
Solucionario de

ejercicios de autoevaluación

Ejercicios de autoevaluación

Unidad de Aprendizaje 1

1. Los datos son conocidos como:

- a. **El petróleo del siglo XXI.**
- b. El oro del siglo XXI.
- c. El Dorado del siglo XXI.
- d. El grial del siglo XXI.

2. ¿Cuál de estos factores no es un “factor decisivo” que confirma la importancia de los datos en los negocios?

- a. Valorar el desempeño.
- b. Mejora de procesos.
- c. Monitorizar sistemas esenciales.
- d. **Obtener un enfoque de eficiencia.**

3. Para administrar de forma eficaz los datos disponibles es importante:

- a. Centrarse en la tecnología.
- b. Clasificar la información, sabiendo cuál es su valor absoluto.
- c. **Enfocarse en la información.**
- d. Ser veloces gracias a la tecnología más novedosa.

4. Dentro de las 5 fases para convertirse en *Data Driven*, la fase *Data Savvy* es:

- a. Aquella donde la empresa entiende los datos y es capaz de extraerles valor e información.
- b. Aquella donde la empresa sabe de la existencia de los datos, pero no qué hacer con ellos.
- c. Aquella donde la empresa sabe de la existencia de los datos, pero únicamente los recopila sin analizarlos.
- d. **Aquella donde la empresa es concedora del poder de los datos y comienza a desarrollar herramientas internas.**

5. Entre los modelos de negocio existentes, en el Modelo A "Recolector de datos y agregador de valor" se realiza la tarea de...

- a. ... recoger datos libres y gratuitos de internet.
- b. ... recoger los datos ofrecidos por los clientes.
- c. ... generar los datos por su cuenta.
- d. ... agregar los datos para crear valor.

6. Determina si la siguiente afirmación es verdadera o falsa: "El *Machine Learning* usa redes neuronales como base de su funcionamiento".

- Verdadero
- Falso

7. En *Data Model Business Canvas*, ¿cuántos bloques forman su plantilla para realizar el modelo?

- a. 7
- b. 9**
- c. 11
- d. 10

8. En *Big Data* se habla de enormes cantidades de datos pero, ¿qué cantidad se estima como mínima para que se pueda considerar que estamos trabajando con datos masivos?

- a. 200-400 GB.
- b. 80-100 TB.
- c. 10-15 PB.
- d. 30-50 TB.**

9. Determina si la siguiente afirmación es verdadera o falsa: "Los datos semiestructurados disponen de un esquema fijo".

- Verdadero
- Falso

10. ¿Cuál de las siguientes palabras no pertenecen a las V del Big Data?

- a. Viabilidad
- b. Versatilidad**
- c. Veracidad
- d. Variedad

Ejercicios de autoevaluación

Unidad de Aprendizaje 2

1. ¿Cuál de estos profesionales es el encargado de realizar la administración de los sistemas de almacenamiento distribuido?

- a. **Data analyst.**
- b. *Data engineer.*
- c. *Data architect.*
- d. *QA engineer.*

2. ¿Qué herramienta *Big Data* trabaja directamente en memoria y no en disco?

- a. *Yarn*
- b. *HBase*
- c. *Mahout*
- d. **Spark**

3. El algoritmo *Random Forest* es un algoritmo que pertenece al tipo:

- a. **Análisis supervisado.**
- b. Análisis no supervisado.
- c. Clustering.
- d. NLP.

4. ¿Cómo se llama el nivel más básico de representación de una red neuronal?

- a. Monocapa
- b. **Perceptrón**
- c. Base radial
- d. No recurrente

5. En Python, existe un tipo de dato compuesto dinámico y actualizable, cuyos datos están separados por comas, al que se denomina:

- a. Tuplas
- b. Diccionario

- c. Listas
- d. Variable

6. Determina si la siguiente afirmación es verdadera o falsa: “La librería *Gensin* de Python se utiliza para trabajar con bases de datos”.

- Verdadero
- Falso

7. No es una característica de los procesos ETL:

- a. Necesitar tiempo para cargar los datos.
- b. Estar basado únicamente en la nube.**
- c. Utilizar datos estructurados.
- d. Emplear tecnologías distintas tanto en la fuente de datos como en el destino de estos.

