|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 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**

¿Permite el algoritmo de Componentes Principales capturar relaciones no lineales entre las variables?

A: Verdadero

Feedback: Componentes Principales hace una combinación lineal de los vectores que conforman la data, por tal motivo, relaciones más complejas entre las variables no se capturan en este método.

\*B: Falso

Feedback: ¡Correcto! Los componentes principales son combinaciones lineales e independientes de los vectores que conforman la data.

**Question 2**

¿Cuántos componentes principales tendrá una matriz X de tamaño ? Suponga que nuestros datos son linealmente independientes y .

A: Tendrá Componentes Principales.

Feedback: Pista: Se tendrán tantos Componentes Principales como eigenvalores.

\*B: Tendrá Componentes Principales.

Feedback: ¡Correcto! Sabemos que tendremos Componentes principales porque la matriz var-cov es simétrica y de tamaño .

C: El número de Componentes Principales varía según la estructura de los datos.

Feedback: Dado que nuestros datos son linealmente independientes, podemos saber con exactitud cuántos Componentes Principales obtendremos.

D: Tendrá el mínimo entre y .

Feedback: El número de observaciones no influye en la cantidad de Componentes Principales.

**Question 3**

¿Es el algoritmo de Componentes Principales es invariante a la escala de las variables?

A: Verdadero

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: Falso

Feedback: ¡Correcto! Por eso es importante siempre estandarizar la data antes de implementar este algoritmo.

**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: Verdadero. 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: Falso, 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.

**Question 6**

De acuerdo al gráfico a continuación, ¿Cuántos componente principales debería utilizar?



A: 1. El primer componente solo explica 50% de la varianza, recuerde el método del codo visto en clase

\*B: 3: Correcto, los 3 primeros componentes explican 90% de la varianza y hay un codo luego del tercer componente

C: 4: El 4to componente explica menos del 10% de la varianza, recuerde el método del codo visto en clase

C: 5: El 5to componente explica menos del 2% de la varianza, recuerde el método del codo visto en clase