimo local.

D: Puede aumentar o disminuir según la topografía del problema.

Feedback: Cada iteración del algoritmo EM aumenta la probabilidad de los datos, a menos que esté exactamente en un óptimo local.