8. ¿Qué tipo de gráfico encuentra su fase en los cuartiles?

- a. *Boxplot***
- b. *Scatter*
- c. *Histograma*
- d. *Matrix plot*

9. Determina si la siguiente afirmación es verdadera o falsa: “En el modelo estrella es necesario realizar *joins* para obtener la información”.

- Verdadero
- Falso

10. Los valores del coeficiente de correlación de *Pearson* pueden variar entre:

- a. -1 y 1**
- b. 0 y 1
- c. -1 y 0
- d. -0,5 y 0,5

Ejercicios de autoevaluación

Unidad de Aprendizaje 3

1. ¿Qué necesita un modelo de *Machine Learning* para obtener una salida?

- a. Datos de entrada.
- b. Un algoritmo.**
- c. Procesar los datos.
- d. Todas las opciones son correctas.

2. Indica cuál de las siguientes suposiciones de la regresión lineal es incorrecta.

- a. Debe existir correlación entre las variables independientes.**
- b. Relación lineal y aditiva.
- c. Términos de error con varianza constante.
- d. Variable dependiente con distribución normal.

3. ¿Cuál de estos algoritmos es considerado de clasificación?

- a. KNN**
- b. PCA
- c. SVD
- d. CNN

4. Identifica cuál de estas características se atribuye a *lazy evaluation*:

- a. Es la más común y habitual en los lenguajes de programación.
- b. Necesita poco tiempo a la hora de ejecutarse.
- c. El algoritmo SVM es uno de sus ejemplos más destacados.
- d. Los valores no se calculan inmediatamente cuando son asignados.**

5. Entre los siguientes algoritmos, ¿cuál no se considera de agrupamiento?

- a. K-Means.
- b. DBSCAN.

- c. **K vecinos más cercanos.**
- d. Agrupamiento jerárquico.

6. Determina si la siguiente afirmación es verdadera o falsa: "En el algoritmo *Random Forest* el problema de sobreajuste es muy difícil de solucionar, considerándose su mayor problema".

- Verdadero
- Falso

7. El cálculo de un residuo es el siguiente:

- a. **Valor observado y - valor predicho y.**
- b. Valor predicho y - valor observado y.
- c. Valor observado y - valor observado x.
- d. Valor predicho y - valor observado x.

8. Respecto a los modelos supervisados, determina a qué métrica de evaluación corresponde la siguiente definición: "Su objetivo mide con qué frecuencia es correcto el clasificador".

- a. Exhaustividad
- b. Precisión
- c. **Exactitud**
- d. Especificidad

9. Determina si la siguiente afirmación es verdadera o falsa: "El algoritmo PCA se utiliza cuando la cantidad de datos disponible es pequeña".

- Verdadero
- Falso

10. ¿Cuál de estos tipos no se considera un *model ensembles*?

- a. *Stacking*
- b. *Bagging*
- c. *Boosting*
- d. ***Tracking***

Ejercicios de autoevaluación

Unidad de Aprendizaje 4

1. ¿Cuál de estas ventajas no corresponde a una red neuronal?

- a. Tolerancia a fallos.
- b. Implementación en problemas sencillos.**
- c. Operación en tiempo real.
- d. Autoorganización.

2. ¿Cuál de estas funciones es más popular en su uso en redes neuronales como función de activación?

- a. Función sigmoide.**
- b. Función logarítmica.
- c. Función exponencial.
- d. Función integral.

3. Una red neuronal profunda se caracteriza por:

- a. Tener varias capas ocultas.**
- b. Tener varias neuronas.
- c. Tener diferentes funciones de activación.
- d. Tener una gran cantidad de entradas.

4. La función de costo o pérdida se define como:

- a. La diferencia entre la salida generada y la entrada.
- b. La diferencia entre la salida generada y la real.**
- c. La diferencia entre la salida real y la salida de la última capa oculta.
- d. La diferencia entre la salida real y la entrada.

5. ¿Qué tipo de redes neuronales son más eficaces en la visión por computador?

