|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 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)* | Generalidades de Componentes principales | | | |
| 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**

¿Qué tipo de algoritmo es Componentes Principales?

A: Es un algoritmo de segmentación para crear grupos homogéneos en su interior y heterogéneos contra los demás segmentos.

Feedback: Componentes Principales es un algoritmo que busca reducir la dimensionalidad de los datos, preservando la máxima varianza posible.

\*B: Es un algoritmo de reducción de dimensionalidad.

Feedback: ¡Correcto! Componentes Principales permite reducir la dimensionalidad de los datos.

C: Es un algoritmo de aprendizaje supervisado donde se busca maximizar la varianza de los datos.

Feedback: Componentes Principales es un algoritmo que busca reducir la dimensionalidad de los datos, preservando la máxima varianza posible.

D: Es una técnica de limpieza de datos para generar modelos predictivos.

Feedback: Componentes Principales es un algoritmo que busca reducir la dimensionalidad de los datos, preservando la máxima varianza posible.

**Question 2**

¿Bajo qué contexto utilizamos Componentes Principales?

A: Cuando queremos reducir la dimensionalidad de nuestros datos para hacer inferencia causal.

Feedback: El resultado de Componentes Principales puede ser utilizado en otros modelos, pero no es el objetivo del algoritmo.

\*B: Utilizamos Componentes Principales como una herramienta descriptiva de nuestros datos.

Feedback: ¡Correcto! El resultado es interpretable y nos sirve para entender mejor nuestros datos y también para usar el resultado del algoritmo como input de algún modelo.

C: Cuando queremos predecir el resultado de una variable .

Feedback: El resultado de Componentes Principales puede ser utilizado en otros modelos, pero no es el objetivo del algoritmo.

D: Cuando queremos aumentar la varianza de nuestros datos.

Feedback: Este algoritmo necesariamente va a reducir la varianza de los datos.

**Question 3**

¿Qué preprocesamiento se le debe hacer a los datos antes de realizar Componentes principales?

A: Ninguno, se pueden ingresar los datos crudos.

Feedback: La matriz de varianza y covarianza de la data es sensible a la escala de las variables por lo que los resultados del algoritmo se sesgan hacia las variables con mayor magnitud.

\*B: Es necesario estandarizar los datos.

Feedback: ¡Correcto! El algoritmo es sensible a la escala de los datos, por eso es importante siempre estandarizar la data antes de implementar este algoritmo.

C: Es necesario eliminar las variables redundantes de la base, aquellas que se correlacionen fuertemente.

Feedback: El algoritmo busca reducir la dimensionalidad de los datos, por ende este elimina la variación redundante.

**Question 4**

¿Es correcto usar el algoritmo de Componentes Principales para reducir la dimensionalidad de una base de datos con variables de tipo numéricas y categóricas?

A: Sí, siempre y cuando se vuelva la variable categórica en un una variable numérica ordenada del estilo: 1, 2, 3, etc.

Feedback: Este preprocesamiento es inadecuado pues supone una distancia entre las categorías y un orden que no necesariamente corresponde a la realidad de los datos. Por ejemplo, no es correcto suponer que una variable categórica como la ciudad de nacimiento tenga un orden.

B: No. No se puede calcular la correlación de Pearson entre una variable categórica y otra numérica.

Feedback: Es cierto que no es correcto calcular la correlación de Pearson entre dos variables categóricas, pero si se puede calcular la covarianza entre dos variables dicótomas.

\*C: Sí, siempre y cuando se partan las variables categóricas en variables dicótomas (one-hot encoder).

Feedback: ¡Correcto! Tener variables dicótomas permite calcular la matriz var-cov de la que parte el algoritmo. No obstante, cuando se tienen conjuntos de datos heterogéneos (es decir con variables de diferentes clases) no es tan recomendable usar PCA pues esta aproximación puede no representar de manera idónea la distancia entre variables. Se recomienda buscar una metodología más idónea para estimar adecuadamente la matriz var-cov.

**Question 5**

¿Están los eigenvalores negativos asociados con una varianza explicada más baja?

A: Sí. Entre más grande sean los eigenvalores, mayor será la varianza explicada. Por eso se deben ordenar los eigenvalores de mayor a menor

Feedback: Se deben ordenar los eigenvalores de mayor a menor en valor absoluto.

\*B: No, depende de la magnitud de los demás eigenvalores.

Feedback: Correcto. Los eigenvalores que mayor porción de la varianza explican son aquellos que sean mayores en valor absoluto relativos a los demás.

 C: La varianza explicada no depende de los eigenvalores sino de la norma de los eigenvectores.

Feedback: Los eigenvalores que mayor porción de la varianza explican son aquellos que sean mayores en valor absoluto relativos a los demás.