- a. Redes neuronales recurrentes.
- b. Redes neuronales generativas adversarias.
- c. Redes neuronales convolucionales.**
- d. Redes neuronales autoasociativas.

6. Determina si la siguiente afirmación es verdadera o falsa: “Los Transformers son modelos de *Deep Learning* usados en procesamiento de lenguaje natural, los cuales necesitan para su buen funcionamiento que los datos se encuentren ordenados”.

- Verdadero
- Falso

7. ¿Qué tipo de redes neuronales son más eficaces en el análisis de series temporales?

- a. **Redes neuronales recurrentes.**
- b. Redes neuronales convolucionales.
- c. Redes neuronales heteroasociativas.
- d. Redes neuronales generativas adversarias.

8. El perceptrón simple se caracteriza por:

- a. Tener una capa oculta.
- b. Ser óptimo en problemas de regresión.
- c. **Ser un clasificador binario.**
- d. Utilizar una función exponencial como función de activación.

9. Determina si la siguiente afirmación es verdadera o falsa: “En una red *Hopfield* los pesos pueden ser calculados previamente y permanecer fijos durante el proceso de aprendizaje”.

- Verdadero
- Falso

10. En el proceso *Dropout* se desactiva un porcentaje de neuronas que se sitúan en la capa:

- a. **Ocultas.**
- b. Entrada.
- c. Salida.
- d. Primera capa oculta.

Ejercicios de autoevaluación

Unidad de Aprendizaje 5

1. ¿Cuál de estas características no pertenece a *Apache Spark*?

- a. Es un marco de computación distribuida.
- b. Permite la escalabilidad.
- c. Es tolerante a fallos.
- d. Es un *framework* costoso.**

2. Uno de estos elementos no forma parte de *Hadoop*. Indica cuál es:

- a. *MapReduce*
- b. RDD**
- c. *Yarn*
- d. HDFS

3. Los bloques en el HDFS son configurables, pero el tamaño por defecto es:

- a. 128 MB**
- b. 64 MB
- c. 256 MB
- d. 66 MB

4. *Spark* puede aumentar la velocidad de ejecución gracias a que...

- a. ... trabaja en HDFS.
- b. ... trabaja en memoria.**
- c. ... trabaja en clúster.
- d. ... trabaja en la nube.

5. ¿Cuál de estos lenguajes de programación no soporta *Spark*?

- a. Scala
- b. Visual Basic**
- c. R
- d. Clojure

6. Determina si la siguiente oración es verdadera o falsa: “El significado de los conjuntos de datos distribuidos que usa *Spark* RDD es *Resilient Distributed Dataframe*”.

- Verdadero
- Falso

7. ¿Qué modo de instalación de *Spark* dispone de un gestor de recursos?

- a. *Modo Standalone*.
- b. *Hadoop V1*.
- c. ***Hadoop V2***.
- d. Las opciones a y b son correctas.

8. ¿Cómo se llama la plataforma analítica de datos basada en *Apache Spark* y desarrollada por ella misma?

- a. ***Databricks***.
- b. *Dataset*.
- c. *DataSpark*.
- d. *Spark Services*.

9. Determina si la siguiente oración es verdadera o falsa: “*Amazon EMR* contiene por defecto la integración de la herramienta *Power BI* para la visualización de datos”.

- Verdadero
- Falso

10. ¿Cuál de estas plataformas no contiene *Apache Spark* en la nube?

- a. Amazon
- b. Google
- c. **Yahoo!**
- d. Microsoft

Ejercicios de autoevaluación

Unidad de Aprendizaje 6

1. Indica cuál de estos tipos no es un *storytelling* usado en publicidad:

- a. *In media res*.
- b. Minigráfico.
- c. Bucles animados.
- d. **Dinámico.**

2. ¿Cuál de estas funciones no puede llevar a cabo un *dashboard*?

- a. Mostrar visualmente KPI.
- b. Medir métricas.
- c. **Trasformar datos.**
- d. Seguimiento de cómo se encuentra la empresa.

3. Uno de estos grupos no existe en la catalogación de las herramientas BI, ¿de cuál se trata?

- a. ***Business Intelligence Open Data.***
- b. *Business Intelligence Cloud.*
- c. *Business Intelligence colaborativo.*
- d. *Business Intelligence integrado.*

4. ¿Qué biblioteca necesita JavaScript y HTML para elaborar visualizaciones interactivas de datos?

- a. *Plotly*
- b. ***Bokeh***
- c. *Nivo.rocks*
- d. *Scikit-learn*

5. ¿Cuál de estas herramientas ofrece servicios *self-services analytics* en *Business Intelligence*?

- a. *QlikSense*
- b. *Microstrategy*
- c. ***Pyramid***
- d. *Tableau*

6. Determina si la siguiente afirmación es verdadera o falsa: “Un ERP es una herramienta útil como fuente de datos en sistemas BI que ayuda y ofrece datos sobre los clientes de esa compañía”.

- Verdadero
- Falso

7. ¿Qué componente es el más adecuado en *Power BI* para una empresa que trabaja con grandes cantidades de datos y necesita un refresco constante de estos?

- a. *Power BI Desktop*.
- b. *Power BI Embedded*.
- c. *Power BI Mobile*.
- d. ***Power BI Pro***.

8. ¿Cuál de estos tipos de archivos no pueden ser importados como fuente de datos en *Power BI*?

- a. XML
- b. **JPG**
- c. JSON
- d. PDF

9. Determina si la siguiente afirmación es verdadera o falsa: “*Power BI* dispone de la herramienta *Power View* para realizar una transformación previa a los datos que se van a utilizar”.

- Verdadero
- Falso

10. ¿Cómo se llama la tecnología sobre la que se basa *Tableau*?

- a. *DAX*
- b. ***VizQL***
- c. *TableauSQL*
- d. *Mesos*

Ejercicios de autoevaluación

Unidad de Aprendizaje 7

1. El acrónimo VUCA significa (en español):

- a. Volatilidad, unión, complejidad, ambigüedad.
- b. Variedad, unión, complejidad, ambigüedad.
- c. Volatilidad, incertidumbre, complejidad, ambigüedad.**
- d. Variedad, incertidumbre, conexión, ambigüedad.

2. ¿Cuál de estos pasos clave no se corresponde con la gestión del cambio?

- a. Crear sentido de urgencia.
- b. Asegurar triunfos a corto plazo.
- c. Formar coalición.
- d. Integración.**

3. ¿Qué tipo de *storytelling* empresarial es el más parecido a la publicidad clásica?

- a. *Storytelling* de consumo.**
- b. *Storytelling* de relación.
- c. *Storytelling* personal.
- d. *Storytelling* marketing.

4. Un plan de difusión de contenidos está compuesto por cuatro fases importantes, que son:

- a. Planificación, Formato, Canales y Medición.**
- b. Planteamiento, Formato, Difusión y Valoración.
- c. Planificación, Canales, Integración y Medición.
- d. Redacción, Medición, Formato y Valoración.

5. ¿Mediante qué herramienta se realiza la gestión de proyectos?

- a. Zoom.
- b. Slack.
- c. Wrike.**
- d. Workplace Facebook.

6. Determina si la siguiente afirmación es verdadera o falsa: "El trabajo en remoto se lleva a cabo fuera de la oficina, pero respetando siempre la jornada laboral".

- Verdadero
- Falso

7. ¿Cuál de estos tipos no es considerado una forma de trabajo en remoto?

- a. Empleado totalmente en remoto.
- b. Empleo flexible.
- c. **Semipresencial.**
- d. Autónomo.

8. ¿Qué herramienta se utiliza para automatizar flujos de trabajo?

- a. *Toggl Plan.*
- b. **Zapier.**
- c. *TeamWork.*
- d. *Mindmaster.*

9. Determina si la siguiente afirmación es verdadera o falsa: "La técnica *pomodoro* consiste en trabajar por tiempos de 50 min de forma continua y posteriormente descansar 5 min".

- Verdadero
- Falso

10. *Teamwork*, *Freedcam* y *nTask* son herramientas utilizadas en:

- a. Automación de flujos de trabajo.
- b. **Gestión de proyectos.**
- c. Generación de contenido.
- d. Inserción de datos.

